

**ENGINEERING  
SUPERCONDUCTIVITY**

# ENGINEERING SUPERCONDUCTIVITY

---

**Peter J. Lee, *Editor***

Applied Superconductivity Center  
University of Wisconsin-Madison



**WILEY-INTERSCIENCE**

**John Wiley & Sons, Inc.**

New York/Chichester/Weinheim/Brisbane/Singapore/Toronto

This book is printed on acid-free paper. ∞

Copyright © 2001 by John Wiley & Sons. All rights reserved.

Published simultaneously in Canada.

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except as permitted under Sections 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 750-4744. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 605 Third Avenue, New York, NY 10158-0012, (212) 850-6011, fax (212) 850-6008, E-Mail: PERMREQ@WILEY.COM.

For ordering and customer service, call 1-800-CALL-WILEY.

***Library of Congress Cataloging-in-Publication Data:***

Engineering superconductivity / Peter J. Lee, editor.

p. cm.

“A Wiley-Interscience publication.”

Includes index.

ISBN 0-471-41116-7 (cloth : alk. paper)

1. High temperature superconductors—Industrial applications.
2. Superconductors—Industrial applications. I. Lee, Peter J.

TK7872.S8 E54 2001

621.3'5—dc21

00-068603

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

# CONTENTS

---

|   |            |  |            |
|---|------------|--|------------|
| <b>PREFACE</b>                                    | vii        | <b>APPLICATIONS AND RELATED TECHNOLOGY</b>     | <b>329</b> |
| <b>SUPERCONDUCTIVITY AND MAGNETISM</b>            | <b>1</b>   | <b>Introduction</b>                            | <b>329</b> |
| Historical Introduction                           | 1          | Capacitor Storage                              | 330        |
| Magnetic Flux                                     | 5          | Diagnostic Imaging                             | 345        |
| Magnetic Levitation                               | 22         | Frequency Converters and Mixers                | 363        |
| Magnetic Materials                                | 30         | Fusion Reactors                                | 375        |
| Magnetic Noise, Barkhausen Effect                 | 41         | Gyrotron                                       | 394        |
| Magnetic Refrigeration                            | 53         | Hall Effect Transducers                        | 401        |
| Magnetic Resonance                                | 58         | High-Energy Physics Particle Detector          |            |
| Superconducting Magnets, Quench Protection        | 64         | Magnets  | 412        |
| Superconductivity: Critical Current               | 91         | Infrared Detector Arrays, Uncooled             | 424        |
| Superconductivity: Electromagnets                 | 107        | Magnetic Sensors                               | 434        |
| Superconductivity: Electronics                    | 116        | Magnetic Separation                            | 448        |
| Superconductivity: Hysteresis and Coupling Losses | 138        | Magnetic Source Imaging                        | 464        |
| Superconductivity: Microwave Technology           | 152        | Magnets for Magnetic Resonance Analysis        |            |
| Superconductivity: Tunneling and Josephson        | 163        | and Imaging                                    | 479        |
| Junctions   | 177        | Microwave Ferroelectric Devices                | 491        |
| Superconductivity: Type I and II                  |            | Peltier Effect                                 | 499        |
|   |            | Power Quality                                  | 506        |
|   |            | SQUIDs   | 517        |
|   |            | Superconducting Cavity Resonators              | 528        |
| <b>SUPERCONDUCTORS</b>                            | <b>187</b> | Superconducting Cyclotrons and Compact         |            |
| Introduction                                      | 187        | Synchrotron Light Sources                      | 540        |
| A15 Superconductors                               | 187        | Superconducting Fault Current Limiters         | 544        |
| Cryogenic Stabilization                           | 188        | Superconducting Filters and Passive Components | 551        |
| Forced Flow Conductor Manufacturing               | 204        | Superconducting Levitation                     | 563        |
| HTS Film Growth                                   | 218        | Superconducting Magnets for Fusion Reactors    | 568        |
| HTS Josephson Junction Development                | 240        | Superconducting Magnets for Particle           |            |
| HTS Processing: Bulk, Thin Film, and Wires        | 248        | Accelerators and Storage Rings                 | 577        |
| HTS Superconductors, Physical Structures, and     | 260        | Superconducting Transformers                   | 596        |
| Role of Constituents                              | 280        | Temperature Sensors                            | 602        |
| Nb-Ti Alloy Superconductors                       | 295        | Thermoelectric Conversion                      | 623        |
| Stability In Forced Flow                          | 306        |  |            |
| Stabilization Against Flux Jumps                  | 317        | <b>INDEX</b>                                   | <b>633</b> |

# PREFACE

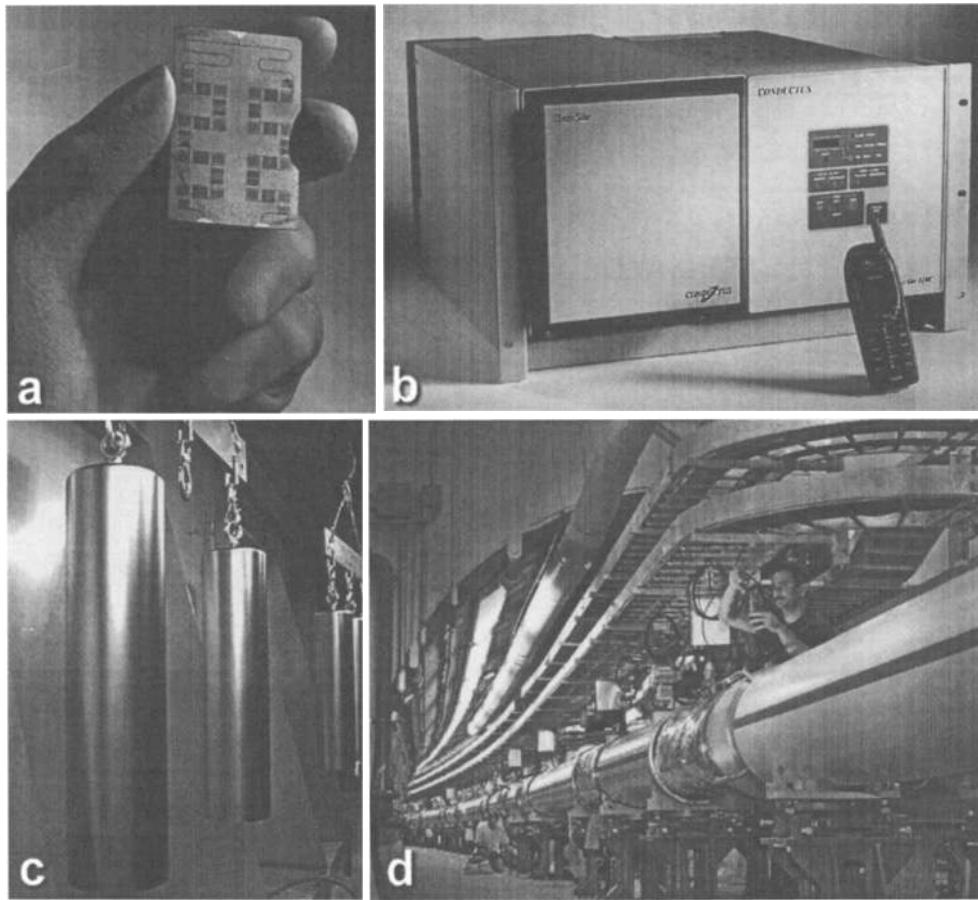
---

The beginning of the twenty-first century sees superconductor applications reaching important milestones (see Fig. 1). The first 900 MHz NMR has been constructed, with Nb<sub>3</sub>Sn superconductor supplying the required field of 21 T. The tau neutrino has been identified at the Fermi National Accelerator Laboratory, near Chicago, using a 4 mi ring of Nb–Ti superconducting magnets. Commercially produced cryogen-free receiver sub-systems are being installed at an accelerated pace in cellular-telephone base stations. Superconducting magnetic resonance imaging (MRI) has become both an essential tool of medicine and a key to unlocking the mysteries of the human body. The remarkable flux sensitivity of superconducting quantum interference devices (SQUIDs) makes both non-invasive magnetic cardiography and even magnetic encephalography possible. Having once been a fascinating but not particularly useful toy of solid-state physicists, superconductors now make possible the most advanced experiments in chemistry, biochemistry, particle physics, and health sciences. They are also becoming a part of the fabric of modern-day life.

This volume looks at the application of superconductors from an engineering perspective. All articles were written by acknowledged experts in their fields for the 24-volume *Wiley Encyclopedia of Electrical and Electronics Engineering* (J. Webster, ed.) and have been updated where necessary by the authors for developments since 1999. Each article has been reviewed by two other experts to ensure that it is clear and precise. The end of each article concludes with references to the current literature. The fundamentals of superconducting behavior are outlined, and the properties and fabrication methods of commercially produced superconductors are explained. Comprehensive articles on the applications of superconductors cover the uses of superconductors as well as competing technologies.

The articles are organized into three sections:

1. Superconductivity and Magnetism
  - Historical Introduction
  - Magnetic Flux
  - Magnetic Levitation
  - Magnetic Materials
  - Magnetic Noise, Barkhausen Effect
  - Magnetic Refrigeration
  - Magnetic Resonance
  - Superconductivity: Critical Current
  - Superconductivity: Electromagnets
  - Superconductivity: Electronics
  - Superconductivity: Hysteresis and Coupling Losses
  - Superconductivity: Magnets, Quench Protection
  - Superconductivity: Microwave Technology
  - Superconductivity: Tunneling and Josephson Junctions
  - Superconductivity: Type I and II
2. Superconductors
  - Introduction
  - A15 Superconductors
  - Cryogenic Stabilization
  - Forced-Flow Conductor Manufacturing
  - HTS Film Growth
  - HTS Josephson Junction Development
  - HTS Processing: Bulk, Thin Film, and Wires
  - HTS Physical Structures and the Role of Constituents
  - Nb-Alloy Superconductors
  - Stability in Forced Flow
  - Stabilization against Flux Jumps
3. Applications and Related Technology
  - Introduction
  - Capacitor Storage
  - Diagnostic Imaging
  - Frequency Converters and Mixers
  - Fusion Reactors
  - Gyrotrons
  - Hall-Effect Transducers
  - High-Energy Physics Particle-Detector Magnets
  - Infrared Detector Arrays, Uncooled
  - Magnetic Sensors
  - Magnetic Separation
  - Magnetic Source Imaging
  - Magnets for Magnetic Resonance Analysis and Imaging
  - Microwave Ferroelectric Devices
  - Peltier Effect
  - Power Quality
  - SQUIDs
  - Superconducting Cavity Resonators
  - Superconducting Cyclotrons and Compact Synchrotron Light Sources



**Figure 1.** Small- and large-scale superconductors and their applications. (a)  $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ -based microwave filter for (b) rack-mount front-end receiver subsystem for cellular telephone base stations (courtesy of Conductus, Inc.). (c) 136 kg Nb-47 wt% Ti alloy billets are destined for accelerator application. Each billet is 750 mm tall by 200 mm diameter (courtesy of Wah Chang). (d) A view of the superconducting magnets at Brookhaven National Laboratory's Relativistic Heavy Ion Collider (RHIC). The 1740 Nb-Ti-based magnets guide and focus gold particles along the collider's 4 km long tunnel at nearly the speed of light. (Courtesy of Brookhaven National Laboratory.)

Superconducting Fault-Current Limiters  
 Superconducting Filters and Passive  
 Components  
 Superconducting Levitation  
 Superconducting Magnets for Fusion Reactors

Superconducting Magnets for Particle Accelerators  
 and Storage Rings  
 Superconducting Transformers  
 Temperature Sensors  
 Thermoelectric Conversion

# CONTRIBUTORS

---

**David E. Andrews**

Oxford Instruments, Inc., Carteret, NJ

**Iltcho Angelov**

Chalmers University of Technology, Göteborg, Sweden

**Daniël H. J. Baert**

University of Ghent, Belgium

**Athanasios Bardas**

State University of New York at Stony Brook

**Henry Blosser**

Michigan State University, East Lansing

**Heinrich J. Boenig**

Los Alamos National Laboratory, NM

**Punit Boolchand**

University of Cincinnati, Cincinnati, OH

**Luca Bottura**

CERN, Geneva, Switzerland

**Pierluigi Bruzzone**

Centre de Recherches en Physique du Plasmas, Villigen, Switzerland

**Thomas F. Budinger**

University of California at Berkeley

**Donald P. Butler**

Southern Methodist University, Dallas, TX

**Zeynep Çelik-Butler**

Southern Methodist University, Dallas, TX

**Kookrin Char**

Conductus, Inc., Sunnyvale, CA

**Durga P. Choudhury**

Northeastern University, Boston, MA

**Terry S. Davies**

University of the West of England, Bristol, England

**Alex de Lozanne**

The University of Texas at Austin

**Arnaud Devred**

Atomic Energy Commission at Saclay, Cedex, France

**Reinhard Dietrich**

Vacuumschmelze GmbH, Hanau, Germany

**F. J. Edeskuty**

Edeskuty Engineering, Los Alamos, NM

**Hal Edwards**

The University of Texas at Austin

**Halit Eren**

Curtin University of Technology, Perth, Australia

**Robert L. Fagaly**

Tristan Technologies, San Diego, CA

**D. K. Finnemore**

Iowa State University, Ames

**Martin G. Forrester**

Northrop Grumman Science and Technology Center, Pittsburgh, PA

**Michael A. Green**

Lawrence Berkeley National Laboratory, Berkeley, CA

**Larry R. Grisham**

Princeton University, NJ

**M. S. Hämäläinen**

Helsinki University of Technology, Finland

**I. D. Hepburn**

University College, London

**John R. Hull**

Argonne National Laboratory, IL

**Brian D. Hunt**

NASA Jet Propulsion Laboratory, Pasadena, CA

**R. Lawrence Ives**

Calabazas Creek Research Inc., Saratoga, CA

**K. Jagannadnam**

North Carolina State University, Raleigh, NC

**Raghavan Jayakumar**

Lawrence Livermore National Laboratory, Livermore, CA

**Colin M. Jefferson**

University of the West of England, Bristol

**Mark Jeffery**

University of California at Berkeley

**Swarn S. Kalsi**

American Superconductor Corporation, Westborough, MA

**R. Kalyanaraman**

North Carolina State University, Raleigh, NC

**Erik L. Kollberg**

Chalmers University of Technology, Göteborg, Sweden

**D. Kumar**

University of Florida, Gainesville, FL

**Peter J. Lee**

University of Wisconsin—Madison

**César Luongo**

Bechtel, San Francisco, CA

**P. Maccioni**

Euratom-CEA Association, Cedex, France

**Vladimir Matijasevic**

Conductus, Inc., Sunnyvale, CA

**Mitsuhiro Motokawa**

Tohoku University, Sendai, Japan

**J. Narayan**

North Carolina State University, Raleigh, NC

**J. T. Nenonen**

Helsinki University of Technology, Finland

**Qian Niu**

The University of Texas at Austin

**Hassan Nouri**

University of the West of England, Bristol, United Kingdom

**S. Oktyabrsky**

North Carolina State University, Raleigh, NC

**Hasan Padamsee**

Cornell University, Ithaca, NY

**Martha Pardavi-Horvath**

The George Washington University, Washington, DC

**D. P. Patel**

Naval Research Laboratory, Washington, DC

**Scott D. Peck**

Houston Advanced Research Center, The Woodlands, TX

**J. M. Pond**

Naval Research Laboratory, Washington, DC

**J. B. L. Rao**

Naval Research Laboratory, Washington, DC

**Christopher M. Rey**

DuPont Superconductivity, Wilmington, DE

**S. T. Ruggiero**

University of Notre Dame, Notre Dame, IN

**Lembit Salasoo**

General Electric Company, Schenectady, NY

**Joel H. Schultz**

Massachusetts Institute of Technology, Cambridge

**D. P. Sekulic**

University of Kentucky, Lexington

**Rajiv K. Singh**

University of Florida, Gainesville, FL

**A. Smith**

University College London, England

**S. Sridhar**

Northeastern University, Boston, MA

**Guru Subramanyam**

University of Dayton, Dayton, OH

**Salvador H. Talisa**

Northrop Grumman Corporation, Baltimore, MD

**John Talvacchio**

Northrop Grumman Science and Technology Center,

Pittsburgh, PA

**B. Turck**

Euratom-CEA Association, Cedex, France

**Robert B. van Dover**

Bell Labs, Lucent Technologies, Murray Hill, NJ

**William H. Warnes**

Oregon State University, Corvallis

**John E. C. Williams**

Massachusetts Institute of Technology, Cambridge, MA

**Jan Zakrzewski**

Silesian Technology, University of Poland

**John C. Zeigler**

Houston Advanced Research Center, TX

**Herbert Zirath**

Chalmers University of Technology, Göteborg,

Sweden



# SUPERCONDUCTIVITY AND MAGNETISM

## HISTORICAL INTRODUCTION

### The Early Years of Superconductivity

In the early years of superconductivity, progress to application was slow and intermittent. On July 10, 1908, Heike Kamerlingh Onnes, professor of experimental physics at the University of Leiden (Holland), was able to liquefy helium for the first time. Not only was he able to determine a boiling point for helium, of 4.3 K (now more precisely 4.2 K), but he was also able to further reduce the temperature to 1.7 K by reducing the pressure on the helium bath.

He soon set about measuring the electrical resistance of metals in the new temperature regime. The resistance of metals is strongly dependent on temperature, and once the dependence has been accurately measured, the resistance can be used as a simple and convenient tool for low-temperature thermometry. The low-temperature behavior of metals was also seen as a tool to study electron theory (for instance by Albert Einstein, who had applied to be an assistant to Onnes in 1901, but had been rejected).

The Leiden group initially extended measurements on platinum wires into the liquid-helium range, and observed that their electrical resistance fell continuously with decreasing temperature to a minimum non-zero value. The minimum resistance decreased as the impurity level of the metal decreased. Looking for a material with a higher available purity, they showed that the resistance of high-purity gold fell to an even lower but measurable value. Seeking an even higher-purity metal, considerable effort was then expended distilling mercury. When the resulting high-purity mercury was tested, the electrical resistance fell steeply but continuously, as expected. At the boiling point of the helium (4.2 K), the resistance of the mercury wire had fallen to 500 times less than it had been at the melting point of the mercury.

What happened next came as a complete surprise. As the mercury wires were slowly cooled below 4.2 K, Gilles Holst (who had been an assistant to Leiden for two years and would later become the first director of the Philips Research Laboratories) measured a sudden and massive drop in the electrical resistance. As best as they could measure, in just a few hundredths of a degree, the resistance dropped to less than one millionth of the melting-point value, and eventually to a billionth of it. In 1912, Onnes termed the new electrical state that the mercury had entered the *superconductive state*.

Having worked so hard to purify the mercury, he was further surprised to find that adding gold and cadmium to the mercury did not stop it from entering the superconducting state. He also observed that very high currents could be passed through the mercury until a threshold current density was reached (as high as 1000 A/mm<sup>2</sup> at 2.45 K), at which point the mercury would return to the normal electrical state (1). This threshold value, which we now term the *critical current*, is perhaps the most important property for practical application.

In December 1912, he discovered that other metals, which could be made into wires at room temperature, could be made superconducting (2), namely tin (3.8 K) and lead (6 K, later raised to 7.2 K). The first two superconducting solenoids were quickly manufactured by G. J. Flim (1875–1900), the chief of the

Technical Department of the (Leiden) Cryogenic Department. At the Third International Congress of Refrigeration, held in Chicago in September 1913, Onnes predicted that superconductivity would enable the production of coils that could generate fields (100,000 G or 10 T) well in excess of that possible with conventional conductors (3). For comparison, the flux density between the poles of a horseshoe permanent magnet is 0.1 T.

Less accurately, he predicted that such a development should not be far away. He was about to discover a roadblock to high-field superconductivity. In a footnote to his Chicago address, he observed that a 0.05 T (500 G) field, developed in a simple superconducting solenoid, was sufficient to revert the superconductor to its normal state. By 1914, Leiden had produced curves of resistance as a function of applied field and had developed an empirical fit to the temperature dependence of the critical field  $H_c$ :  $H_c(T) \approx H_{c0}[1 - (T/T_c)^2]$ . It seems surprising to us now that it was not until 1916 that the interdependence of critical current and critical field was shown from an analysis of the Leiden data by F. Silsbee of the National Bureau of Standards in America (4). For many years, it appeared that the low critical currents and the suppression of superconductivity with very small applied fields would make superconductors impractical for any application other than laboratory studies of solid-state physics.

### Higher Critical Magnetic Fields and Critical Temperatures

Leiden enjoyed a monopoly of liquid-helium research until after the First World War. In 1923, a helium liquefier, based on the Leiden design, started operation at the University of Toronto. Four years later, a helium liquefier capable of 10 L/h was started at the Physikalisch-Technische Reichsanstalt (PTR) near Berlin under the direction of Walther Meissner. In successive years from 1928 to 1930, the PTR identified three important new superconductors: tantalum ( $T_c = 4.4$  K), thorium ( $T_c = 1.4$  K) and niobium ( $T_c = 9.2$  K) (5). An alloy of niobium, Nb–47 wt% Ti, is now by far the most important commercial superconductor, with widespread use in magnets for magnetic resonance imaging (MRI) systems in hospitals as well as many other applications (see Ductile Superconductors). Nb-based technology is also the current standard for digital superconducting circuits (see Superconducting Electronics). Meissner's group would go on to find that most of the transition elements in groups IV and V were superconducting. In Fig. 1 we show a listing of the elemental superconductors with their locations in the periodic table.

In the same period a further important discovery came from the group of Wander Johannes de Haas, who became codirector of the Leiden laboratory in 1924 (with Willem Hendrik Keesom). It was found that a solid solution of 4% bismuth in gold was found to be superconducting at 1.9 K (6) despite neither of the components being superconducting at ambient pressure. A similar result was found later the same year when copper sulfide ( $T_c = 1.1$  K) was examined by Meissner (7). This time an insulator (sulfur) had been combined with a very good normal conductor (Cu) to produce a superconductor. The Meissner group went on to find a large number of carbides and nitrides with high transition temperatures, in particular NbC ( $T_c > 10$  K).

|                   | IA    | IIA           | IIIB  | IVB            | VB            | VIB             | VII B          | VIII          | VIII              | IB              | IIIB  | IIIA           | IVA           | VA            | VIA    | VIIA   | 0              |       |
|-------------------|-------|---------------|-------|----------------|---------------|-----------------|----------------|---------------|-------------------|-----------------|-------|----------------|---------------|---------------|--------|--------|----------------|-------|
| 1                 | 1 H   |               |       |                |               |                 |                |               |                   |                 |       |                |               |               |        |        | 2 He           |       |
| 2                 | 3 Li  | 4 Be<br>0.026 |       |                |               |                 |                |               |                   |                 |       | 5 B            | 6 C           | 7 N           | 8 O    | 9 F    | 10 Ne          |       |
| 3                 | 11 Na | 12 Mg         |       |                |               |                 |                |               |                   |                 |       | 13 Al<br>1.175 | 14 Si         | 15 P          | 16 S   | 17 Cl  | 18 Ar          |       |
| 4                 | 19 K  | 20 Ca         | 21 Sc | 22 Ti<br>0.40  | 23 V<br>5.40  | 24 Cr           | 25 Mn          | 26 Fe         | 27 Co             | 28 Ni           | 29 Cu | 30 Zn<br>0.85  | 31 Ga<br>1.10 | 32 Ge         | 33 As  | 34 Se  | 35 Br          | 36 Kr |
| 5                 | 37 Rb | 38 Sr         | 39 Y  | 40 Zr<br>0.61  | 41 Nb<br>9.25 | 42 Mo<br>0.912  | 43 Tc<br>7.80  | 44 Ru<br>0.49 | 45 Rh<br>0.0003   | 46 Pd           | 47 Ag | 48 Cd<br>0.517 | 49 In<br>3.4  | 50 Sn<br>3.72 | 51 Sb  | 52 Te  | 53 I           | 54 Xe |
| 6                 | 55 Cs | 56 Ba         | 57 La | 58 Hf<br>0.128 | 59 Ta<br>4.47 | 60 W<br>0.0154  | 61 Re<br>1.697 | 62 Os<br>0.66 | 63 Ir<br>0.113    | 64 Pt<br>0.0019 | 65 Au | 66 Hg<br>4.15  | 67 Tl<br>1.70 | 68 Pb<br>7.2  | 69 Bi  | 70 Po  | 71 At          | 72 Rn |
| 7                 | 87 Fr | 88 Ra         | 89 Ac |                |               |                 |                |               |                   |                 |       |                |               |               |        |        |                |       |
| Lanthanide series |       |               |       | 58 Ce          | 59 Pr         | 60 Nd           | 61 Pm          | 62 Sm         | 63 Eu             | 64 Gd<br>1.083  | 65 Tb | 66 Dy          | 67 Ho         | 68 Er         | 69 Tm  | 70 Yb  | 71 Lu<br>0.100 |       |
| Actinide series   |       |               |       | 90 Th<br>1.38  | 91 Pa<br>1.4  | 92 U<br>0.6/1.8 | 93 Np          | 94 Pu         | 95 Am<br>1.1/0.79 | 96 Cm           | 97 Bk | 98 Cf          | 99 Es         | 100 Fm        | 101 Md | 102 No | 103 Lr         |       |

Figure 1. Elemental superconductors.

A further breakthrough occurred in 1929 when de Haas and Voogd found that the magnetic threshold for  $\text{Bi}_5\text{Tl}_3$  at 3.4 K exceeded their available applied field of 0.53 T. The following year they were able to borrow a more powerful electromagnet and showed that they could maintain superconductivity in a eutectic Pb–Bi alloy at 2 K in an applied field of 1.9 T (8). Finally, Kamerlingh Onnes' dream of high-power superconducting magnets was possible. Unfortunately, the first attempts to make coils exploiting these results (at the Clarendon Laboratory in Oxford under Kurt Mendelssohn, by Keesom at Leiden, and by L. V. Shubnikov's group at the Kharkov Institute in the Ukraine) proved to be so disappointing that they did not inspire the necessary effort to resolve their limitations.

**The Meissner Effect and Type II Superconductivity.** In 1933 an important discovery was made by Meissner and his student Robert Ochsenfeld, using cylinders of single-crystal and polycrystalline lead. Investigating the transition to the superconducting state in magnetic field, they found that whatever path was used to apply the field, magnetic flux was expelled from the interior below the critical field  $H_c$  (9). The path independence results from the thermodynamic reversibility of the superconducting transition. If field is applied when the superconductor is in the superconducting state, surface currents which counteract the external field are produced. Just two years later, the brothers Fritz and Heinz London developed a set of electrodynamic equations, now known as the London equations, which described the Meissner effect by supplementing Maxwell's equations (10). A consequence of these equations is the London penetration depth,  $\lambda_L$ , which is the skin depth

that a magnetic field can penetrate into a type I superconductor. See the articles MAGNETIC NOISE, BARKHAUSEN EFFECT (under "Flux Pinning and Losses in Superconductors") in this section and HTS SUPERCONDUCTORS, PHYSICAL STRUCTURES, AND ROLE OF CONSTITUENTS (under "Common Features of High- $T_c$  Superconductors") in the next section titled "Superconductors."

The same year as London and London's paper was published, the group of L. V. Shubnikov (Figure 2) at Kharkov showed that single crystals of  $\text{PbTl}_2$  had two distinct critical fields (11). Up to a lower critical field ( $H_{c1}$ ), the flux is excluded; above that field the flux begins to penetrate and increases in its penetration until an upper critical field ( $H_{c2}$ ) is reached, when the flux completely penetrates and superconductivity is extinguished. The superconductors that show this characteristic would come to be known as *type II* superconductors. This class includes all the technically useful superconductors, including all alloy and compound superconductors as well as the elements niobium, vanadium, and technetium. See SUPERCONDUCTIVITY: TYPE I AND II, and SUPERCONDUCTIVITY: CRITICAL CURRENT. Unfortunately, the importance of the work at Kharkov was not fully appreciated outside Soviet Union, as Shubnikov's group was caught up in one of Stalin's purges and Shubnikov himself died in prison in 1945. Only with his posthumous exoneration was it possible for his Soviet colleagues to openly acknowledge his contributions to their work.

Although the Second World War resulted in a hiatus in superconductor activity for most groups, there was one significant advance. In Germany, Ascherman and coworkers found that niobium nitride had a  $T_c$  of 15 K (12), the first time the liquid-helium region had been surpassed. [The boiling



**Figure 2.** Lev Vasil'evich Shubnikov, discoverer of type II superconductivity (courtesy of the Kharkov Institute of Science and Technology, Ukraine).

temperature for a cryogen can be fixed at any temperature between the triple point and the critical point by maintaining the corresponding system pressure. Above helium (which does not have a true triple point), the next cryogen is liquid hydrogen, which has a triple point of 13.80 K and a critical point of 32.98 K.]

### Ginzburg, Landau, and Abrikosov

Despite the huge potential of superconductivity, the difficulty in obtaining liquefied helium had limited research on it to a handful of laboratories worldwide. However, just as the advances in cryogenics at the turn of the twentieth century had made the discovery of superconductivity possible, so had advances in cryogenic technology, particularly the commercialization of the Collins liquefier, broaden the availability of the superconducting state after the Second World War. 1950 saw the publication of Vitalii Ginzburg and Lev Landau's landmark modifications to the London equations (13). The resulting Ginzburg–Landau equations are described in the articles SUPERCONDUCTIVITY: TYPE I AND II and SUPERCONDUCTIVITY: CRITICAL CURRENT and in the article HTS SUPERCONDUCTORS, PHYSICAL STRUCTURES AND ROLE OF CONSTITUENTS in the next section titled “Superconductors.” An important new material parameter, the coherence length  $\xi$ , was defined as the distance over which the density of the superconducting electrons decreases at a superconducting–normal

interface. The Ginzburg–Landau formulations were able to predict the conditions under which type I and type II behavior would occur using the parameter  $\kappa = \lambda/\xi$ . From this work came the now familiar flux-line lattice description of type II superconductors by Landau's student Aleksei Abrikosov, eventually published in 1957 (14). See Superconducting Critical Current. In type II superconductors, the flux tube has a radius  $\lambda$  with an internal normal core of radius  $\xi$ .

In 1953 Bernd Matthias at the Bell laboratories raised the  $T_c$  ceiling for superconductors to 17.86 K with NbN–NbC. That discovery was followed the same year by John Hulm's group at the University of Chicago with another 17 K superconductor  $V_3Si$ , but of a new crystal structure, A15, that would eventually supply a series of important superconductors. Another A15,  $Nb_3Sn$ , would be added the following year at Bell Labs, with a further increase in  $T_c$  to 18 K.

### The Beginning of Engineering Superconductivity

It was not until 1954 that the first successful superconducting magnet was made (by G. B. Yntema at the University of Illinois), thereby ushering in the age of engineering superconductivity. Yntema used Nb wire, which had been shown (by D. Shoenberg at Cambridge University) to have a markedly better critical field than any of the other known superconductors. The resulting magnet produced a field of 0.71 T at 4.2 K. He also discovered that cold-working the strands markedly increased the current density that they could carry. It was beginning to be clear that critical current was, to a major extent, a property that could be increased independently of the intrinsic bulk properties  $H_{c2}$  and  $T_c$ . By August 1960 Stan Autler (at MIT Lincoln Laboratory) had produced a 2.5 T field at 4.2 K. Even more significantly, he had applied the persistent current in a solenoid to provide the magnetic field for a solid-state maser, perhaps the first application of superconductivity. A flurry of activity followed, focused on the high- $T_c$ , high- $H_{c2}$  A15 compound  $Nb_3Sn$ .  $Nb_3Sn$ , at that time, however, was difficult to fabricate into magnets because of its brittleness and because the simple elemental-powder-in-tube process required a 1000°C heat treatment. It was soon displaced by two ductile alloy superconductors, first Nb–Zr ( $T_c \approx 12$  K) and then Nb–Ti ( $T_c = 7$  K to 10 K).

Whereas the Ginzburg–Landau theory coupled with Abrikosov's work provided an enduring phenomenological description of superconductivity, it did not provide a microscopic description. That was to be supplied in 1957 when John Bardeen, Leon Cooper, and Robert Schrieffer of the University of Illinois at Urbana published their Nobel Prize-winning theory of superconductivity (15). The superconducting carriers of the phenomenological description are two electrons [Cooper pairs (16)] with equal or opposite spin and momentum. The coherence length is the size of the Cooper pair, and the order parameter is proportional to the electron energy gap, which itself is proportional to  $T_c$ . Three years later, Lev Gorkov showed how to unify the phenomenological and microscopic models and to incorporate strong magnetic fields (17). The key to Gorkov's description was a variation in the energy-gap parameter with position.

The next major theoretical advance came in 1962, from a graduate student at Cambridge University, Brian D. Josephson. He predicted that superconducting current would tunnel

through a thin insulating layer, or *weak link*, separating two superconducting electrodes, and that a phase difference would be produced between the superconducting electrons in the two electrodes. The phase difference generates a voltage difference between the two electrodes. The Josephson effect, as it is now known, is the basis for superconducting electronic devices such as the Superconducting quantum interference device (SQUID) and very high-precision voltage standards. It would earn Josephson a Nobel Prize.

### High Field Magnets

Much of the theoretical development in the 1960s and 1970s was in the area of flux pinning. Increasing the critical current density in superconductors reduces costs, because less superconductor is required; it also makes it possible to operate magnets at higher magnetic fields. When current flows through a superconductor it produces a Lorentz force acting to move the flux lines. If the Lorentz force is allowed to move the flux-line lattice freely within the superconductor, then power is dissipated, and eventually the resulting heating may drive the superconductor normal. Introducing microstructural features, such as nonsuperconducting precipitates and grain boundaries, that pin the flux lines in place can, however, restrict movement of the flux-line lattice. Understanding the nature of flux pinning is key to understanding how to improve the critical current density in superconductors. The mechanisms and theory of flux-pinning are comprehensively reviewed in the corresponding sections below.

From the late 1960s onwards, the needs of the high-energy physics community propelled considerable advances in superconducting strand technology, initially for bubble-chamber and then for accelerator magnets. In March 1983, the first superconducting accelerator ring was completed at Fermi National Accelerator Laboratory. With 774 dipole magnets 6 m long and 210 quadrupole magnets covering a 4 mi circle, it exemplifies the progress that had been made. The key features were now in place: the superconducting strand was in the required form of a composite of fine ( $>30\ \mu\text{m}$  diameter) filaments in a high-purity, high-normal-conductivity Cu (or Al) matrix for stability (see Superconductors, Cryogenic Stabilization), and the strand was twisted in order to reduce filament coupling currents. Superconducting strands were cabled together to form a thick ribbon-like conductor. Increased understanding of the microstructural development of the superconductor and its role in flux pinning would further increase the critical current density, making possible the next generation of accelerators; see article "Nb-Alloy Superconductors." Four years later, the largest superconducting magnet yet was fabricated for the DELPHI project at the CERN particle accelerator laboratory. The 7.4 m long, 6.2 m diameter, 84 tonne magnet survived a 1600 km trip to CERN by road, ship, and barge.

In addition to magnet technology, the high-energy physics community also benefited from superconducting cavity technology (see Superconducting Cavity Resonators). When the Large Electron-Positron (LEP) Collider was initially run in 1989, with 128 conventional copper accelerating cavities, they provided enough energy to take the energy of each beam to 50 GeV. After upgrading with 272 superconducting cavities, the LEP ring was eventually able to reach 104 GeV per beam in April 2000.

In 1986, Alex Müller and Georg Bednorz, at the IBM Research Laboratory in Rüschlikon, Switzerland, made a ceramic perovskite of lanthanum, barium, copper, and oxygen that superconducted at 35 K (18). In fact, small amounts of this material were later found to be superconducting at 58 K due to lead impurities. The impact of this discovery can be gauged by the almost immediate awarding of the Nobel Prize to the two discoverers. The following year the research groups of Paul Chu at the University of Houston and Maw-Kuen Wu at the University of Alabama at Huntsville substituted yttrium for lanthanum and produced a ceramic that superconducts at 92 K (19). Now in the space of a year the highest  $T_c$  had been raised from 23.2 K (for  $\text{Nb}_3\text{Ge}$  discovered in 1973 by John Gavaler) through 35 K to 92 K, well above the temperature of liquid nitrogen (77 K). A further huge jump in  $T_c$  came in 1988 from Allen Hermann and Z. Z. Sheng of the University of Arkansas with a 120 K Tl-Ca-Ba-Cu-O superconductor (20). In 1993 A. Schilling, M. Cantoni, J. D. Guo, and H. R. Ott from Zurich, Switzerland, measured a  $T_c$  of 133 K in  $\text{HgBa}_2\text{Ca}_2\text{Cu}_3\text{O}_8$  (21). The partial substitution of thallium in this high- $T_c$  mercury-based oxide by P. Dai and B. C. Chakoumakos (Oak Ridge National Laboratory), G. F. Sun and K. W. Wong (University of Kansas), and Y. Xin and D. F. Lu (Midwest Superconductivity Inc.) increased the  $T_c$  to 138 K for a nominal composition of  $\text{Hg}_{0.8}\text{Tl}_{0.2}\text{Ba}_2\text{Ca}_2\text{Cu}_3\text{O}_{8+\delta}$  (22).

We have come to expect continual advances in the high properties of both high-temperature superconductors (HTSs) and low-temperature superconductors (LTSs), and have yet to be disappointed. In the years since the discovery of the HTSs we have seen steady improvements in their current densities, on both the laboratory and the industrial scale, as well as in the production piece lengths.

The advent of HTSs accelerated the application of LTSs as well. Not only was there a new market for LTS-based test facilities to measure the properties of new HTSs, but there was increased investment in refrigeration technology and renewed public interest. Cryocoolers now allow for the cryogen-free use of LTS-based magnets, increasing public acceptance of the technology for everyday use. HTSs in the form of current leads (there are now several commercial vendors) make large-scale LTS applications much more energy-efficient. HTSs also make new applications possible, such as microwave filters, and make old ideas, such as power transmission lines, more promising.

### BIBLIOGRAPHY

1. H. Kamerlingh Onnes, Further experiments with liquid helium. H. On the electrical resistance of pure metals etc. VII. The potential difference necessary for the electric current through mercury below 4.19 K (continuation), *Comm. Phys. Lab. Leiden*, **133b**, 29, 1913.
2. H. Kamerlingh Onnes, Further experiments with liquid helium. H. On the electrical resistance of pure metals etc. (continued). VIII. The sudden disappearance of the ordinary resistance of tin, and the super-conductive state of lead, *Comm. Phys. Lab. Leiden*, **133d**, 51, 1913.
3. H. Kamerlingh Onnes, Report on the researches made in the Leiden cryogenics laboratory between the second and third international congress of refrigeration: Superconductivity, *Comm. Phys. Lab. Leiden Suppl.*, **34b**: 55-70, 1913.
4. F. B. Silsbee, A note on electrical conduction in metals at low temperatures, *Washington Acad. Sci. J.* **6**: 597-602, 1916.

5. W. Meissner and H. Franz, Messungen mit Hilfe von flüssigen Helium. VIII. Supraleitfähigkeit von Niobium, *Physikalisch-Technische Reichsanstalt Mitteilug*, 558–559, 1930.
  6. W. J. De Haas, E. van Aubel, and J. Voogd, A superconductor consisting of two nonsuperconductors, *Akad. Wetenschappen Amsterdam Proc.*, **32**: 730, 1929.
  7. W. Meissner, Messungen mit Hilfe von flüssigem Helium. V. Supraleitfähigkeit von Kupfersulfid, *Physikalisch-Technische Reichsanstalt Mitteilug*, 571, 1929.
  8. W. J. de Haas and J. Voogd, The influence of magnetic fields on supraconductors, *Akad. Wetenschappen Amsterdam Proc.* **33**: 262–270, 1930.
  9. W. Meissner and R. Oschenfeld, Ein neuer Effect bei Eintritt der Supraleitfähigkeit, *Naturwiss.*, **21**: 787–788, 1933.
  10. F. London and H. London, The electromagnetic equations of the supraconductor, *Proc. Roy Soc. London Ser. A*, 71–88, 1935.
  11. J. N. Rjabinin and L. V. Schubnikov, Magnetic properties and critical currents of superconducting alloys, *Phys. Z. Sowjetunion*, **6**: 605–607, 1935.
  12. G. Aschermann et al., Supraleitfähige Verbindungen mit extrem hohen Sprungtemperaturen (NbH und NbN), *Phys. Z.* **42**: 349–60, 1941.
  13. V. L. Ginzburg and L. D. Landau, On the theory of superconductivity, *Zh. Eksp. Teor. Fiz.* **20**: 1064–1082, 1950.
  14. A. A. Abrikosov, On the magnetic properties of superconductors of the second group, *Sov. Phys. JETP*, **5**: 1174–1182, 1957.
  15. J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Theory of superconductivity, *Phys. Rev.*, **108**: 1175–1204, 1957.
  16. L. N. Cooper, Bound electron pairs in a degenerate Fermi gas, *Phys. Rev.*, **104**: 1189–1190, 1956.
  17. L. P. Gorkov, Theory of superconducting alloys in a strong magnetic field near the critical temperature, *Sov. Phys. JETP*, **10**: 998–1004, 1960.
  18. G. Bednorz and K. A. Müller, Possible high  $T_c$  superconductivity in the Ba–La–Cu system, *Z. Phys. B*, **64**: 189–197, 1986.
  19. M. K. Wu et al., Superconductivity at 93 K in a new mixed-phase Y–Ba–Cu–O compound system at ambient pressure, *Phys. Rev. Lett.*, **58**: 908–910, 1987.
  20. Z. Z. Sheng and A. M. Hermann, 90 K Tl–Ba–Cu–O and 120 K Tl–Ca–Ba–Cu–O bulk superconductors, *Proc. 1988 World Congress on Superconductivity*, Singapore: World Scientific, pp. 365–376, 1988.
  21. M. Cantoni et al., Characterisation of superconducting Hg–Ba–Ca–Cu-oxides. Structural and physical aspects, *Physica C*, **215**: (1–2): 11–18, 1993.
  22. P. Dai et al., Synthesis and neutron powder diffraction study of the superconductor  $\text{HgBa}_2\text{Ca}_2\text{Cu}_3\text{O}_{8+\delta}$  by Tl substitution, *Physica C*, **243**(3–4): 201–206, 1995.
  23. G. Hammerl et al., Enhanced supercurrent density in polycrystalline  $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$  at 77 K from calcium doping of grain boundaries, *Nature*, **407**: 162–164, 2000.
- A. C. Rose-Innes and F. H. Rhoderick, *Introduction to Superconductivity*, Oxford, UK: Pergamon, 1969.
- M. Tinkham, *Introduction to Superconductivity*, New York: McGraw-Hill, 1975.
- T. P. Orlando and K. A. Delin, *Foundations of Applied Superconductivity*, Reading, MA: Addison-Wesley, 1991.
- A. M. Campbell and J. E. Evetts, *Adv. Phys.*, **21**: 199–428, 1972.

## MAGNETIC FLUX

The magnetic flux is related to the number of magnetic lines of force crossing a given area. It is therefore analogous to the flux of a flowing quantity. The magnetic flux is defined as the integral of the product of an elemental area and the magnetic induction perpendicular to it. If the magnetic induction is constant over a given area, then the flux is the product of the magnetic induction and the area. The magnetic induction is referred to as the magnetic flux density, since it is the magnetic flux per unit area.

Since the flux represents the number of lines of force and the lines of force are fixed in a infinitely conducting medium, the flux is conserved in an ideal conducting medium. Finite high-conductivity media such as copper and high-temperature plasma conserve (enclosed) flux when an external field change occurs or the cross section is changed over a time short compared to the time scale in which the lines of force (and the magnetic field) can diffuse across the medium.

The magnetic flux is related to the energy stored in the magnetic field and represents the capability of a primary magnet to induce voltage in a coupled secondary circuit over a time duration. When a magnet current is changed the flux (magnetic induction) changes, and by Lenz's law (discussed later) it induces a voltage in a secondary circuit. The magnitude of the voltage depends upon the rate of change of flux. The duration over which this voltage can be maintained is thus proportional to the flux:

$$\Phi = \int (\mathbf{B} \cdot d\mathbf{S})$$

where the integral of the magnetic flux density  $B$  is over the area of interest  $S$ . Since the surface area vector is normal to the surface, the integral gives the flux that intersects the surface. If  $B$  is constant over  $S$ ,

$$\Phi = BS$$

Since the flux density is related to the magnetic field intensity  $H$  by the relation  $B = \mu_r \mu_0 H$ , where  $\mu_r$  is the relative permeability of the medium and  $\mu_0$  is the permeability of vacuum,

$$\Phi = \mu_0 \int (\mu_r \mathbf{H} \cdot d\mathbf{S})$$

The MKS unit for magnetic flux is Tesla =  $\text{m}^2$  or Webers or volt-seconds.

### Reading List

In 1986 the Applied Superconductivity Conference celebrated the 75th Anniversary of the discovery of superconductivity with a symposium on the history of superconductivity. The symposium is published in full in *IEEE Trans. Magn.*, **23**: 354–415, 1986.

A recent text that is both informative and entertaining is Per Fridtjof Dahl, *Superconductivity: Its Historical Roots and Development from Mercury to the Ceramic Oxides*.

Classical and still excellent introductions to the phenomenon of superconductivity are:

### Magnetic Flux and Vector Potential

The vector potential is a quantity closely related to flux and is defined as

$$\mathbf{B} = \text{curl } \mathbf{A}$$

The flux is then given by

$$\Phi = \int (\text{curl } \mathbf{A} \cdot d\mathbf{S})$$

Applying Stoke's theorem,

$$\Phi = \int (\mathbf{A} \cdot d\mathbf{l})$$

where the integral is over the closed loop enclosing the surface area of interest. Therefore the flux is the line integral of the vector potential around the perimeter of the area of interest. For example, for a coil which has only an azimuthal component  $A_\phi$ , the flux enclosed by a circle of radius  $r$  is  $2\pi r A_\phi$ . The equation is commonly used to create flux plots (lines of force) by using constant-vector-potential lines and to calculate the flux enclosed by a given area, since many electromagnetic and magnetostatic problems are solved by solving for the vector potential.

### Magnetic Flux and Inductance

The inductance  $L$  of an electrical circuit depicts the ability of the circuit to oppose a change of current in its own circuit or a mutually (magnetically) coupled circuit. More fundamentally it is the ability of a circuit to oppose the change in magnetic flux enclosed by the current circuit. The flux enclosed by a circuit is proportional to the current in the circuit or the circuit that is mutually coupled to it, and the proportionality constant is the self-inductance  $L$  or the mutual inductance  $M$ , or

$$\Phi = LI \quad \text{or} \quad \Phi' = MI'$$

where  $I$  is the current in the circuit whose flux is of interest and  $I'$  is the current in the mutually coupled circuit. The mutual inductance of two circuits with self-inductance  $L_i$  and  $L_j$  is given by

$$M_{ij} = K(L_i L_j)^{1/2}$$

where  $K$  is the coupling coefficient. The mutual inductance may be positive or negative.

For a coil with  $N$  turns, the total inductance is given by

$$L_{\text{total}} = \sum_{i=1}^N \sum_{j=1}^N (L_i + M_{ij})$$

If the turns are identical and fully coupled ( $K = 1$ ) to each other,

$$L_{\text{total}} = \sum_{i=1}^N [L_i + (N-1)L_i] = N^2 L_i = N^2 \Phi / I$$

where  $\Phi$  is the flux due to one turn when a current  $I$  passes through it. Since the flux  $\Phi_N$  induced by  $N$  turns is  $N\phi$ ,

$$L_{\text{total}} = N\Phi_N / I$$

### Magnetic Energy Density

The volume permeated by a magnetic field stores energy and therefore any device which generates a magnetic field also stores energy. The energy density associated with the magnetic field of a region is given by

$$u = (\mathbf{B} \cdot \mathbf{H}) / 8 = B^2 / (8\mu_r \mu_0)$$

so that the energy stored in a flux tube cylinder with a cross sectional area  $A$  (perpendicular to  $\mathbf{B}$ ) and length  $l$  is

$$E_f = \Phi^2 l / (8\mu_r \mu_0 A)$$

### Lenz's Law and Flux Conservation

As stated previously, an electrical circuit such as a loop of wire or a metallic cylinder opposes a change in the flux enclosed by itself. A voltage is induced in the circuit in a direction such that the voltage can drive a current that opposes the change in the flux. Therefore Lenz's law states that the induced voltage is given by the time rate of change of the flux, that is,

$$V = -d\Phi / dt \quad (1)$$

The voltage drives a current  $I$  given by

$$V = L dI / dt + IR \quad (2)$$

where  $L$  and  $R$  are the inductance and resistance of the circuit. Equating Eqs. (1) and (2) and integrating over time,

$$LI + \Phi = \text{const} - \int (IR dt)$$

Now, the left-hand side is the total flux (sum of the initial flux and the induced flux). Therefore the flux is conserved if the electrical circuit has zero resistance, that is, for an ideal electrical circuit such as a loop of conductor or a cylinder with zero resistance, the flux enclosed by it does not change when the flux density (magnetic field) or the area enclosed by the circuit is changed (Fig. 1). Analogously, the flux enclosed by a circuit is conserved if the current or the inductance of the circuit is changed. However, in nonideal conductors with nonzero resistance the current induced by the changing flux would decay with a time constant of  $L/R$  and the flux would change with the same time constant.

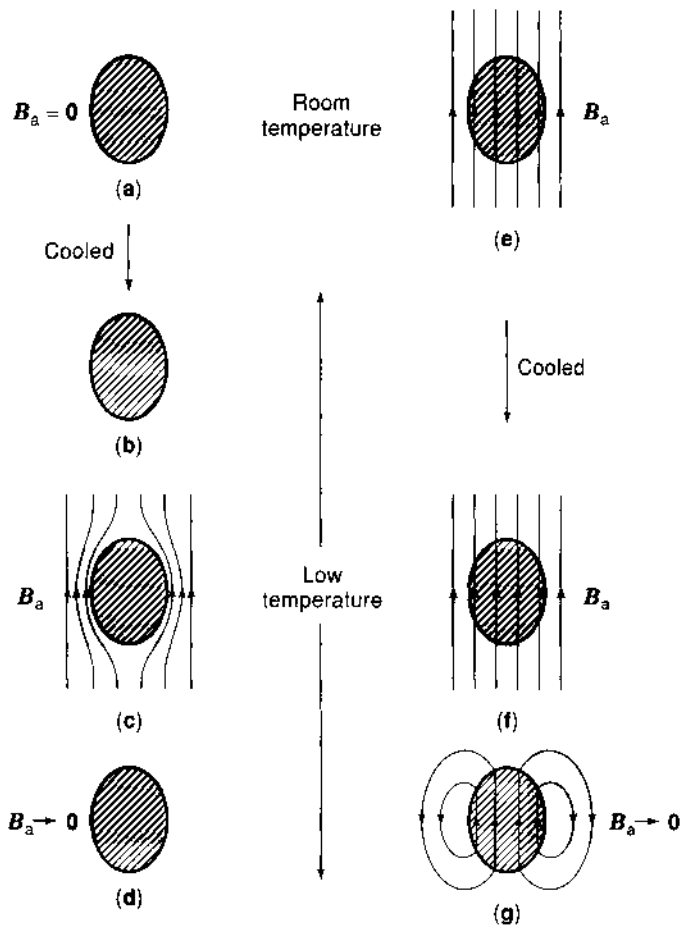
In mutually coupled circuits Eq. (2) is modified to include voltage induced by the mutually inductance. Therefore for two circuits  $p$  and  $q$ ,

$$V_p = L_p dI_p / dt + M_{pq} dI_q / dt + I_p R_p$$

and the same flux conservation concept would apply if the resistance is zero, that is,

$$L_p I_p + M_{pq} I_q + \Phi = \text{const when } R_p = 0$$

The currents induced for the conservation of flux are called diamagnetic (1) or eddy currents and in resistive conductors, such currents cause losses in the conductors when the field (flux) is changed.



**Figure 1.** Magnetic behavior of a "perfect" conductor. (a) and (b) Specimen becomes resistanceless in absence of field. (c) Magnetic field applied to resistanceless specimen. (d) Magnetic field removed. (e) and (f) Specimen becomes resistanceless in applied magnetic field. (g) Applied magnetic field removed (subscript 'a' refers to applied field) (Ref. 4).

It can be easily shown that a diamagnetic material placed inside a coil reduces the inductance of the coil (the total magnetic flux in the coil is reduced), while a paramagnetic and ferromagnetic material placed inside a coil increases its inductance.

### Poynting Flux

When the magnetic field is not constant in time, by Maxwells Law

$$\text{curl } \mathbf{E} = -d\mathbf{B}/dt$$

Therefore, an electric field  $\mathbf{E}$  is always associated with a time varying magnetic field. The medium therefore stores both electric and magnetic energy and this energy is, in general, time dependent and in addition, as the changing fields penetrate the volume, there may be energy dissipation in volume  $V$  due to resistive currents driven by the electric field at a rate given by

$$dE_{\text{dis}}/dt = \text{Int}[\mathbf{J} \cdot \mathbf{E}] dV$$

where  $\mathbf{J}$  is the current density in the volume. Maxwells Law can be written as

$$dE_{\text{dis}}/dt = \int [\mathbf{E} \cdot \text{curl } \mathbf{H} - \mathbf{E} \cdot d\mathbf{D}/dt] dV$$

where  $\mathbf{D}$  is the electric displacement vector. Using vector identity and assuming linear properties, this can be written as

$$dE_{\text{dis}}/dt = \int [\mathbf{J} \cdot \mathbf{E}] dV = - \int [du/dt + \text{div}(\mathbf{E} \times \mathbf{H})]$$

where now the energy density of the region,

$$u = (\mathbf{E} \cdot \mathbf{D} + \mathbf{B} \cdot \mathbf{H})/2$$

In differential form this leads to the energy conservation equation

$$du/dt + \text{div}(\mathbf{E} \times \mathbf{H}) + \mathbf{J} \cdot \mathbf{E} = 0$$

The first term represents the rate of change of the energy density, the term  $\mathbf{S} = (\mathbf{E} \times \mathbf{H})$  represents energy flow in or out of the volume and the last term represents the energy dissipation.  $\mathbf{S}$  is called the Poynting vector and is particularly relevant to electromagnetic fields and waves.

### Flux Penetration and Diffusion

As stated in the previous sections, if a magnetic field is applied to (or changed on) the exterior of a material (the flux enclosed by the area of the material is changed), the material gets an induced voltage that drives diamagnetic currents opposing the change in flux. If the material has a finite resistance, the currents will then decay and the flux will penetrate into the material. The flux will penetrate diffusively much like the diffusion of heat over time. The following relations illustrate this phenomenon.

Maxwells law gives

$$\text{curl } \mathbf{E} = -\partial\mathbf{B}/\partial t$$

$$\text{curl } \mathbf{H} = \mathbf{J} + \frac{\partial\mathbf{D}}{\partial t}$$

and Ohm's law gives

$$\mathbf{E} = \rho\mathbf{J}$$

where  $\mathbf{E}$  and  $\mathbf{J}$  are the induced electric field and (eddy) current density, respectively, and  $\rho$  is the resistivity of the material.

Therefore,

$$\text{curl}(\rho/\mu)\text{curl } \mathbf{B} = -\frac{\partial\mathbf{B}}{\partial t}$$

Since  $\text{div } \mathbf{B} = 0$  and for uniform resistivity  $\rho$ ,

$$(\rho/\mu)\nabla^2\mathbf{B} = \partial\mathbf{B}/\partial t$$

which is the equation for the diffusion of the flux into the material and  $\rho/\mu_0 = D_m$  is the magnetic diffusion coefficient. The flux and the field diffuse into the material thickness of  $l$  in a time given by  $l^2/D_m$ .

If the conducting material is in motion such as a moving plasma, an additional induced electric (convective) field  $\mathbf{v} \times \mathbf{B}$  is present. The net time rate of change of the magnetic field is given by

$$d\mathbf{B}/dt = (\rho/\mu)\nabla^2\mathbf{B} + \text{curl}(\mathbf{v} \times \mathbf{B})$$

and in analogy with viscous flow, a “magnetic” Reynolds number can be defined as

$$R_M = Lv/D_m$$

where  $L$  is the characteristic dimension of the flow. The magnetic Reynold’s number can vary from a value much less than 1 for laboratory devices to values on the order of 100 for fusion plasmas, while for geophysical or astronomical conditions,  $R_m$  can be as high as  $10^6$  to  $10^{10}$ . Therefore the flux can be diffused by the flow as it penetrates the conducting material. This convective flow can be a mechanism for converting one type of flux into another (see the section entitled “Flux Conversion”).

## FLUX LINE AND FLUX TUBE

The flux lines are directed lines of force (LOF) and lie in the direction that a north (mono-) pole would point to when placed in the magnetic field. The LOF is defined by the equation

$$dx/B_x = dy/B_y = dz/B_z$$

where  $B_x$ ,  $B_y$ , and  $B_z$  are components of the flux density in the directions  $x$ ,  $y$ , and  $z$ . The equation may be integrated to give surfaces of the type (2)

$$\begin{aligned} f(x, y, z) &= a \\ g(x, y, z) &= b \end{aligned}$$

the intersection of which gives a specific line of force. In this case, the local unit vector of the LOF is given by

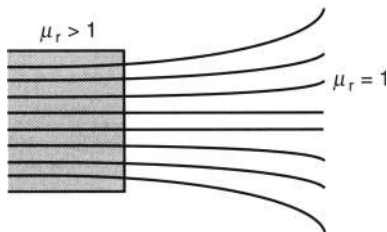
$$\mathbf{k} = \text{grad } f \text{ grad } g / (|\text{grad } f \text{ grad } g|^2)^{1/2}$$

which involves components of tensorial products. A tube of force is a collection or a group of lines of force. Since  $\text{div } \mathbf{B} = 0$ , the flux in a tube is conserved as the lines of force diverge and converge. If a tube branches into a number of tubes, then the sum of fluxes remains the same.

Since the flux in a tube of force is conserved, the cross section of the tube of force traversing through materials of different permeability would be inversely proportional to the permeability (Fig. 2); however, continuity equations require this variation in cross section to be gradual.

A useful concept is the specific volume of a magnetic tube of force given by

$$U = dV/d\Phi$$



**Figure 2.** Spreading of field lines (LOF) in low permeability region for the same flux.

where now  $V$  is the volume of the tube and  $\Phi$  is the flux enclosed by it.

Since the flux in a tube is conserved,

$$U = \iint (dS dl) / (B dS)$$

or

$$U = \int (dl/B)$$

where  $dS$  and  $dl$  are the cross sectional area and length of a volume element,  $B$  is the flux density, and the integral is over the whole tube.

## FLUX AND FIELDS IN A TOROIDAL GEOMETRY

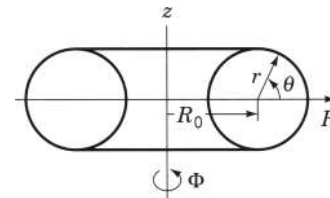
The toroidal geometry (Fig. 3) has applications especially to plasma-confinement devices, and the topology of the field and the constituent magnetic flux are of specific interest in such devices and astrophysics. In a toroidal geometry, a pure toroidal field (field lines going around the major circumference) or a pure poloidal field (field lines going around the minor circumference) would give closed field lines. In most plasmas of interest, both fields would be present and the toroidal and poloidal fluxes are also called longitudinal and azimuthal fluxes. The toroidal or azimuthal flux  $\chi$  is the flux enclosed by the surface  $\phi = \text{const}$ , where  $\phi$  is the azimuthal angle around the major axis of the toroid. The poloidal or the longitudinal flux  $\Phi$  is the flux enclosed by the surface  $\theta = \text{const}$  (Fig. 4). If both fields are present, the lines of force go around helical paths around the torus. In general, a line of force starting at a certain poloidal angle will arrive at a different poloidal angle after one traverse or more around the major circumference. The rotational transform is defined as the change in angle averaged over a large number of transits around the major circumference:

$$t = \lim_{n \rightarrow \infty} \sum_{k=1}^n \iota_k / n$$

For a toroid with a toroidal current of  $I_\phi$  and a uniform toroidal field of  $B_0$ , the rotational transform at the minor radius  $r$  is given by

$$\begin{aligned} \iota &= B_\phi(r) 2\pi R / r B_0 \\ &= 2\pi d\chi / d\Phi \end{aligned}$$

where  $B_\phi(r) = \mu_0 I_\phi / 2\pi r$  is the poloidal field at minor radius  $r$  and  $R$  is the major radius. The quantity  $q = 2\pi / \iota$  is known as the factor of safety in fusion-device terminology.



**Figure 3.** Toroidal geometry,  $R_0$  is the major radius,  $\Phi$  the azimuthal angle,  $\gamma$  the minor radius, and  $\theta$  the poloidal angle.



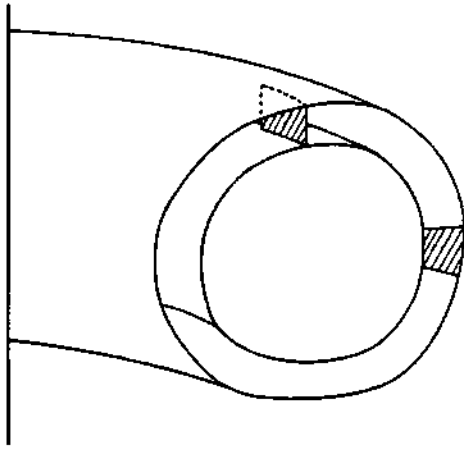


Figure 4. Illustration of the poloidal and toroidal surface elements  $dS_p$  and  $dS_t$ .

The surface on which the helical lines that close on themselves after a number of transits is called a rational surface.

### MAGNETOMOTIVE FORCE AND RELUCTANCE

These terms are defined analogously to electrical circuits. A magnetic circuit consists of flux threading the circuit, analogous to current. The flux is "driven" by the magnetomotive force (mmf)  $E_M$ , and the flux  $\Phi$  is limited by the reluctance, so that reluctance is analogous to resistance in an electrical circuit. In most applications the mmf can be defined as ampere-turns, that is, the product of the instantaneous current and the number of turns. The reluctance of a circuit element is given by

$$\mathcal{R} = E_M / \Phi$$

The reluctance of an element is related to the characteristics of the element by

$$\mathcal{R} = L / \mu_r \mu_0 A$$

where  $\mu_r$  and  $\mu_0$  are the relative permeability of the circuit element material and the permeability of vacuum, respectively, and  $L$  and  $A$  are the length and cross-sectional area of the circuit element. Therefore materials with high permeability such as iron have low reluctance, and vacuum or air has high reluctance. The concept of reluctance can be used in magnetic circuits analogous to electric circuits. If the mmf is analogous to the emf (applied voltage), then flux is analogous to resistive current, and the reluctance is analogous to electrical resistance (with permeability being equivalent to electrical conductivity). For example, the flux generated by a coil with ampere turns  $NI$  and threading two adjacent (in series) volumes with reluctances  $R_1$  and  $R_2$  is given (in one dimensional approximation) by

$$\Phi = NI / (R_1 + R_2)$$

The magnetization in a material is given by

$$M = B - \mu_0 H = (\mu_r - 1) \mu_0 H$$

Thus low-reluctance materials also have high magnetization. For  $\mu_r \gg 1$ ,

$$M \sim H(L/\mathcal{R}A)$$

### Ferromagnetic Materials and Shielding

Ferromagnetic materials (e.g., iron) have high permeability and therefore low reluctance. Therefore in magnetic devices, where the flux is to be linked effectively between two electrical circuits, e.g., transformers and motors, a ferromagnetic path is usually employed. Conversely, to shield regions from magnetic fields, a low reluctance magnetic path may be provided for the field so that the field lines prefer to pass through the ferromagnetic region rather than the region that has to be shielded. Such a shielding iron may cover the source or cover the region to be shielded.

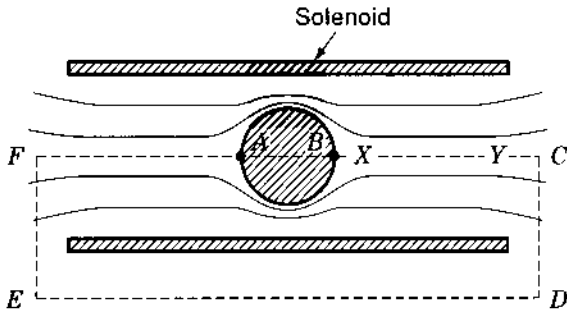
### DEMAGNETIZATION FACTOR

While the flux inside a perfect conductor is conserved, it must be remembered that the magnetic field intensity  $H$  is not. In fact, the field intensity inside the diamagnetic material can be shown to increase by a factor depending upon the geometry of the material, for the same conditions of field excitation, for example, magnetic current. (For paramagnetic and ferromagnetic materials, the field intensity decreases by some factor.) The demagnetization factor is 2 for a sphere and is 1.5 for a cylindrical cross section. This fact can be explained as follows. (For an alternate description, see Ref. 3 on the analogous characteristic of depolarizing factor.)

Consider a long solenoidal magnet (4) that produces a nearly uniform external field intensity  $H_e$  in the direction  $x$  at the center of the solenoid. Now if a sphere of diamagnetic material is placed at the center and the solenoidal field is established, the diamagnetic material will exclude this flux from inside and it can be shown (solution to the Laplace's equation) that the field lines will be as shown in Fig. 5. While the field intensity  $H_e$  outside the sphere is unaffected far away from the sphere, the field intensity  $H_i$  is zero just outside of the sphere along the diameter parallel to the field direction, and near that region the external field intensity  $H_e$  will be less than the value in the absence of the sphere. Therefore the field intensity  $H_e$  outside the sphere will be less than or equal to the field intensity  $H_e$  in the absence of the sphere. Now, the  $\int \mathbf{H} \cdot d\mathbf{l}$  along the closed path  $ABCDEF$  gives (by Ampere's law) the total ampere turns in the solenoid, which was held constant when the sphere was placed. Therefore

$$\begin{aligned} & \int (\mathbf{H}_i \cdot d\mathbf{l})_{AB} + \int (\mathbf{H}_e \cdot d\mathbf{l})_{BC} + \int (\mathbf{H}_e \cdot d\mathbf{l})_{CDEF} + \int (\mathbf{H}_e \cdot d\mathbf{l})_{FA} \\ &= \int (\mathbf{H}_i \cdot d\mathbf{l})_{AB} + \int (\mathbf{H}'_e \cdot d\mathbf{l})_{BC} + \int (\mathbf{H}'_e \cdot d\mathbf{l})_{CDEF} + \int (\mathbf{H}'_e \cdot d\mathbf{l})_{FA} \end{aligned}$$

Since,  $H'_e$  is less than or equal to  $H_e$  along  $BC$  and  $FA$  and the integral over  $CDEF$  is unaffected (too far away),  $\int (\mathbf{H}'_i \cdot d\mathbf{l})_{AB}$  is larger than  $\int (\mathbf{H}_i \cdot d\mathbf{l})_{AB}$  to preserve the sum of the integrals,



**Figure 5.** A diamagnetic sphere in a solenoid. The field strength at a point close to the sphere, such as *X*, is less than it would be if the sphere were absent, while the field strength at a point far away, such as *Y*, is essentially unchanged. The line integral of *H* around the broken line is independent of whether the sphere is present or not, so the field strength inside the sphere must exceed the applied field  $H_a$  (Ref. 4).

which essentially means that  $H_i$  is larger than  $H_e$  along *CD*, that is, inside the sphere. This effect was first noted for paramagnetic and ferromagnetic materials for which the relative permeability is greater than 1, and therefore the field intensity inside the sphere would be less than the field in the absence of the sphere. Since this effect was first noted in such materials, the effect is considered to be a demagnetization and can be stated as

$$H_i = H_a - nI_m$$

where  $H_i$  is the field intensity inside the object with magnetization  $I_m$  and  $H_a$  is the applied field intensity. The quantity  $n$  is called the demagnetization factor and depends on the geometry of the object. For spheres  $n = \frac{1}{3}$  and for cylinders  $n = \frac{1}{2}$ . Clearly if the applied field is perpendicular to a thin cylindrical wire, because of the volume average,  $n = 0$ . It must be remembered that for paramagnetic and ferromagnetic materials,  $I_m$  is positive and the field intensity inside the material decreases, while in diamagnetic materials,  $I_m$  is negative and the field intensity inside the material increases. This demagnetization factor has important consequences for nonlinear magnetization and critical characteristics of materials such as iron and superconductors.

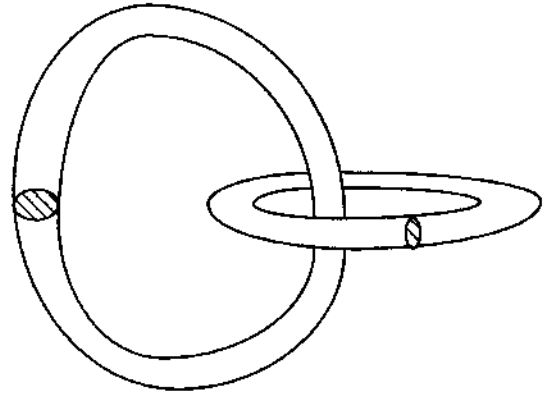
**MAGNETIC HELICITY**

The topology of magnetic surfaces and the complexity of the structure of the magnetic field can be described by a quantity known as magnetic helicity, which is defined as

$$H = \int (\mathbf{B} \cdot \mathbf{A} dV)$$

where the integral is over the volume of interest. The magnetic helicity describes the linking of field lines and tubes of force. Considering the two linked tubes in Fig. 6, the helicity can be written as

$$H = \iint [d\mathbf{S} \cdot d\mathbf{l}(\mathbf{B} \cdot \mathbf{A})]$$



**Figure 6.** Linkage of flux tubes, tubes with cross sections 1 and 2 are threaded by flux  $\Phi_1$  and  $\Phi_2$  respectively.

where *S* is the cross section of tube 1 and *l* is the length of tube 1. Since *B* is approximately normal to *S*, this can be written as

$$H = \oint (\mathbf{A} \cdot d\mathbf{l}) \int (\mathbf{B} \cdot d\mathbf{S})$$

The surface integral is the flux  $\Phi_1$  enclosed by tube 1 while the line integral is the flux  $\Phi_2$  enclosed by tube 2. The linked system helicity is then equal to the sum of the two helicities equal to  $2\Phi_1\Phi_2$ . If the tubes are linked *N* times the helicity will be equal to  $2N\Phi_1\Phi_2$ . In many systems of interest, the helicity is conserved as the magnetic configuration evolves.

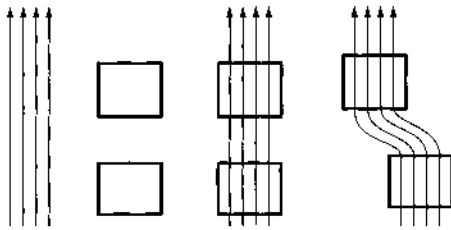
**Flux Conversion**

If the field configuration is confined in a closed, perfectly conducting and nonpermeable surface (the normal component of **B** and velocity of any conducting medium *v* are zero), then helicity is conserved. This means that in such a configuration, although individual fluxes of different tubes (components of flux density), for example, toroidal and poloidal fluxes, are not conserved independently, the product is conserved. The symmetry then permits conversion of one type of flux into another. Such flux conversions are observed in plasma devices and in geomagnetic phenomena (5-7). The presence of turbulent structures and coherent magnetic field fluctuations may provide a mechanism for the conversion of flux and geomagnetic phenomena (8).

A simple generation of flux conversion is illustrated by using the diffusion time for flux lines in a good conductor. Consider a magnetic field applied externally to a pair of conducting materials [Fig. 7(a)]. After a certain time, depending on the conductivity of the material, the flux will diffuse in the two conductors [Fig. 7(b)]. Now, if one of the conductors is moved fast compared to the diffusion time, the flux lines will be bent and appear as shown in Fig. 7(c) until the lines can redistribute themselves inside the conductor. It is clear that in this process, a portion of the magnetic field that was previously in the vertical direction has been converted into a horizontal field.

**Dynamo Action and Geomagnetism**

The fact that convective motion of conducting fluids can generate magnetic fields has been invoked in explaining spontaneous flux generation from seed magnetic flux and is considered to be the source of the dynamo action in the earth's core, which produces magnetic fields. In perfectly conducting fluids, the lines



**Figure 7.** Conversion of vertical field to horizontal field. (a) Field before penetrating two blocks of conductor. (b) Fields after penetration. (c) Fields after the lower conductor is moved—a horizontal component is created in the gap between the two conductors.

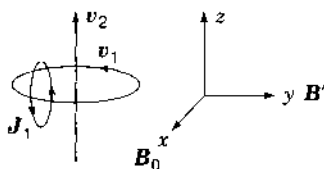
of force are frozen (see the section on plasma equilibrium and Ref. 9). While the earth's core is conducting, any generated magnetic field must have short decay time due to the finite resistivity of the earth's melted core. Therefore a continuous dynamo action is necessary to maintain this field. Such a dynamo action is caused by the correlation between velocity and field perturbations in the turbulent motion of the core (10,11). Two effects, the  $\alpha$  effect and the  $\omega$  effect, are invoked to explain the dynamo action (12).

The  $\alpha$  effect is a direct result of the Faraday effect. Consider Ohm's law,

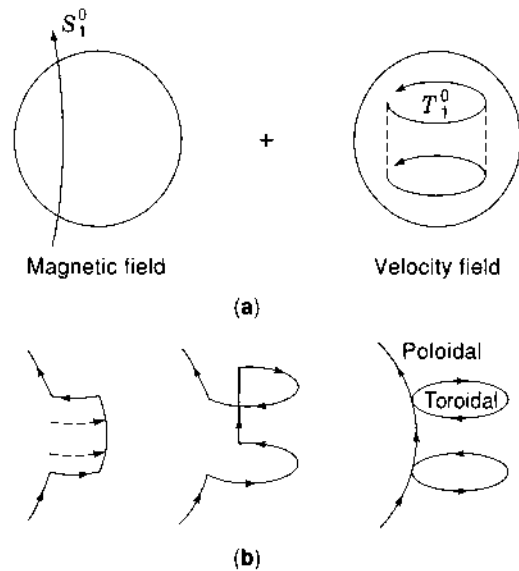
$$\mathbf{J} = \sigma \mathbf{E} + \sigma(\mathbf{v} \times \mathbf{B})$$

where the second term is due to the induced Faraday emf. If a turbulent system is present, so that  $\mathbf{v} = \mathbf{v}_0 + \mathbf{v}'$  and  $\mathbf{B} = \mathbf{B}_0 + \mathbf{B}'$ , the average induced electric field  $\mathbf{E} = \mathbf{v}_0 \times \mathbf{B}_0 + \mathbf{v}' \times \mathbf{B}'$ , since the averages of  $\mathbf{v}'$  and  $\mathbf{B}'$  are zero in turbulent perturbations. Therefore an additional emf  $\mathbf{E}' = \mathbf{v}' \times \mathbf{B}'$  associated with the correlated velocity and magnetic fields occurs. In specific systems, this electric field can be written as  $\mathbf{E}' = \alpha \mathbf{B}_0$ , where  $\alpha$  is a tensor in general. This electric field then has a component of current which maintains the dynamo action. Consider Fig. 8, where the turbulent velocity of the fluid can be resolved into an axial component  $v_2$  and a rotational component  $v_1$ . If the initial magnetic field is in the  $x$  direction, the  $v_1$  component will produce an electric field  $\mathbf{v}_1 \mathbf{B}_0$  and a current  $\mathbf{J}_1$  perpendicular to the  $y$  axis. This current will then produce a magnetic field  $\mathbf{B}'$  in the  $y$  direction. An electric field  $\mathbf{E}' = v_2 \mathbf{B}'$  will be produced in the  $x$  direction (parallel to the original magnetic field), as stated previously.

The  $\omega$  effect is caused by convective effects illustrated in the preceding section. In the illustration shown in Fig. 9, a radial or poloidal field is convected by a toroidal flow. When a toroidal



**Figure 8.** Illustration of the  $\alpha$  effect. Consider a right-handed helical velocity field depicted by  $v_1$  and  $v_2$  in the presence of a field  $\mathbf{B}_0$  aligned along the  $x$  axis. This will produce current loops such as  $\mathbf{J}_1$ , lying in the  $x$ - $z$  plane. Associated with the current loop  $\mathbf{J}_1$ , is a field  $\mathbf{B}'$  aligned parallel to the  $y$  axis. This new field  $\mathbf{B}'$  interacts with  $v_2$  to produce an electric field parallel to the  $x$  axis.



**Figure 9.** Production of a toroidal magnetic field in the core. (a) An initial poloidal magnetic field passing through the Earth's core is shown on the left, and an initial cylindrical shear velocity field,  $T_1^0$ , is shown on the right. (b) The interaction between the velocity and the magnetic field in (a) is shown at three successive times moving from left to right. The velocity field is only shown on the left by dotted lines. After one complete circuit two new toroidal magnetic field loops of opposite sign ( $T_2^0$ ) have been produced. After Ref. 9.

flow is impressed upon a poloidal field, the velocity field shears the magnetic field and produces a toroidal magnetic field such that the direction at the top and the bottom are opposite, preserving helicity (13). The dynamo is again due to the correlation between the turbulent velocity field and the turbulent magnetic field.

## ELECTRICAL MACHINES

### Transformers

Transformers essentially use Lenz's law. In transformers a "primary" coil is supplied with a time varying current and a "secondary" coil mutually coupled to the primary coil receives an induced voltage that can then be used to drive a current into another circuit (Fig. 10). This then permits isolating the secondary circuit electrically from the primary circuit while enabling the indirect use of the source that powers the primary circuit. In addition, the transformer permits the "stepping up or down" of the voltage, that is, the secondary voltage can be larger or smaller than the primary voltage by the ratio of the number of turns in the primary and secondary coils. In a transformer the secondary coil is made to link nearly all the flux due to the primary coil by placing the primary and secondary coils around iron, which provides a closed low-reluctance path for the magnetic flux. The changing current in the primary coil causes a change in the flux and the secondary coil receives an induced voltage that opposes this changing flux.

$$V_p = N_p d\Phi/dt$$

$$V_s = -N_s d\Phi/dt$$

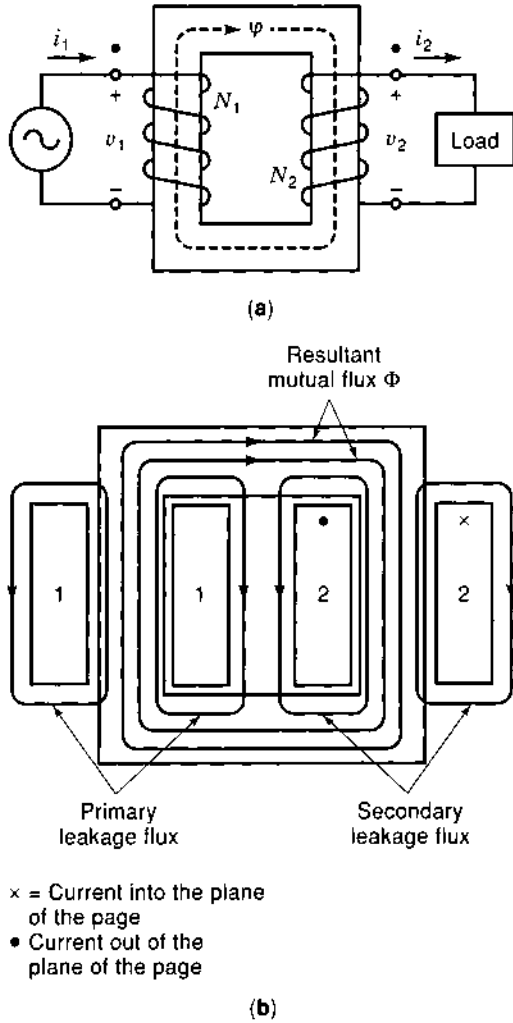


Figure 10. (a) Ideal transformer and load. (b) Component fluxes.

The negative sign indicates that the secondary coil opposes the change in flux caused by the primary coil

$$V_s/V_p = -N_s/N_p$$

Since the same flux is linked and the flux path is the same (the reluctance and the flux are equal in the primary and secondary circuits),

$$N_s I_s = N_p I_p \quad I_s/I_p = N_p/N_s$$

or

$$V_s I_s = V_p I_p$$

However in a nonideal transformer (14), part of the voltage applied to the primary coil is expended in generating the flux in the core and part is expended for compensating for eddy currents in the coil and iron and losses in iron due to hysteresis. The flux generated in the core by the primary current links the secondary current as a mutual flux and the remaining current leaks out into the air (which is outside the iron core and there is no linkage with the secondary current) as leakage flux. Similarly the flux due to current in the secondary coil (under load conditions) also has two parts: mutual and leakage flux. As is

evident from the terminology, the mutual flux of the primary and secondary coils are equal and the leakage flux is dependent on the core size and permeability—the larger the area and permeability, the smaller the fraction of leakage flux.

The mutual flux is subject to saturation effects in iron. The reduced permeability of the iron at high excitation currents (flux densities) causes a smaller increase in mutual flux for an increase in the current, and the induced voltage exhibits saturation. This increases the high-harmonic components in the secondary voltage. Since the leakage flux is in the air, it is proportional to the current.

The total primary and secondary flux can be written as

$$\Phi_{tp} = N_p \Phi_{lp} + N_p \Phi_m$$

$$\Phi_{ts} = N_s \Phi_{ls} + N_s \Phi_m$$

where subscript l refers to the leakage flux and m refers to the mutual flux. Similarly the total voltage is also the sum of that induced by (or used in creating) the leakage flux and that induced by the mutual flux. The leakage flux

$$N_p \Phi_{lp} = N_p^2 I_p / \mathcal{R}_p$$

$$N_s \Phi_{ls} = N_s^2 I_s / \mathcal{R}_s$$

where \$\mathcal{R}\_p\$ and \$\mathcal{R}\_s\$ are the reluctances of the primary and secondary leakage paths, respectively. The corresponding voltages associated with the leakage paths can be defined as

$$V_{lp} = L_{lp} dI_p/dt$$

$$V_{ls} = -L_{ls} dI_s/dt$$

where \$L\_{lp} = N\_p^2/\mathcal{R}\_p\$ and \$L\_{ls} = N\_s^2/\mathcal{R}\_s\$ are the leakage inductances. Taking into account the resistance of the coils, \$r\_p\$ and \$r\_s\$, the total voltage is then given by

$$V_p = E_p + L_{lp} dI_p/dt + I_p r_p$$

$$V_s = E_s - L_{ls} dI_s/dt - I_s r_s$$

where \$E\_p\$ is the voltage inducing the mutual flux in the primary and \$E\_s\$ is the voltage induced by the mutual flux. In well-designed transformers, the leakage and resistive terms are usually negligible.

Since

$$V_p = N_p d\Phi_{tp}/dt$$

$$\Phi_{tp} = (1/N_p) \int (V_p dt)$$

For a sinusoidal voltage with a frequency \$f = \omega/2\pi\$, \$V\_p = V\_0 \sin(\omega t + \alpha)\$

$$\Phi_{tp} = (V_0/\omega N_p) \cos(\omega t + \alpha) + \Phi_c$$

where \$\Phi\_c\$ is a transient flux that decays after switching on due to eddy currents and hysteresis losses. Therefore the flux induced in the transformer is inversely proportional to the frequency of the applied voltage and lags in phase angle by \$\pi/2\$.

An approximate equivalent circuit of the transformer can be constructed in a single circuit taking into account the mutual coupling, where the circuit consists of primary inductance and

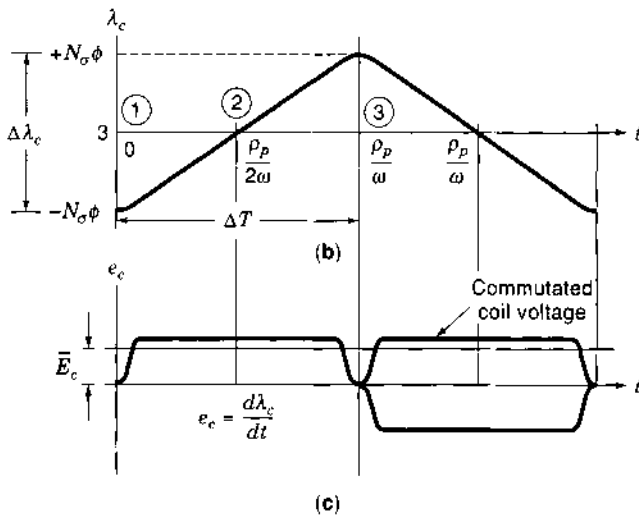
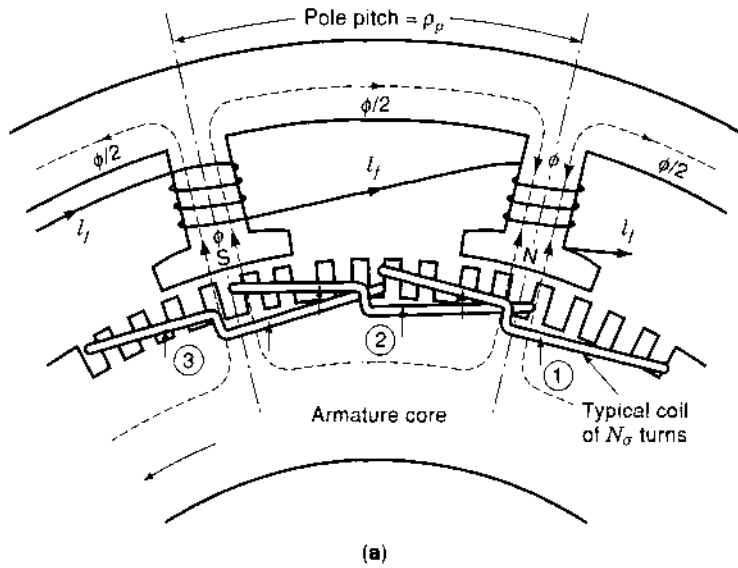


Figure 11. Schematic of a generator.

resistance, the mutual coupling inductance and magnetization, the leakage flux, and the secondary impedance (inductance and resistance referred to the primary). Other nonideal effects such as saturation of the iron core, ac losses in the core, and eddy currents can be taken into account in such a circuit (15).

**Dc Electric Generators**

In a generator (Fig. 11 from Ref. 14), a coil of conductors on the armature (rotor) moves across the north and south magnetic poles (stator). If the coil has  $N_c$  turns and the poles generate a flux  $\Phi$ , the coil will link a flux  $\Phi$  under the north pole, then zero flux between poles, and a flux  $-\Phi$  under the south pole. Therefore the voltage induced in the coil is

$$V_c = \Delta\Phi / \Delta t$$

where  $\Delta\Phi = 2N_c\Phi$  is the change in the flux seen by the coil and  $\Delta t$  is the time over which the flux change occurs. If the coil is rotating at a rate of  $n$  rotations per second and there are  $p$

poles in the stator, then  $\Delta t = 1/np$ , so that

$$V_c = 2N_cpn\Phi$$

If  $C$  coils are connected in series and  $a$  are connected in parallel, the generated voltage is

$$V_g = 2CN_cpn\Phi/a = K_a\Phi\omega$$

where  $\omega$  is the angular frequency of rotation and  $K_a = CN_c/\pi ap$  is known as the armature constant. As shown later, the voltage induced is alternating, and dc generators require so called commutators to change the brush polarity alternately to generate dc voltage.

**Generalized ac Machines**

**Induced Voltage in an ac Generator.** In ac motors and generators, a number of multipole excitation coils are placed in a stationary high permeability core and a set of secondary coils are placed in a rotary core. The secondary and the primary excitation coils are placed around a common axis and have a

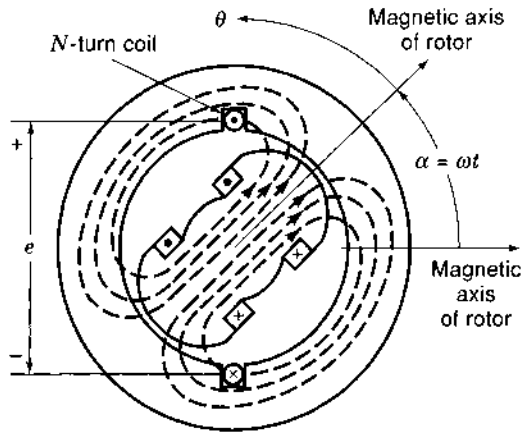


Figure 12. Elementary two-pole ac machine with stator coil of  $N$  turns.

small gap. (A simple example of a two-pole ac machine is shown in Fig. 12.) At an arbitrary angle  $\theta$  between the rotor and the stator with  $N$  turns, the flux linked by the stator is

$$N\Phi = N \int_{-\pi/2}^{\pi/2} [B_{\max}(\cos \theta)l r d\theta] = 2NB_{\max}lr$$

where  $l$  is the length of the rotor (normal to the figure) and  $r$  is the radius of the stator at the gap. For  $p$  poles, the flux is  $(2/p)(2NB_{\max}lr)$ . If the rotor spins with an angular velocity  $\omega$ , the flux links changes with time as

$$\Phi' = N\Phi \cos(\omega t)$$

The voltage induced due to the time variation of the flux is given by

$$e = -d\Phi'/dt = \omega N\Phi \sin(\omega t) - N \cos(\omega t) d\Phi/dt$$

If the flux produced by the coils is independent of time, the second term is zero, but it is clear that the generated voltage is alternating.

**Rotating Magnetic Field.** In three-phase ac machines (where three legs of the ac supply each have a phase difference of  $120^\circ$ ), three sets of stator coils are connected to the three phases (Fig. 13), so that the currents in the coils are given by

$$\begin{aligned} I_a &= I_{\max} \cos(\omega t) \\ I_b &= I_{\max} \cos(\omega t - \pi/3) \\ I_c &= I_{\max} \cos(\omega t - 2\pi/3) \end{aligned}$$

where  $\omega = 2\pi f$  and  $f$  is the frequency of the ac supply. In such a case, the total instantaneous force on the armature at an arbitrary angle  $\theta$  due to the three coils is proportional to the flux linked, which, in turn, is proportional to the current and is given by

$$\begin{aligned} \Phi(\theta, t) &= \Phi_{\max} \cos \theta \cos(\omega t) + \Phi_{\max} \cos \theta \cos(\omega t - \pi/3) \\ &\quad + \Phi_{\max} \cos \theta \cos(\omega t - 2\pi/3) \\ &= 1.5\Phi_{\max} \cos(\theta - \omega t) \end{aligned}$$

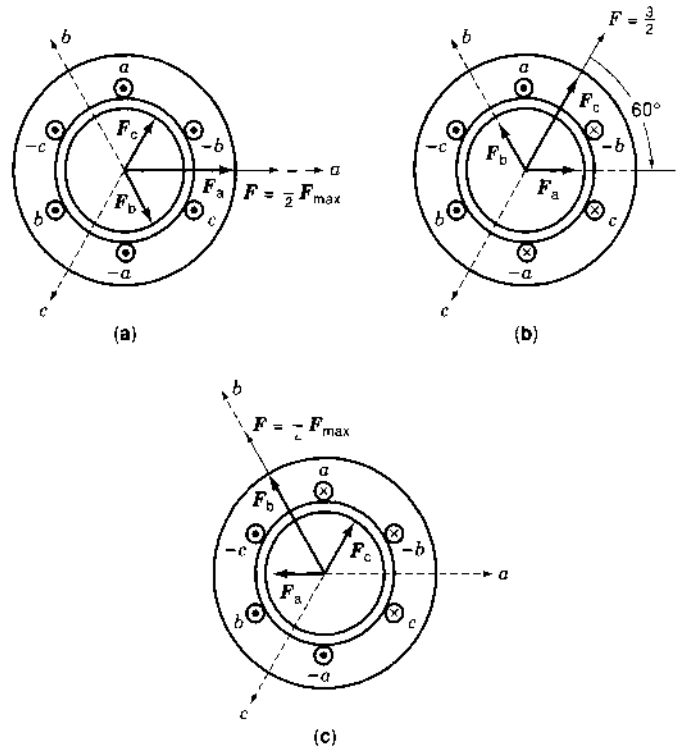


Figure 13. Production of a rotating magnetic field by means of three currents.

which represents a traveling wave of flux (also an mmf or force in a motor, or induced emf in a generator). If at  $t = 0$  the peak of the flux was at  $\theta = \theta_0$ , then in a time  $t_0$  the peak moves to  $\theta = \theta_0 - \omega t$  and therefore the field appears to rotate in time. Figure 13 shows this rotation, where  $F$  is the force (proportional to the flux linked) experienced by the armature.

If the armature also rotates but with an angular velocity  $\omega_a$ , the linkage is given by

$$\Phi(t) = 1.5\Phi_{\max} \cos(\omega_a - \omega)t$$

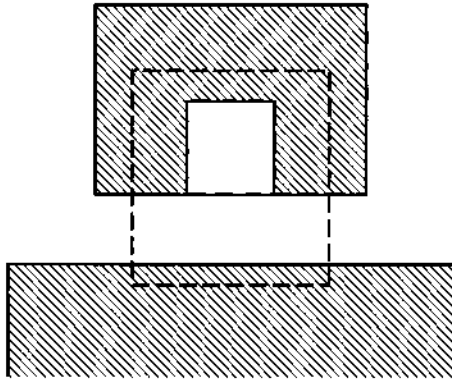
so that when  $\omega_a = \omega$ , the linked flux appears to be a constant and the motor or generator is *synchronous*.

This description can also be represented by a coupled-circuit description (14) using the stator and rotor inductance and flux linkages and resistance of the coils. The circuit description then leads to a set of two differential equations with time derivatives of current. Solutions of these equations give the instantaneous values of current and magnetic energy in the machine. The derivative of this energy with respect to the mechanical angle gives the torque produced.

**Sinusoidally Wound Stators.** The windings are arranged in such a fashion that the number of turns in the excitation and primary coils is a sinusoidal function of the angle, that is

$$N_i = N_{i0} \sin(p\phi) = N_{i0} \sin(p\pi x/L)$$

where  $i$  refers to the excitation or the secondary coil,  $N_0$  is the maximum number of turns,  $p$  is the number of pairs of poles,  $x$  is the position along the circumference, and  $L$  is total circumferential length. [If the coils are not arranged in a sinusoidal



**Figure 14.** Flux in the gap between the stator and the rotor. Most of the reluctance is in the gap since the stator and the rotor have high permeability iron path.

fashion and are as shown in Fig. 12, then the fundamental component is given by  $N_i = (2/\pi)N_{i0} \sin(p\phi/2)$ .

Now if the field in the gap is  $H_g$ , the integral of the field around a closed loop enclosing a coil (see Fig. 14) has two legs of the loop in the iron core that contribute negligibly if the permeability is very high and has two legs that cross the gap. Because the field direction remains along the integration direction, these add and the integral

$$\int (\mathbf{H} \cdot d\mathbf{x}) = 2H_g g$$

where  $g$  is the gap. But, by Ampere's law, the integral is also equal to the total current (ampere turns enclosed),  $IN_0 \sin(p\pi x/L)$ . Therefore

$$H_g = (IN_0/2g) \sin(p\pi x/L)$$

Now the flux coupled to the coil,

$$\Phi = l \int (\mathbf{B} \cdot d\mathbf{x}) = (\mu_0 l IN_0/2g) \int [\sin(p\pi x/L) dx]$$

where  $l$  is the length of the coil (or the area under consideration). Over one length of the pole,  $p/L$ , the integral gives

$$\Phi_0 = (\mu_0 l IN_0 L/2gp\pi) IN_0 G_g$$

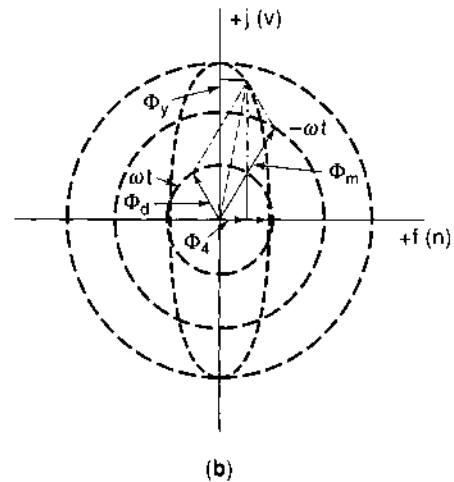
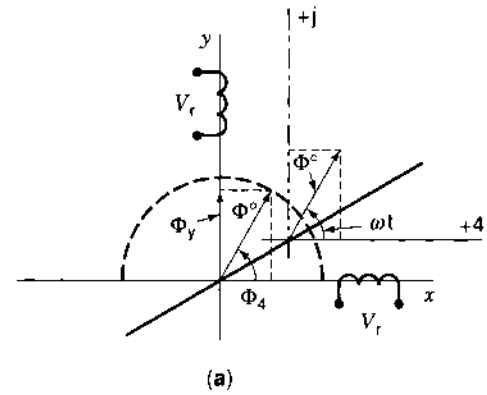
where  $G_g = (\mu_0 l L/2gp\pi)$  is the gap permeance per pole (or is the inverse of the reluctance of the gap per pole).

Now, if the current is alternating and the rotor is in motion, at any instance the flux coupled to the coil is given by

$$\begin{aligned} \Phi &= I_0 N_0 G_g \sin(\omega t + f) \\ &= IN_0 G_g \sin(\omega t) \cos \phi + IN_0 G_g \cos(\omega t) \sin \phi \\ &= \Phi_x \sin(\omega t) + \Phi_y \cos(\omega t) \end{aligned}$$

where  $\phi = p\pi x/L$  represents the angle or the spatial phase of the rotor at time  $t$ .

If we use the designations  $j$  and  $J$  to distinguish between and separately account for the rotational time dependence and



**Figure 15.** (a) Vector diagrams to illustrate spatial flux vector and (b) time vector diagram to illustrate geometrical meaning of the symmetrical component decomposition.

the angular position

$$\sin \phi = (e^{j\phi} - e^{-j\phi})/2j$$

$$\sin(\omega t) = (e^{j\omega t} - e^{-j\omega t})/2j$$

The above expression can be rewritten as

$$\Phi = \Phi_{td} + \Phi_{top} = (j\Phi_y - j\Phi_x)e^{j\omega t} + (j\Phi_y + j\Phi_x)e^{-j\omega t}$$

which represents two counterrotating components of flux, one direct and another opposite, that is, two components of flux with a phase difference of  $\pi$ . The two components are illustrated in Fig. 15.

The preceding description of the flux is useful in the design of devices such as sine-cosine transformers (SCT), remote and point control systems, tachogenerators, and servomotors.

## CHARGED-PARTICLE MOTION IN A MAGNETIC FIELD

A charged particle is deflected from its original path by a magnetic field if it has a velocity component perpendicular to the magnetic field (that is, charged particles with velocity in the direction of the magnetic field do not experience a force). The particle moves in a direction perpendicular both to the initial velocity and the magnetic field. Since the motion is perpendicular

to the magnetic field, no work is done by the magnetic field and the particle energy does not change. It can be seen then that the particle exercises circular motion around the field direction (flux lines), and if the particle has a parallel velocity (which remains unaffected by the field), the particle executes spiral motion. The radius of the circular motion is called the Larmor or gyro radius and the rotational frequency is called the Larmor or gyro frequency.

It can be shown that if the field varies slowly in space and in time, the flux enclosed by the charged particle is constant. This conservation of flux is true in an "adiabatic" sense and leads to other adiabatic constants of motion, which enable the development of magnetic traps for plasmas and particle beams as well as particle accelerators and particle detectors.

The equation of motion of the charged particle in an electric and a magnetic field is given by

$$d\mathbf{p}/dt = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$$

where  $\mathbf{p} = \gamma m \mathbf{v}$  and  $m$ ,  $\mathbf{v}$ ,  $\gamma$ , and  $q$  are the particle momentum, mass, velocity, relativistic factor and charge, respectively, and  $\mathbf{E}$  and  $\mathbf{B}$  are the electric and magnetic fields. The equation can be written in component form for  $\mathbf{E} = 0$  as (for simplicity shown only for the magnetic field in the  $z$  direction, i.e.,  $\mathbf{B} = B\mathbf{e}_z$ ).

$$\begin{aligned} dv_x/dt &= (qB/\gamma m)v_y \\ dv_y/dt &= -(qB/\gamma m)v_x \\ dv_z/dt &= 0 \end{aligned}$$

which are the equations for circular or spiral orbits with the gyro frequency  $\Omega_g = qB/\gamma m$ . Solving the equations for displacements, one gets

$$\begin{aligned} x &= (v_p/\Omega_g) \sin(\Omega_g t) \\ y &= \pm (v_p/\Omega_g) \cos(\Omega_g t) \end{aligned}$$

where the + or - sign corresponds to the positive or negative charge (which may be dropped if the gyrofrequency includes the sign of the charge) and  $v_p = (v_x^2 + v_y^2)^{1/2}$ . Therefore  $r_g = v_p/\Omega_g$  is the Larmor radius. (For a full relativistic treatment of the charged particle motion, see Ref. 16.)

### Motion in a Time-Varying Field

If a particle is performing gyro orbits in a time-varying magnetic field, the energy is not a constant, since there is an associated electric field given by

$$\text{curl } \mathbf{B} = -\partial \mathbf{E}/\partial t$$

The energy gain is given by (2)

$$\Delta U_p = \int (q d\mathbf{r} \cdot \mathbf{E})$$

where the integral is around the orbit and  $\mathbf{r}$  is the displacement along the path of the orbit. For this approximately closed path,

the Stoke's theorem gives

$$\begin{aligned} \Delta U_p &= \int (q d\mathbf{S} \cdot \text{curl } \mathbf{E}) \\ &= \int (q d\mathbf{S} \cdot \partial \mathbf{B}/\partial t) \\ &\sim q\pi r_g^2 dB/dt \end{aligned}$$

For time scales much larger than the time period of the gyro motion,

$$dU_p/dt \sim \Omega_g \Delta U_p/2\pi = (q\Omega_g r_g^2/2) dB/dt$$

which gives

$$(1/U_p) dU_p/dt = (1/B) dB/dt$$

or

$$d\mu/dt = d(U_p/B)/dt = 0$$

where  $\mu$  is known as the magnetic moment of the particle,  $\mu = q\Omega_g r_g^2/2$ .

Substituting for  $\Omega_g$ ,  $\mu = (q^2/2\pi m)(\pi r_g^2 B) = (q^2/2\pi m)\Phi_\gamma$ , where  $\Phi_\gamma$  is the flux enclosed by the circular orbit. Since  $d\mu/dt = 0$ ,  $d\Phi_\gamma/dt = 0$ , the flux enclosed by the particle orbit is conserved if the rate of change of the magnetic field is adiabatic, that is, the change occurs over a period much larger than the gyro time period.

### Motion in an Inhomogeneous Magnetic Field

The flux enclosed by a particle orbit also remains constant if the spatial variation of the magnetic field is adiabatic, that is, the scale length of variation is much larger than the gyro radius of the particle orbit. This can be shown simply by the fact that the situation is essentially same as for slow time variation of the field.

The magnetic field variation experienced by the charged particle as it moves in an inhomogeneous magnetic field with a velocity  $v$  is given by

$$dB/dr = v_r dB/dt$$

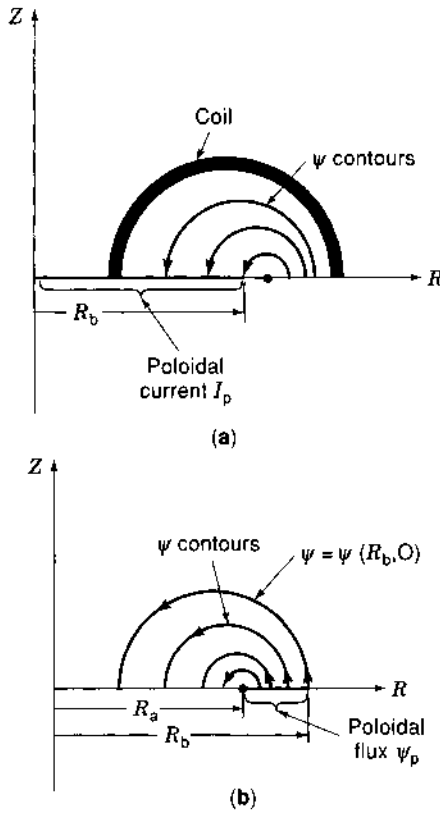
where  $v_r$  is the component of the velocity  $v$  in the direction  $r$ . Again, as shown before, in such a case, the magnetic moment is conserved and therefore the flux enclosed by the particle orbit is conserved.

Other adiabatic invariants such as the bounce invariant in trapped orbits and the line integral of the canonical angular momentum in a periodic motion (17) are also the result of flux conservation.

## PLASMA EQUILIBRIUM AND FLUX SURFACES

The most common devices for nuclear fusion and plasma applications employ a toroidal geometry, where the plasma carries toroidal and poloidal currents (see the section entitled "Flux and Fields in Toroidal Geometry") and are confined by toroidal and poloidal fields. In such cases, the equilibrium is obtained as a balance between the Lorentz body force, which is generated by the interaction of the plasma current with the magnetic





**Figure 16.** (a) Disk-shaped surface through which the total (plasma plus coil) poloidal current  $I_p$  flows. (b) Washer-shaped surface through which the poloidal flux  $\psi_p$  passes.

field, and the pressure force due to gradients in pressure. Such a confinement scheme is used in the Z, theta, and screw pinches, tokamaks, spheromaks, stellarators, and compact toroids. In many of these applications the primary configuration of the plasma is axisymmetric (except for, e.g., helical devices), that is, the variation of the current, magnetic field, pressure, and plasma properties are small and only appear as perturbations. Plasmas in such toroidal geometries attain equilibria (position and shape of the plasma, conditions of magnetic field and plasma current profiles, etc.) based on the solution to the Grad-Shafranov equation. It can be shown that the poloidal flux [see Fig. 16(b)] is constant on specific surfaces. While it is obvious that in the absence of pressure, the surfaces of constant flux are concentric, they are not so when the plasma pressure is finite. Since the outermost flux surface is usually fixed by a flux-conserving boundary or by an external vertical field, this means that the center of the plasma is shifted from the minor axis of the toroid by the so-called Shafranov shift. An equilibrium pressure limit (the so-called equilibrium  $\beta$  limit, where  $\beta$  is the ratio of the plasma pressure to the pressure due to the magnetic field) is obtained when the shift exceeds the radius.

For the geometry shown in Fig. 3, the primary coordinates are the major radius  $R$ , the azimuthal angle  $\phi$  around the major axis, and the vertical coordinate  $z$ . Additional coordinates are the minor radius  $r$  and the poloidal angle  $\theta$ . We limit ourselves to axisymmetric equilibria so that  $\partial/\partial\phi = 0$ .

Maxwell's equations are

$$\text{div } \mathbf{B} = 0 \quad (3a)$$

$$\text{curl } \mathbf{B} = \mu_0 \mathbf{J} \quad (3b)$$

and the plasma force balance equation is

$$\mathbf{J} \times \mathbf{B} = \text{grad } p \quad (3c)$$

where  $\mathbf{B}$  is the magnetic field,  $\mathbf{J}$  the current density and  $p$  the pressure. Expanding the first equation,

$$(1/R)(\partial/\partial R)(RB_R) + (1/R)\partial B_\phi/\partial\phi + \partial B_Z/\partial Z = 0$$

where the second term is zero due to axisymmetry. If we define a flux function  $\psi$ , such that

$$B_Z = (1/R)\partial\psi/\partial R$$

$$B_R = -(1/R)\partial\psi/\partial Z$$

then

$$\mathbf{B} = \mathbf{B}_\phi + \mathbf{B}_p = B_\phi \mathbf{e}_\phi + (1/R)\text{grad } \psi \times \mathbf{e}_\phi \quad (4)$$

The poloidal flux

$$\Phi_p = \int (\mathbf{B}_p \cdot d\mathbf{S}) = \int_{R_0}^R [2\pi R(1/R)(\partial\psi/\partial R) dR] = 2\pi \psi$$

so that the flux function is essentially equal to the poloidal flux except for a constant of  $2\pi$ .

Now taking a scalar product of Eq. (3c) and (3b),

$$\mathbf{B} \cdot \text{grad } p = 0$$

$$(B_\phi/R)(\partial p/\partial\phi) + (1/R)\text{grad } \psi \times \mathbf{e}_\phi \cdot \text{grad } p = 0$$

The first term is zero by axisymmetry; therefore

$$\text{grad } \psi \times \text{grad } p \cdot \mathbf{e}_\phi = 0$$

which shows that the pressure is constant if  $\psi$  is constant or  $p = p(\psi)$ , that is, the flux surfaces are constant-pressure surfaces. This is an important result that says the solution of flux surfaces gives the plasma equilibrium.

Substituting the expression for  $\mathbf{B}$  as in Eq. (4) in Maxwell's Eq. (3b) and using the axisymmetric condition,

$$\begin{aligned} \mu_0 \mathbf{J} = \text{grad}(RB_\phi) \times \mathbf{e}_\phi / R - (1/R)[R(\partial/\partial R)(1/R)(\partial\psi/\partial R) \\ + \partial^2 \psi / \partial Z^2] \mathbf{e}_\phi \end{aligned}$$

The total current density can be divided into poloidal and toroidal components

$$\mu_0 \mathbf{J} = \mu_0 \mathbf{J}_p + \mu_0 \mathbf{J}_\phi$$

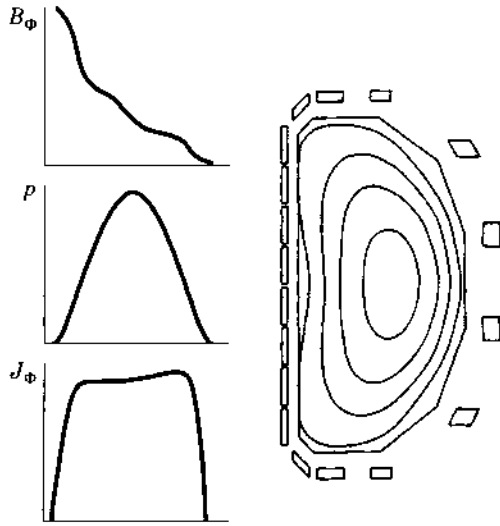
$$\mu_0 \mathbf{J}_p = \text{grad}(RB_\phi) \times \mathbf{e}_\phi / R \quad (5)$$

$$\mu_0 \mathbf{J}_\phi = \Delta^* \psi / R$$

where

$$\Delta^* \psi = R(\partial/\partial R)(1/R)(\partial\psi/\partial R) + \partial^2 \psi / \partial Z^2$$

The quantity  $RB_\phi$  is designated  $F(\psi)$ , which can be shown to be proportional to the total poloidal plasma current enclosed by



**Figure 17.** Numerically computed equilibrium of the noncircular, high- $\beta$  tokamak DIII-D located at GA Technologies. Shown are flux surface plots and midplane profiles. Courtesy J. Helton, GA Technologies.

the flux surface,  $\psi(R, \theta) = \text{const}$ ,

$$I_p = \int (\mathbf{J}_p \cdot d\mathbf{S}) = \int dR \int (R d\theta [\text{grad}(RB_\phi) \times \mathbf{e}_\phi]_z) \\ = 2\pi \int (dR \partial F / \partial R) = 2\pi F(\psi)$$

Now taking a scalar product of Eq. (3c) with  $\text{grad } \psi$

$$\text{grad } \psi \cdot (\mathbf{J} \times \mathbf{B} - \text{grad } p) = 0$$

which after using Eqs. (4) and (5) gives

$$\Delta^* \psi = -\mu_0 R^2 dp/d\psi - F dF/d\psi \quad (6)$$

where the property  $\text{grad } p = dp/d\psi \cdot \text{grad } \psi$  is used.

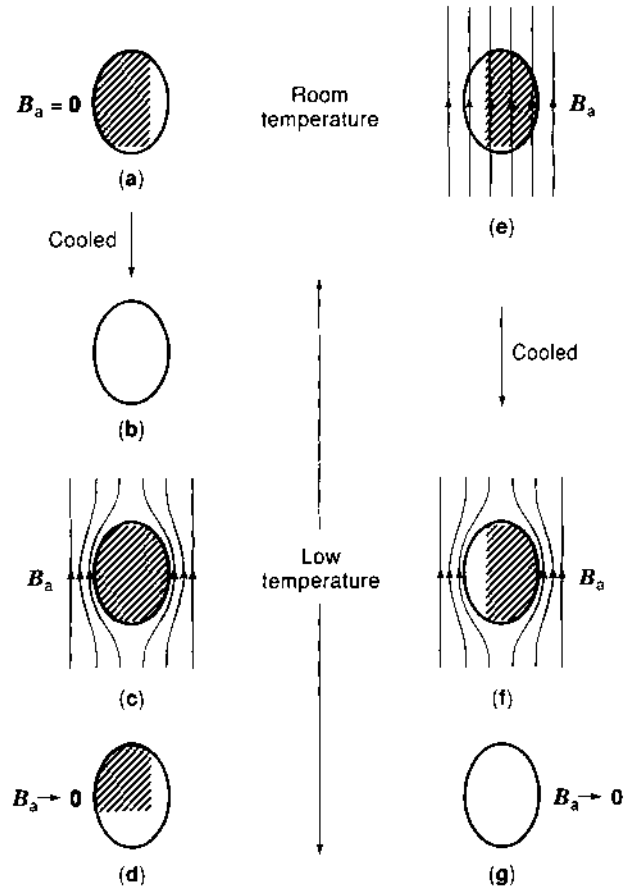
Equation (6) is known as the Grad-Shafranov equation. With appropriate boundary conditions, the equation can be solved to obtain plasma position and equilibrium. The solution is obtained as the solution to the shapes and locations of different flux surfaces (surfaces of constant  $\psi$ ), and since each flux surface has an associated pressure, the flux surfaces define the plasma shape and location. Figure 17 shows an example of the equilibrium for the Doublet IIID tokamak.

Additional discussions on the applications and solutions to the Grad-Shafranov equation can be found in Refs. 4 and 18.

## SUPERCONDUCTORS AND MAGNETIC FLUX

### Superconducting Properties

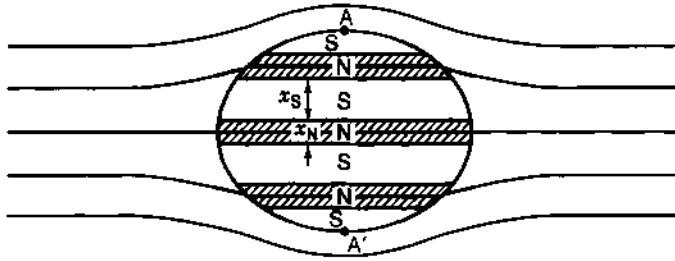
Superconductors are materials that have special properties below the so-called critical temperature, critical field, and critical current density. When such materials are superconducting, they have zero resistivity and in addition they exhibit the Meissner effect (19,20). Earlier it was shown that Maxwell's equations lead to the fact that perfect conductors with zero resistivity exclude flux when the magnetic field (flux) is increased from zero to some finite value. Such perfect conductors maintain the initial flux, and diamagnetic currents cancel any



**Figure 18.** Magnetic behaviour of a superconductor. (a) and (b) Specimen becomes resistive in absence of magnetic field. (c) Magnetic field applied to superconducting specimen. (d) Magnetic field removed. (e) and (f) Specimen becomes superconducting in applied magnetic field. (g) Applied magnetic field removed.  $B_a$  is the applied magnetic flux density (Ref. 4).

change in the flux. However, in 1933, Meissner and Ochsenfeld observed that superconductors that are in the Meissner regime (e.g., lead) exclude all flux whether it was initially present or not (see Fig. 18). This is a significant characteristic of superconductors that distinguishes them from perfect conductors. These so-called type I superconductors receive induced surface currents, called Meissner currents, in the presence of a magnetic field which cancel all the flux inside the superconductor volume, independent of whether the initial flux (flux prior to the material becoming superconductor) was zero or finite. Another way of stating this is that superconductors are not just diamagnetic materials but have a relative permeability  $\mu_r$  that is equal to zero, so that the magnetization is equal and opposite to the applied magnetic induction, that is,  $M = -\mu_0 H$ .

The superconducting property arises from the fact that below the critical temperature and field, the free energy for such materials is lower in the superconducting state compared to the normal (nonsuperconducting) state. This is due to the formation of Cooper pairs of superconducting electrons, which, on average, do not lose any energy by collisions with the lattice ions. The density of superconducting electrons and the free-energy gap at  $T = 0$  and  $H = 0$  are properties of the material. (See Ref. 21 for detailed references on energy-gap measurements.) As the temperature is raised, the free-energy gap is



**Figure 19.** Ellipsoid split into normal and superconducting laminae in a magnetic field.

reduced, and above the critical temperature the superconducting state has an unfavorable free energy, and therefore the material would be normal. As the magnetic field is increased, the total energy, which includes the energy due to magnetization, is increased until again at the critical field, the normal state with zero magnetization is favored and the material would be normal.

While the description of the field being excluded from the volume of the superconductor is reasonable, in reality the external field penetrates to a small depth, the so-called London penetration length  $\lambda_L = (m_e/\mu_0 n_s e^2)^{1/2}$ , where  $e$  and  $m_e$  are the electron charge and mass and  $n_s$  is the density of superconducting electrons (20,22).

However, it must be noted that there is a class of alloy superconductors, known as type II superconductors, which are commonly used in electrical and magnetic applications, the flux (field) is allowed to penetrate into the superconductor above the thermodynamic critical field, while preserving the superconducting (zero resistance) property.

### Intermediate State

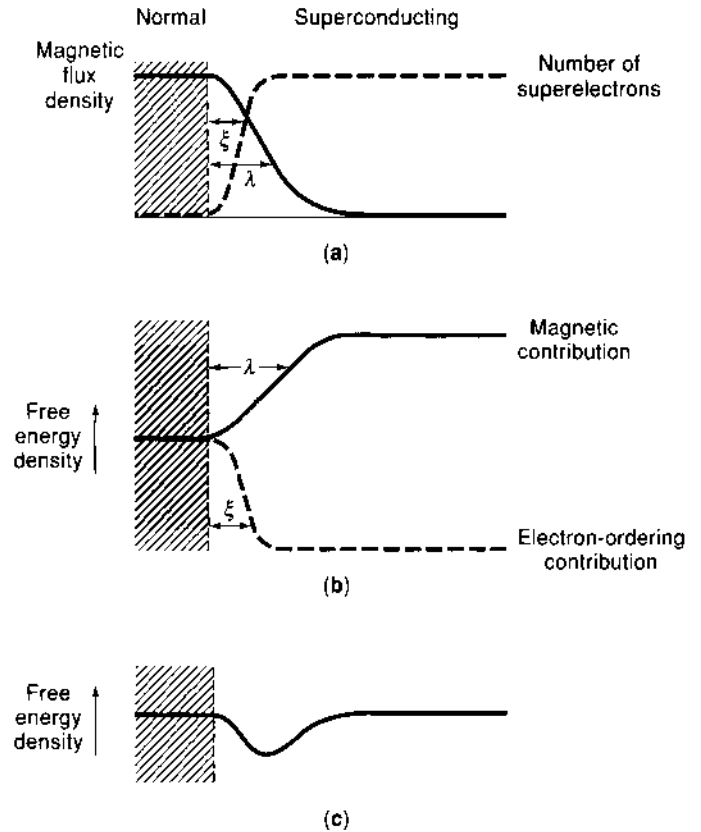
As was noted in the section entitled "Demagnetization Factor," the field intensity in a diamagnetic material is higher than the applied field intensity. Since ideal superconductors exhibiting the Meissner effect have  $I = -H_i$  ( $M = -B_a$ )

$$H_i = H_a/(1 - n)$$

Now, since the superconductor would become normal at  $H_i = H_c$  where  $H_c$  is the critical field intensity, this means that the applied field is less than the critical field. This is a paradoxical situation, since this means that as the material would become normal at  $H_a < H_c$ , which in turn would make  $I = 0$  and we would have the material in a normal state for  $H_a < H_c$ . This is resolved by the realization of the fact that normal and superconducting regions coexist inside the material for  $H_a > (1 - n)H_c$  (Fig. 19). The cross-sectional area of normal material is such that the average magnetization is such as to satisfy the boundary condition  $H_i = H_a/(1 - n)$  for  $H_i < H_c$  and  $H_i = H_a$  for  $H_a = H_c$ . This condition is obtained if

$$H_i = H_a/[1 + n(f - 1)]$$

where  $f$  is the fraction of normal cross section. This state is known as the intermediate state in superconducting materials and is analogous to an equilibrium of solid and liquid phases of matter near transition conditions. For actual observations on the intermediate state see Refs. 23 and 24.



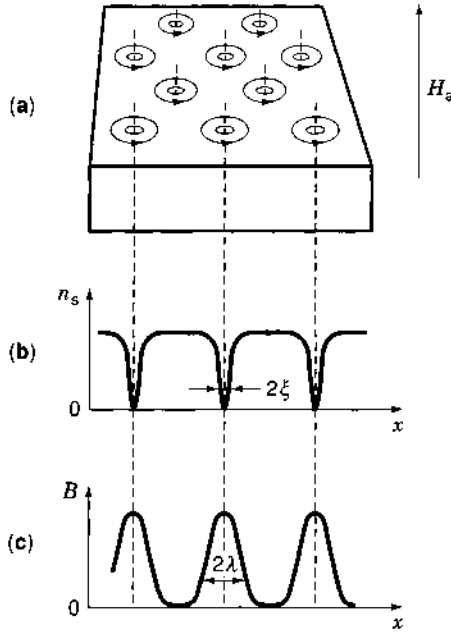
**Figure 20.** Negative surface energy; coherence range less than penetration depth. (Compare this with Fig. 6.9.) (a) Penetration depth and coherence range. (b) Contributions to free energy. (c) Total free energy.

### Type II Superconductors

In a concept proposed by Pippard in 1953, the density of superelectrons cannot change abruptly and changes gradually only over a distance called the coherence length, that is, there cannot be a sharp boundary between normal and superconducting regions. The coherence length is a property of the material and if impurities are present, it is considerably reduced (by an order of 10 or more) to the geometric mean of the pure coherence length and the electron mean free path.

If the coherence length is shorter than the penetration length (see the section entitled "Superconducting Properties"), the formation of coexisting normal and superconducting zones is favored, since then the total free energy of the material is reduced because the surface energy of the boundaries between the normal and superconducting zones is negative for short coherence length (Fig. 20). For the Ginsburg-Landau constant  $\kappa = \lambda/\xi > 0.71$ , where  $\lambda$  is the penetration length and  $\xi$  is the coherence length, the material favors a mixed state of normal and superconducting regions over a fully normal state for applied fields greater than the thermodynamic critical field. Intrinsic superconductors, such as niobium, have  $\kappa > 0.71$  (0.78 to 0.9) even without impurities, but alloys such as niobium-titanium have even higher values of  $\kappa$ .

Therefore in type II superconductors, once the applied field exceeds the thermodynamic (first) critical field, small zones of normal state are formed and the excess field lines are localized along these cores of normal zones, which have circulating currents on their surfaces that preserve the superconducting



**Figure 21.** Mixed state in applied magnetic field of strength just greater than  $H_{c1}$ . (a) Lattice of cores and associated vortices. (b) Variation with position of concentration of superelectrons. (c) Variation of flux density.

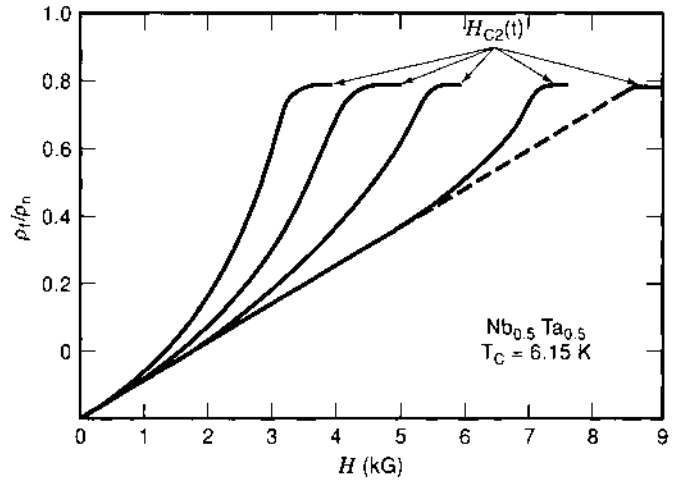
state of the regions outside the cores. Since the surface energy is negative, the formation of the smallest and maximum number of “flux” cores is favored to maximize the total surface area of such cores. Since the coherence length is small in such materials, there can be many fluxons that require sharp transition zones. The flux core is therefore of such a size as to give the minimum flux, that is, the flux of a so called “fluxon” or flux core,  $\Phi_0 = 2.07 \times 10^{-15}$  W (25). The material acquires a lattice of fluxons as shown in Fig. 21. The number of fluxons depends on the amount of flux that needs to pass through the material. Because such fluxons are maintained by circular currents around the flux cores, the fluxons are also called vortices).

Under such conditions, the superconductor does not become normal until the fluxons with the transition regions fill up the area of cross section and the new critical field called  $H_{c2}$  is given by

$$H_{c2} \sim 1.41\kappa H_c$$

where  $H_c$  is the thermodynamic critical field at which the magnetic energy is equal to the difference between the free energy in the normal state and the superconducting state.

**Flux Flow in Type II Superconductors.** While the foregoing is true for a superconductor with no transport current (e.g., current from an external circuit), the amount of current the superconductor can carry in a magnetic field or the critical current requires additional considerations. When a superconducting fluxon lattice is also carrying current, the fluxons experience a Lorentz body force per unit volume of the conductor equal to the vector cross product of the current density and the magnetic field threading the fluxon (in most cases the applied field) (26–28). These forces would move the fluxons perpendicular to both the current density and the applied field, for example, in a wire with a transverse magnetic field, the fluxons would move



**Figure 22.** Variation of the flux-flow resistivity of a NbTa specimen with field at various constant reduced temperatures. (right to left—0, 30%, 50%, 60%, and 70% of the critical temperatures)

radially perpendicular to the field. But these vortices or fluxons are pinned by imperfections in the lattice. Such imperfections in the lattice may be created due to working of the metal or impurities in the metal. Therefore, the superconducting state will be maintained as long as the pinning force per unit volume is larger than the Lorentz force.

As the Lorentz force approaches and exceeds the pinning force, the flux cores start moving and there will be some viscous resistance to such motion. Such a resistance would require work to be performed and energy to be supplied. This power requirement would then manifest itself as an electrical resistance and an associated voltage drop. The critical pinning force is not a constant and increases from zero with field and then reduces again to drop to zero at the upper critical field. Figure 22 (29) shows the increase of the flux flow resistivity with increasing field (30). This is called flux flow, and in this regime the flux cores move with a velocity relative to the electrons carrying the transport current. Since the cores are carried forward (in the direction of the transport current) as well, the net motion of the cores is at an angle to the transport current, but in most cases the angle is close to  $90^\circ$ . The electric field induced by the motion is given by

$$E = n_f \Phi_0 v_f$$

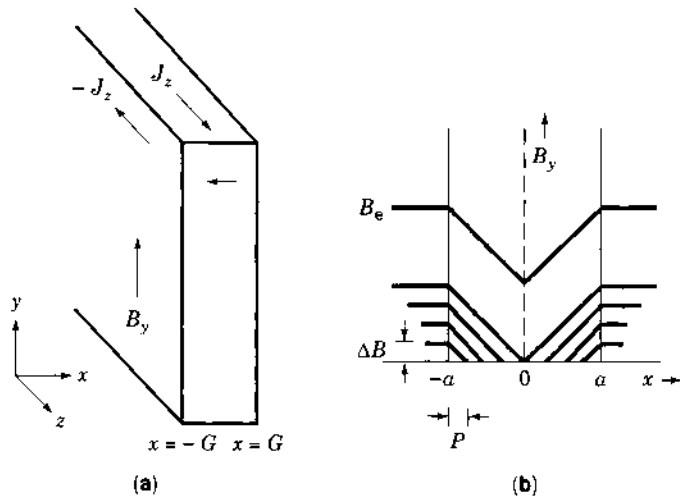
where  $n_f$  is the number of fluxons and  $v_f$  is the velocity of the fluxons (the voltage is given by the product of the electric field and length along the current direction).

Superconducting applications require zero or infinitesimally small resistance and therefore the regime of flux flow is required to be as close to the critical current as possible. Therefore, most applications require a high  $n$  value, which is given by

$$E \sim (I/I_c)^n$$

where  $I$  is the transport current and  $I_c$  is the critical current.

**Flux Penetration and Flux Jump in a Type II Superconductor.** According to the critical state model, when a field is applied

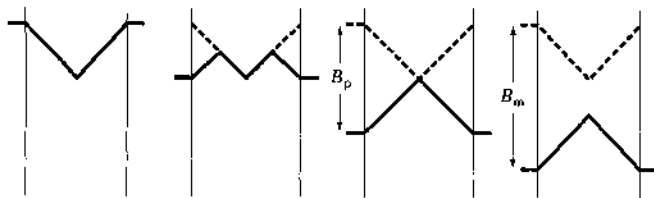


**Figure 23.** (a) Screening currents induced to flow in a slab by a magnetic field parallel to the slab surface; (b) Magnetic field pattern across the slab showing the reduction of internal field by screening currents.

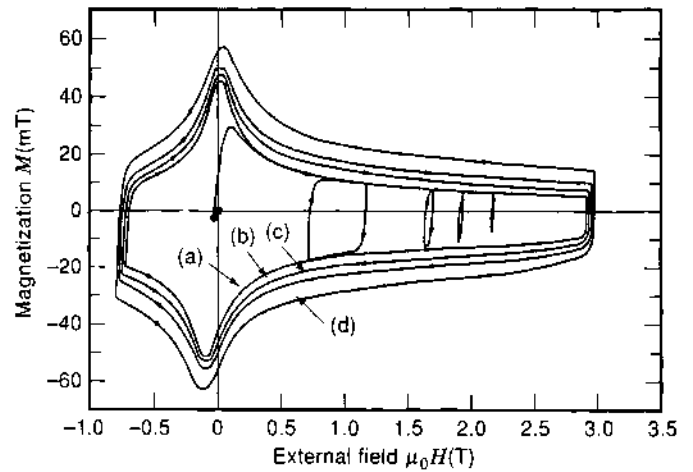
in the exterior of the superconductor, screening currents would be induced to exclude the field. The cross section (proportional thickness in the slab shown in Fig. 23) of the current flow is equal to the total current divided by the critical current of the specimen. The current density in the material is always equal to the critical current density. As the field is increased, so is the thickness of the current sheet and after current flows throughout the cross section of the material, the field is fully penetrated and then increases inside the material as the field is raised.

Unlike in normal conducting materials, the penetration of the field in the superconductor does not reverse when the field change direction is reversed. If a field has penetrated well inside the material during one direction of change (say the field is increased) and if the field change direction is reversed (say the field is decreased), the field inside the superconductor initially decreases on the edge of the superconductor while the field inside the superconductor remains unaffected. As the field change is continued further this reduction continues into the thickness. Figure 24 shows this phenomenon schematically. This behavior of the diamagnetism causes the superconductor magnetization to be hysteretic (31). The magnetization hysteresis for a typical superconductor is shown in Fig. 25.

Now the critical state can be unstable because if there is perturbation in the form of a temperature increase, the critical current density is reduced, which then causes the fields in the superconductor to redistribute requiring motion of flux in the superconductor. This flux motion generates heat, resulting in the further increase of temperature. The larger the sensitivity



**Figure 24.** (a) Field pattern within a superconducting slab subjected to large field change; (b) as the field is reduced; (c) when the field change penetrates to center of slab; (d) when the field reaches a minimum value before rising again.



**Figure 25.** Magnetization of a 361-filament NbTi/copper composite with a twist pitch of 25.4 mm, measured at (a)  $0.0075 \text{ T} \cdot \text{s}^{-1}$ , (b)  $0.0375 \text{ T} \cdot \text{s}^{-1}$ , (c)  $0.075 \text{ T} \cdot \text{s}^{-1}$ , and (d)  $0.15 \text{ T} \cdot \text{s}^{-1}$ .

of the critical current density to the temperature the larger the heat produced. The smaller the specific heat of the material, the larger the temperature increase for a given heat generated. Therefore, if conditions are unfavorable, the superconducting material will run away in temperature, resulting in the material quenching and the flux jumping inside the material. The condition beyond which such a flux jump would occur is given by

$$\mu_0 J_c^2 a^2 / [3\rho_m C (T_c - T_0)] < 1$$

where  $\rho_m$ ,  $T_c$ ,  $J_c$ ,  $C$ , and  $a$  are the density, critical temperature, critical current density, specific heat, and thickness of the superconductor, and  $T_0$  is the bath (initial) temperature (32). The result of the flux jump is that filaments of a superconductor cannot be larger than a certain size and when a superconductor is made of a large number of superconducting filaments of wire, these must be twisted to cancel the diamagnetic currents over short distances (33).

### Superconductor Performance under ac Conditions

When superconductors are operated under ac conditions (applied magnetic field and currents), the superconductor response is significantly different from that of a good conductor. There are two reasons for this difference: (a) the magnetization of the superconductor is hysteretic, (b) there are saturation effects due to the criticality with respect to magnetic field and superconducting current capability. In addition, since usual superconductors are composites of multiple superconducting filaments as well as normal stabilizing conductors like copper, there are coupling effects due to mutual inductance and cross conductance. The subject is wide, and details are strongly dependent upon the details of the superconductors and applications (34,35).

Under alternating magnetic field conditions, diamagnetic currents induce reverse magnetization, and since this magnetization is hysteretic, in a full cycle, an energy loss per unit volume of  $\int [\mathbf{M} \cdot \mathbf{B}] d\mathbf{B}$  is incurred. This loss is similar to the hysteresis loss in ferromagnetic materials.

With an alternating magnetic field transverse to the superconductors, diamagnetic saddle currents flow similar to the

normal conductors, but these currents are not limited by conductivity but rather by the critical current. If the superconductor is a monofilament, large currents would mostly flow in the skin of the superconductor, and the superconductor would be unstable. In order to limit the current, the superconductor is made of several filaments, twisted together and separated by a conducting matrix, for example, of copper. The induced voltage is then limited by the twist length, and the currents are distributed over all the filaments. However, for cryogenic stability and stability against flux jump, the matrix has to be made of high conductivity material, and this leads to the fact that the filaments couple somewhat through the matrix. Thus currents flowing along the filaments cross over through the matrix, and such currents induce ac losses in the superconductors. Such ac losses have to be limited to prevent heating of the conductor and subsequent quenching (transition to normal state).

Another issue under alternating field and current conditions is that the filaments or a cable of superconductors are not exactly identical due to differences in superconducting characteristics, twist pitches and end effects, and therefore currents may not be shared equally. A strong nonuniform current distribution can result from very small differences. The superconductor performance under those conditions would be significantly poorer than the sum of the individual filaments or superconductors.

#### MEASUREMENT OF FLUX

It is usually only necessary to measure magnetic fields, magnetization, etc. However, in special cases where the measurement of flux is desired in a static field region, a good way is to place a coil and move it transverse to the field direction so that the coupled flux changes. The time integral of induced voltage then gives the change in flux over the amplitude of motion. In instances where the flux is changed over time, a stationary coil enclosing the flux can be used. It is also appropriate to use flux density probes such as "Hall" probes over the area of interest and integrate the flux density over the area.

#### BIBLIOGRAPHY

1. A. H. Morrish, *Physical Principles of Magnetism*. New York: Krieger, 1980.
2. P. A. Sturrock, *Plasma Physics*, New York: Cambridge Univ. Press, 1994, pp. 19–25.
3. J. A. Stratton, *Electromagnetic Theory*, New York: McGraw-Hill, 1941, p. 206.
4. K. Miyamoto, *Plasma Physics for Nuclear Fusion*, Cambridge, MA: MIT Press, 1976.
5. H. A. B. Bodin and A. A. Newton, *Nucl. Fusion*, **20**: 1255, 1980.
6. L. Woltjer, *Proc. Natl. Acad. Sci.*, **44**: 489, 1958.
7. D. R. Wells and J. Norwood, Jr., *J. Plasma Phys.*, **3**: 21, 1969.
8. A. Janos, *Proc. 6th U.S. Symp. Compact Toroid Res.*, Princeton, NJ, 1984, pp. 97–102.
9. A. Janos et al., Princeton Plasma Physics Rep. No. PPPL-2214 (1985).
10. M. Steenbeck et al., *Z. Naturforsch.*, **21a**: 369–376, 1966.
11. H. K. Moffatt, *Magnetic Field Generation in Electrically Conducting Fluids*, New York: Cambridge Univ. Press, 1978, p. 343.
12. R. T. Merrill et al., *The Magnetic Field of the Earth*, Vol. 63 of

- International Geophysics Series, New York: Academic Press, 1996, pp. 317–320.
13. E. N. Parker, Hydromagnetic dynamo models, *Astrophys. J.*, **122**: 293–314, 1955.
14. G. McPherson, *Introduction to Electrical Motors and Transformers*, New York: Wiley, 1981.
15. R. Lee, *Electronic Transformers and Circuits*, New York: Wiley, 1947.
16. J. D. Jackson, *Classical Electrodynamics*, 2nd ed., New York: Wiley, 1975, pp. 571–612.
17. H. P. Furth and M. N. Rosenbluth, *Proc. Conf. Plasma Physics Controlled Nuclear Fusion*, Vol. 1, Vienna: IAEA, 1962, p. 821.
18. J. Freidberg, *Ideal MHD Equilibrium and Stability*, New York: Plenum, 1987.
19. A. C. Rose-Innes and E. H. Rhoderick, *Introduction to Superconductivity*, New York: Pergamon, 1978, pp. 19–34.
20. P. G. de Gennes, *Superconductivity of Metals and Alloys*, New York: Benjamin, 1966, p. 4.
21. D. H. Douglas, Jr. and L. M. Falicov, *Low Temperature Physics*, C. G. Gorter (ed.), Amsterdam: North-Holland, 1964.
22. A. L. Schawlow and G. Devlin, *Phys. Rev.*, **113**: 120–126, 1959.
23. W. de Sorbo, *Phys. Rev. Lett.*, **4**: 406–408, 1960.
24. A. L. Schawlow, *Phys. Rev.*, **101**: 573–579, 1956.
25. A. Abrikosov, *Sov. Phys.—JETP*, **5**: 1174–1198, 1957.
26. C. P. Bean, *Rev. Mod. Phys.*, **36**, 31, 1964.
27. A. M. Campbell and J. E. Evetts, *Adv. Phys.*, **21**: 90, 1972.
28. P. W. Anderson, *Phys. Rev. Lett.*, **9**: 309–311, 1962.
29. Y. B. Kim et al., *Phys. Rev.*, **139**: 1163–1172, 1965.
30. E. J. Kramer, *J. Appl. Phys.*, **44**: 1360–1370, 1973.
31. M. N. Wilson, *Superconducting Magnets*, Oxford: Clarendon, 1983, p. 162.
32. P. S. Swartz and C. P. Bean, *J. Appl. Phys.*, **39**: 4991, 1968.
33. M. N. Wilson, *Superconducting Magnets*, Oxford: Clarendon, 1983, p. 135.
34. W. J. Carr, Jr., *Ac Loss and Macroscopic Theory of Superconductors*, New York: Gordon and Breach, 1983.
35. Ref. 31, pp. 159–199.

RAGHAVAN JAYAKUMAR  
Lawrence Livermore National  
Laboratory

#### MAGNETIC LEVITATION

To see gravity defied by the use of magnetic levitation in any form usually brings a sense of magic to the observer. Although mention of magnetic levitation these days probably brings to mind either a small permanent magnet floating above a superconductor or a high-speed magnetically levitated (maglev) train flying above a metal guideway, accounts of magnetic levitation go back to at least the 11th century (1). As we will see, magnetic levitation is robust, in that nature has provided many different ways to achieve it.

#### MAGNETIC LEVITATION METHODS

##### Magnetic Forces

The basic forces involved in magnetic levitation are derived from the basic laws of electricity and magnetism. The first basic

force is the force between magnetic poles. Although isolated magnetic poles are not known to exist in nature, they constitute a convenient model and are discussed in most elementary science texts. For a pair of magnetic poles, opposite poles attract, like poles repel, and the force between point poles is proportional to the inverse square of the distance between them.

The second basic force involved in magnetic levitation is the force that occurs when a conductor moves in a magnetic field. This force is derived from Faraday's law and Lenz's law, which teach that a changing magnetic flux produces a voltage in a circuit in such a way that it opposes the change in flux. This law is responsible for jumping rings and eddy currents (2).

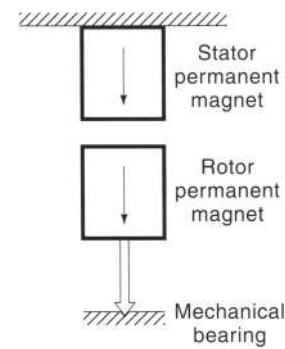
The response of a material to magnetic fields also produces a magnetic levitation force. A diamagnetic material, such as graphite, aluminum, or a superconductor, will be repulsed by a magnetic pole, whereas a paramagnetic material, such as oxygen or water, will be attracted to a magnetic pole. Ferromagnetic materials, such as iron, are strongly attracted to magnetic poles.

### Earnshaw's Theorem

Earnshaw's theorem is an important theorem that affects the static levitation of magnetic systems. The theorem is developed from the property of curl- and divergence-free fields that precludes the existence of local, detached, scalar-potential maxima or minima. According to this classical theorem, it is impossible to attain stable equilibrium in a system in which only inverse-square-law electrostatic or magnetostatic forces are acting (3). Braunbek deduced that electric or magnetic suspension is not possible when all materials have  $\epsilon_r > 1$  or  $\mu_r > 1$ , but that it is possible when materials with  $\epsilon_r < 1$  or  $\mu_r < 1$  are introduced (where  $\epsilon_r$  is the relative electrical permittivity and  $\mu_r$  is the relative magnetic permeability) (4). Earnshaw's theorem is grasped intuitively by most people when they release a permanent magnet next to the ferromagnetic door of their refrigerator. The magnet either moves to stick to the door, or it falls on the floor; it does not hover in space near the point where it was released. Braunbek (4), and later Boerdijk (5), experimentally demonstrated the stable levitation of small pieces of bismuth and graphite (these are slightly diamagnetic materials with  $\mu_r < \approx 1$ ) in strong magnetic fields. Arkadiev (6) first reported the stable levitation of a strongly diamagnetic superconductor ( $\mu_r \ll 1$ ). However, it was not until the discovery of superconductors with critical temperatures (i.e., temperatures at which a material enters the superconducting state) above that of liquid nitrogen that passively stable levitation became a common laboratory occurrence. We note that Earnshaw's theorem only applies to conditions of static stability and does not apply to dynamic systems.

### Levitation of Permanent Magnets

Although magnetic levitation of one permanent magnet by another is not stable according to Earnshaw's theorem, this arrangement is still useful if stability is provided by other means. A simple rendition of such a levitation, originally proposed by Evershed (7), is shown in Fig. 1. Here the levitated object, that is, the rotor, consists of a permanent magnet below which hangs a rigid rod with a point on the bottom. Each permanent magnet is magnetized, as shown by the dark arrow in Fig. 1, with its north pole down. The permanent magnet of the rotor experiences an attractive force toward the stationary permanent



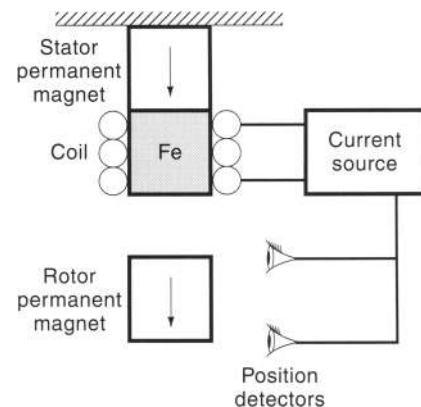
**Figure 1.** Evershed bearing design, in which most of the weight is provided by attraction between permanent magnets, and the remaining weight and vertical stability are provided by a small mechanical bearing.

magnet immediately above it. This system is statically stable in the radial direction but unstable in the vertical direction. The gap between the two magnets is adjusted in such a manner that the attractive force between the pair of magnets is just less than 100% of the rotor weight, so that the rotor would tend to fall. The remainder of the weight is provided by the small mechanical bearing at the bottom, which supplies sufficient stiffness for vertical stability and some additional radial stability.

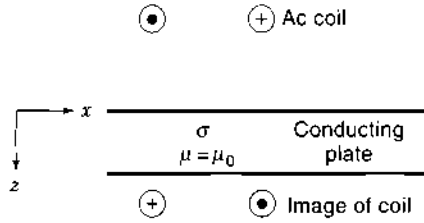
The Evershed design, either in the simple form shown in Fig. 1, or in some modification, is used in applications ranging from simple toys and watt-hour meters to high-speed centrifuges. This basic design may be combined with other magnetic levitation techniques and is often referred to as magnetic biasing.

### Electromagnetic Levitation

Truly contact-free magnetic levitation can be attained by replacing the mechanical bearing in Fig. 1 with a small electromagnet and a feedback system, as shown in Fig. 2. As before, most of the levitation force is provided by the attractive force between the two permanent magnets. The coil in Fig. 2 is wired to magnetize the iron in the same sense that the permanent magnets are magnetized. Then, as in Fig. 1, when the coil is energized, the system is stable in the radial direction but unstable in the vertical direction. The feedback system consists of



**Figure 2.** Attractive levitation of a permanent magnet with feedback control of an electromagnet.



**Figure 3.** Repulsive levitation of an ac coil above a conducting sheet using method of images.

two sensors that detect the top and bottom edge of the levitated permanent magnet. For example, the sensors may be photocells that detect light from a light source to the left of the magnet. When light to the bottom sensor is obscured by the falling magnet, the current to the coil is turned on, and the attractive force increases. When light to the top sensor is obscured by the rising magnet, current to the coil is turned off, and the attractive force decreases. In such a manner, the magnet may be stably levitated.

This basic design forms the basis of most active magnetic-bearing concepts. As discussed in the applications section, the actual control algorithms and sensors may be more complicated than the simple system illustrated in Fig. 2.

### AC Levitation

The ability of electromagnetic forces to impart significant levitation forces through Faraday's law is well known from jumping-ring experiments (2) that are performed in many introductory physics classes. Consider the system shown in Fig. 3, in which a coil is stationed above a conducting plate. If the coil is energized with a pulse so that current flows as shown, by Faraday's law, eddy currents will be induced in the plate in such a way that magnetic flux is expelled from the plate. The effect of these eddy currents is that the coil will see its mirror image, as shown in Fig. 3, and the interaction between the coil and its image produces a repulsive levitation force. Because of the finite electrical conductivity of the plate, the eddy currents in the plate that arise from a current pulse in the coil will exponentially decay in time, as will the levitation force, after the pulse has ended. In the jumping ring experiment, the conducting plate consists of a ring that sits on top of the coil. When the coil is pulsed, the mutual repulsive force causes the ring to accelerate upward.

To provide a more continuous levitation force, it is necessary to provide another current pulse in the coil before the initial eddy currents have completely died away. This is most conveniently accomplished by supplying the coil with an alternating current. The general phenomenon can be quantitatively understood by considering an alternating magnetic field,  $\mathbf{B} = B\hat{i}$  (where  $\hat{i}$  is the unit vector in the  $x$  direction and  $B = B_0 \exp(j\omega t)$ ) incident on the surface of a half space  $z > 0$ , with electrical conductivity  $\sigma$  and magnetic permeability  $\mu$ , where  $j$  is the square root of  $-1$ ,  $\omega$  is the radial frequency, and  $t$  is time. One then must solve the magnetic diffusion equation

$$\partial^2 \mathbf{B} / \partial z^2 = \mu \sigma \partial \mathbf{B} / \partial t \quad (1)$$

The solution to Eq. (1) in the conducting half space is

$$\mathbf{B} = B_0 \exp(-z/\delta) \exp[j(\omega t - z/\delta)] \quad (2)$$

where the skin depth  $\delta$  is given by

$$\delta^2 = 2/\mu\sigma\omega$$

The current density  $\mathbf{J}$  in the half space, given by Maxwell's equation  $\mathbf{J} = \nabla \times \mathbf{H}$ , is in the  $-y$  direction; its magnitude is given by

$$J = (B_0/\mu\delta)(1+j) \exp(-z/\delta) \exp[j(\omega t - z/\delta)] \quad (3)$$

Thus, we see that current density has the same exponential decay as the magnetic field but is phase-shifted by  $45^\circ$ . Here, the force per unit volume, given by  $\mathbf{F} = \text{Re}\{\mathbf{J}\} \times \text{Re}\{\mathbf{B}\}$ , is in the  $z$  direction and its magnitude is

$$F(z) = (B_0^2/\mu\delta) \exp(-2z/\delta) [1/2 - 2^{-1/2} \sin(2\omega t - 2z/\delta - \pi/4)] \quad (4)$$

The force consists of a time-independent part plus a sinusoidal part that is twice the applied frequency. The mean force on the plate is downward, as expected, with a corresponding force upward on an ac coil above the plate. The pressure  $P$  at the surface of the plate is given by

$$P = \int_0^\infty F(z) dz \quad (5)$$

$$P = (B_0^2/4\mu)[1 + \cos(2\omega t)] \quad (6)$$

The average levitation pressure is independent of frequency and proportional to the square of the applied magnetic field. To achieve a relatively constant levitation height, it is desirable that the period of the applied field  $\tau = 2\pi/\omega$  be much smaller than the characteristic time of the mechanical motion. However, the frequency cannot be made arbitrarily high, because this type of levitation is associated with joule heating; the heating rate  $Q$  per unit volume is given by

$$Q = J^2/\sigma \quad (7)$$

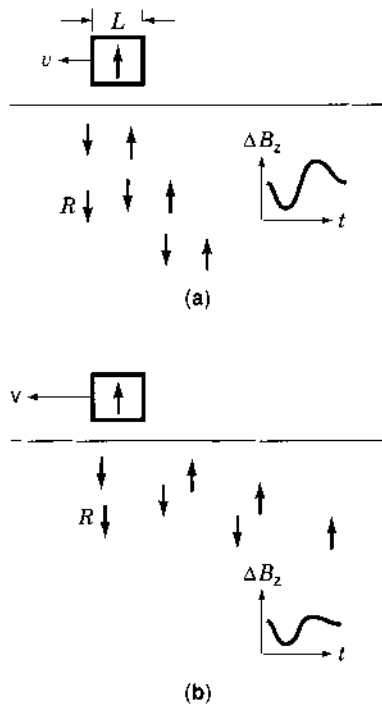
From Eq. (3), we surmise that the maximum heating rate occurs at the surface and is proportional to the frequency, with the total heating rate proportional to the square root of the frequency.

### Electrodynamic Levitation

In electrodynamic levitation, a moving magnet (permanent magnet, electromagnet, or superconducting magnet) interacts with a conducting sheet or a set of coils to produce a levitation force. A drag force, typically much higher than that associated with ferromagnetic suspensions, is associated with the eddy currents. However, above some speed, the drag force decreases as unity divided by the square root of the velocity. The system is passively stable, that is, no feedback is required. The disadvantage is that there is a minimum speed below which the levitation force is not sufficient, so some mechanical support is needed on startup.

The subject of eddy currents is a separate article in this encyclopedia; hence, here we limit the discussion to eddy currents caused by a moving magnet. The phenomenon can be understood by applying the principle of images, as shown in Fig. 4. In





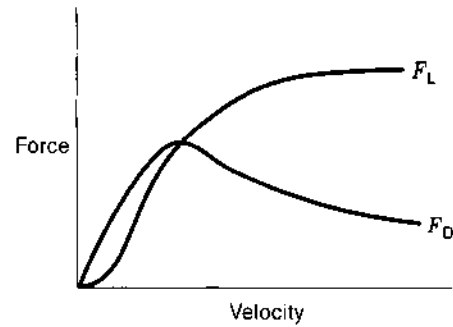
**Figure 4.** Maxwell's eddy current model applied to a magnet moving over a conducting plane: (a) low velocity; (b) high velocity. The vertical component of the induced images and the magnetic field due to the eddy currents are shown in each case.

the case of a plane conducting sheet, the imaginary system on the negative side of the sheet is not the simple image, positive or negative, of the real magnet on the positive side, but consists of a moving train of images (2,8–11). According to this model, when a magnet passes a point on the conducting plane, it induces first a "positive" image, then a "negative" image. These images propagate downward at a velocity  $R$ , which is proportional to the specific resistivity (and to the reciprocal thickness if the sheet is thin when compared with the skin depth).  $R$  is also the electrical resistance of a square portion of the conducting sheet; its value  $R = \rho/2\pi h$  ( $\rho$  = resistivity,  $h$  = thickness) is independent of the size of the square. In electromagnetic units,  $R$  has the dimensions of velocity.

Two examples that apply Maxwell's model are shown in Fig. 4. In the first example, the velocity  $v$  of the magnet is  $< R$ . The positive image has moved down a distance  $Rdt = RL/v$  when the negative image appears at the same location. Then, as the two images move away head-to-tail, the induced field falls toward zero. Because of the eddy currents, the vertical component of magnetic field  $\Delta B_z$  is in one direction near the leading edge of the moving magnet and the opposite direction at the trailing edge, which leads to the waveform shown at the right of the figure.

In the second example [Fig. 4(b)], the velocity is considerably greater than  $R$ . The positive image has moved only a small distance  $RL/v$  away when the negative image appears, and the two images nearly cancel each other thereafter. Now, the magnetic field due to the eddy currents  $\Delta B_z$  is predominantly in one direction, as shown in Fig. 4(b).

Reitz (12) solved Maxwell's equation for several types of moving magnets with the geometry shown in Fig. 4. In each case, he obtained a "wake of images," similar to those shown in



**Figure 5.** Velocity dependence of lift force  $F_L$  and drag force  $F_D$ .

Fig. 4, moving into the plate with a velocity  $w = 2\rho/\mu_0 h$ , which is Maxwell's  $R$  expressed in rationalized mks units.

The force on a magnet moving over a nonmagnetic conducting plane can be conveniently resolved into two components: a lift force perpendicular to the plane and a drag force opposite to the direction of motion. At low velocity, the drag force is proportional to velocity  $v$  and considerably greater than the lift force, which is proportional to  $v^2$ . As the velocity increases, however, the drag force reaches a maximum (referred to as the drag peak) and then decreases as  $v^{-1/2}$ . On the other hand, the lift force, which increases with  $v^2$  at low velocity, overtakes the drag force as velocity increases and approaches an asymptotic value at high velocity, as shown in Fig. 5. The lift-to-drag ratio, which is of considerable practical importance, is given by  $F_L/F_D = v/w$ .

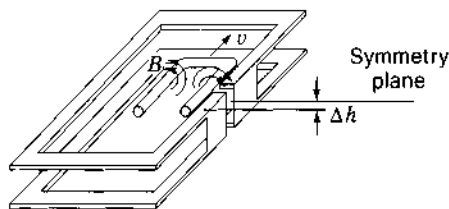
Qualitatively, these forces can be understood by considering the diffusion of magnetic flux into the conductor. When a magnet moves over a conductor, the flux tries to diffuse into the conductor. If the magnet is moving rapidly enough, the flux will not penetrate very far into the conductor, and the flux compression between the magnet and the conductor causes a lift force. The flux that does penetrate the conductor is dragged along by the moving magnet, and the force required to drag this flux along is equal to the drag force. At high speeds, less of the magnetic flux has time to penetrate the conductor. At high speed, the lift force that is a result of flux compression approaches an asymptotic limit, and the drag force approaches zero.

The lift force on a vertical dipole of moment  $m$  moving at velocity  $v$  at a height  $z_0$  above a conducting plane can be shown to be (12)

$$F_L = 3\mu_0 m^2 / 32\pi z_0^4 [1 - w(v^2 + w^2)^{-1/2}] \quad (8)$$

At high velocity, the lift force approaches the ideal lift from a single image:  $3\mu_0 m^2 / 32\pi z_0^4$ ; at low velocity, the factor in the brackets is approximately equal to  $v^2/2w^2$ , so the lift force increases as  $v^2$ .

The drag force, as already pointed out, is  $w/v$  times the lift force, so the drag force is proportional to  $v$  at low velocity. According to the thin-plate model that we have been discussing thus far, the drag force should fall off with  $1/v$  as the lift force reaches its high-speed limit. However, at high velocity, penetration of the eddy currents and magnetic fields is limited to the skin depth, which is proportional to  $v^{-1/2}$ . As a first approximation, one might replace plate thickness by skin depth at high speed. The transition from thin-plate to skin-depth behavior



**Figure 6.** Null-flux geometry, showing full stationary null-flux coil and one half of levitated coil moving with velocity  $v$  and displacement from the symmetry plane by  $\Delta h$ .

should occur at about 30 m/s in a 1-cm-thick aluminum plate, for example.

One may improve on the basic system of a magnet moving over a conducting plate by using the null-flux-geometry system (13) shown in Fig. 6. When the moving magnet or coil is in the symmetry plane, no net flux threads the track loop so that lift and drag forces approach zero. Lift forces increase linearly and drag forces as the square of the (small) displacement  $\Delta h$  from the symmetry plane. The velocity dependence of null-flux systems is the same as that of the eddy current systems.

### Levitation by Tuned Resonators

Several magnetic bearing concepts involve passive techniques. These concepts have the advantage of simplicity and the lack of a control system. One system achieves a stable stiffness characteristic by using an  $LC$ -circuit excited slightly off resonance (14). The  $LC$ -circuit is formed with the inductance of the electromagnetic bearing coil and a capacitor. The mechanical displacement of the rotor changes the inductance of the electromagnet. The  $LC$  circuit is operated near resonance and tuned in such a way that it approaches resonance as the rotor moves away from the electromagnet. This increases the current from the ac voltage source and thus pulls the rotor back to its nominal position. The low forces and stiffnesses of this system, coupled with the necessity for continuous ac energization of the coils, are disadvantages. The system is also subject to a low-frequency, negative-damping instability, and auxiliary damping is usually required for stability.

### Miscellaneous Levitation Methods

The Levitron<sup>®</sup> is a toy, manufactured by Fascinations in Seattle, Washington, that consists of a spinning permanent magnet in the form of a top. The spinning magnet is repelled by a magnet of opposite polarity in a stationary base beneath the top. The levitation height is such that the top is statically stable in the vertical direction and unstable in the radial direction. As designed, the top is also statically unstable to rotation of the magnetic moment. The concept of the Levitron is similar to that of other passive magnetic bearing concepts, in that its stability is dynamical (1,15,16). It is distinguished from other passive magnetic bearing concepts in that its stability is the result of the tendency of its precession axis to align with the local field direction: that is, when the top moves radially away from equilibrium, the magnetic moment of the top will rotate so as to point in the direction of the magnetic field from the stationary magnet. This rotation of the moment adds a term to the potential energy of the system that is not present if the magnetic moment is decoupled from the radial excursion. There is a rotational frequency range for the dynamic stability. The

top must be spinning fast enough so the gyroscopic action prevents it from flipping over. However, if the top is spinning too fast, the gyroscopic action will maintain the direction of the moment vertical, and it cannot reorient fast enough to point in the direction of the magnetic field. If the top is able to precess several times around the field line during the time scale of a radial excursion, the system is said to be adiabatically stable. The same concept is used to trap cold neutrons and produce Bose-Einstein condensation (16). Under presently investigated conditions, the frequency range in which the top is stable is rather small, and the stability is rather fragile. It is not clear whether this concept would be useful in practical magnetic bearings.

Somewhat related to the Levitron levitation technique is the levitation of weakly diamagnetic materials in high magnetic fields. For example, a 16 T steady magnetic field has been used to levitate a living frog (17). The magnetic field induces a magnetic moment  $m$  in the diamagnetic material that is in the opposite direction to the applied field  $B$ , where the magnitude of the moment is given by

$$m = |\chi|VB/\mu_0$$

where  $\chi$  is the susceptibility of the material (water, which constitutes the majority of the mass of most living creatures, is weakly diamagnetic with  $\chi = -1.3 \times 10^{-5}$ ),  $V$  is the material volume, and  $\mu_0$  is the permeability of free space. The magnetic moment of the diamagnetic material interacts with the gradient of the applied magnetic field to produce a levitation force. The vertical component of this force against gravity is given by

$$F_z = m dB/dz$$

For the levitation to be stable, the levitated object must be placed where the energy is a minimum. This is only possible with diamagnetic objects. Further, the levitation is stable because, like the Levitron, the magnetic moment of the diamagnetic object rotates with the applied magnetic field.

Another magnetic levitation technique that is dynamically stable is analogous to the strong-focussing technique used in particle accelerators (18). In this system the static stability of the magnetic system alternates as the levitated part rotates. During one portion of the rotation, the system is radially stable and vertically unstable. During the next portion of the rotation, the system is vertically stable and radially unstable. For some frequency ranges, such a system can be dynamically stable.

### Magnitude of Levitation Pressure

The levitation pressure of a magnetic system is considerably smaller than that of most mechanical systems. Here, we consider two systems: two magnetized objects, such as a pair of permanent magnets or a permanent magnet and a ferromagnet; and a coil traveling at some velocity  $v$  over a conducting sheet.

The maximum magnetic pressure  $P$  between two magnetized objects of magnetization  $M_1$  and  $M_2$  occurs at zero gap between the two objects and is given by

$$P = M_1 M_2 / 2\mu_0$$

where  $\mu_0 = 4\pi \times 10^{-7}$ . As convenient reference, for  $\mu_0 M_1 = \mu_0 M_2 = 1.0$  T, the pressure is 400 kPa. In a sintered NdFeB permanent

magnet,  $M$  is typically between 1 and 1.5 T; ferromagnetic materials may achieve magnetizations up to about 2.5 T.

For a set of dc coils of alternating polarity of spatial period  $L$  moving at a height  $h$  over a conducting sheet, the maximum levitation pressure is given by the pressure of the image force and is

$$P_1 = (1/2\pi\mu_0)B_0^2 \exp(-4\pi h/L)$$

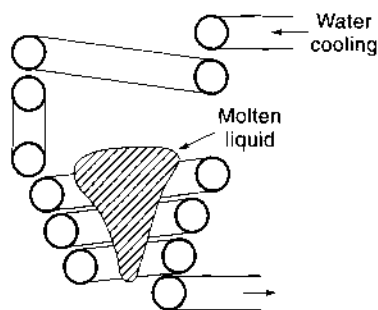
where  $B_0$  is the rms (spatially averaged) value of the magnetic induction in the plane of the magnets. With a NbTi superconducting coil,  $B_0$  can easily be 5 T.

## APPLICATIONS

### Levitation Melting

We have seen earlier that an ac coil may levitate over a conducting plate by means of eddy currents induced in the plate. If an ac coil is wound in approximately the shape shown in Fig. 7, it may support a conducting liquid. In the case shown in Fig. 7, the coil supports the sloping sides of the levitated liquid, but there is a "magnetic hole" at the bottom of the coil. The levitated liquid is prevented from leaving the magnetic trap through this hole by its own surface tension. The size of the melt is then determined by the surface tension, and levitation melters are usually only capable of supporting masses substantially less than 1 kg. Such levitators are often used to conduct reactions in which contamination of the sample with material from the crucible walls cannot be tolerated. In such a case, the original charge is often solid and the induction heating from the eddy currents melts the charge. Such devices are also used when a highly homogeneous final product is desired, because eddy currents cause significant stirring of a molten metal. The levitation melter is also used to accurately determine the surface tension of liquids, equilibria between liquid metals and gases, and thermal diffusion.

Efforts to increase the mass of molten material to be levitated have involved the use of a water-cooled copper jacket with thin vertical slots surrounded by a conventional solenoidal coil. Such a device, usually referred to as a cold crucible (19), can levitate kilogram quantities of molten metal. The shape of the metal that is levitated in cold crucibles is mostly spherical. Attempts to levitate low-aspect-ratio sheets of molten metal must accommodate the tendency of the free surface to undergo a Rayleigh–Taylor instability (20).



**Figure 7.** Typical construction of water-cooled coil system for levitation of molten metals. The upper coil is in series opposition to the lower coil. The molten metal assumes approximately the shape and position shown.

### Electromagnetic Casting

The basic principles of electromagnetic levitation can be applied to create vertical walls of molten metal. These principles have been applied for many years in the moldless casting of aluminum ingots, in which the top of the ingot is surrounded by an electromagnetic fence and the molten metal is cooled from below. The ingot is moved downward and new molten metal is added to the top at the same rate as the metal solidifies. The advantage to this system is that the outer skin of the ingot is free of contamination and mold marks. Recently, similar principles have been applied to create electromagnetic edge dams for twin-roll casting of steel.

### Maglev Transport

Maglev transport involves the levitation of vehicles by one of the magnetic levitation principles discussed earlier. There is then no mechanical contact between the vehicle to be transported and the guideway that directs its travel. This application has been considered primarily for high-speed transport of people via trains, where damage to the tracks by a high-speed wheel-on-steel-rail system can be significant. However, maglev transport as an application is also being applied in clean-room environments where electronic dimensions are becoming increasingly smaller and contamination caused by rubbing or rolling contact cannot be tolerated.

The history of maglev transport has been detailed in Ref. 2 and several of the reading list entries. As early as 1907 Robert Goddard, better known as the father of modern rocketry, but then a student at Worcester Polytechnic Institute, published a story in which many of the key features of a maglev transportation system were described. In 1912, a French engineer Emile Bachelet proposed a magnetically levitated vehicle for delivering mail. His vehicle was levitated by copper-wound electromagnets moving over a pair of aluminum strips. Because of the large power consumption, however, Bachelet's proposal was not taken very seriously, and the idea lay more or less dormant for half a century.

In 1963 J. R. Powell, a physicist at Brookhaven National Laboratory, suggested using superconducting magnets to levitate a train over a superconducting guideway. Powell and G. R. Danby proposed in 1967 a system that used a less expensive conducting guideway at room temperature. Later, they conceived the novel idea of a "null-flux" suspension system that would minimize the drag force and thus require much less propulsion power (13).

During the late 1960s, groups at the Stanford Research Institute and at Atomic International studied the feasibility of a Mach-10 rocket sled that employed magnetic levitation. The maglev principle was later applied to high-speed trains by Coffey et al. and Guderjahn. In 1972 the group at Stanford Research Institute constructed and demonstrated a vehicle that was levitated with superconducting magnets over a continuous 160-m-long aluminum guideway. The vehicle, weighing several hundred kilograms, demonstrated both passive and active electromagnetic damping of the vehicle's motion (21).

At about the same time, a team from MIT, Raytheon, and United Engineers designed the magneplane system, in which lightweight cylindrical vehicles, propelled by a synchronously traveling magnetic field, travel in a curved aluminum trough. One advantage of the curved trough is that the vehicle is free

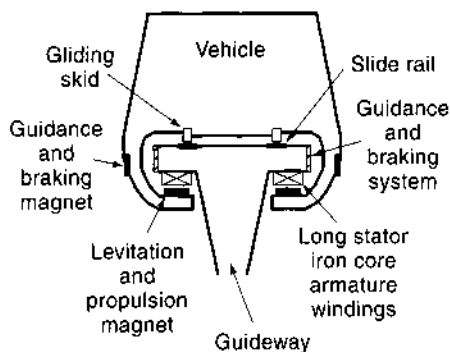
to assume the correct bank angle when negotiating curves, but the guideway itself is banked at only approximately the desired angle. The magneplane concept was tested with a 1/25-scale model system that used both permanent magnets and superconducting coils for levitation above a 116-m-long synchronized guideway.

Research groups at the Ford Motor Company Scientific Laboratories (12,22), the University of Toronto, and McGill University (23,24) carefully studied magnetic levitation and electromagnetic propulsion; although they did not construct test vehicles, these groups contributed immensely to our understanding of the basic physics and engineering principles involved. The initial research effort in the United States ended about 1975. Interest revived again around 1989, when four major conceptual designs were funded for several years by the federal government (25), after which government funding again disappeared.

Maglev systems have been studied in several other countries, most notably Japan, Germany, and the United Kingdom. The only maglev train ever in commercial service was a 600-m-long route in Birmingham, England that connected the airport to a conventional rail line. This service operated from 1984 to 1995. Its service was terminated mainly because of lack of parts in a one-of-a-kind technology.

Research in Germany and Japan has continued to the present, and full-scale vehicles have been tested in both countries. The two main maglev technologies that are being pursued at this time may be broadly classified as electromagnetic systems (EMSs), in which active feedback and electromagnets are used, and electrodynamic systems (EDSs), in which repulsive forces are generated by eddy currents that are induced in the guideway by the passage of a superconducting magnet.

EMSs depend on the attractive forces between electromagnets and a ferromagnetic (steel) guideway, as shown in Fig. 8. This system is under intensive investigation in Germany. The magnet-to-guideway spacing must be small (only a few centimeters at most), and this requires that the straightness tolerances of the guideway must be relatively small. On the other hand, it is possible to maintain magnetic suspension even when the vehicles are standing still, which is not true for EDSs. In the system shown in Fig. 8, a separate set of electromagnets provides horizontal guidance force, and the levitation magnets, acted on by a moving magnetic field from the guideway, provide the propulsion force. The German Transrapid TR-07 vehicle is designed to carry 200 passengers at a maximum speed of 500 km/h. The levitation height is 8 mm, and power consumption is estimated to be 43 MW at 400 km/h. A full-scale prototype



**Figure 8.** Schematic diagram of Transrapid maglev system (Germany).

of the Transrapid has run for many years on a 30 km test track in Emsland, Germany. Plans call for the Transrapid to enter commercial service in 2005 on a 292 km route that connects Hamburg to Berlin.

EDSs depend on repulsive forces between moving magnets and the eddy currents they induce in a conducting aluminum guideway or in conducting loops, as shown in Fig. 6. This system is being intensively investigated in Japan. The repulsive levitation force is inherently stable with distance, and comparatively large levitation heights (20 cm to 30 cm) are attainable by using superconducting magnets. Various guideway configurations, such as a flat horizontal conductor, a split L-shape conductor, and an array of short-circuit coils on the sidewalls, have been investigated. Each has its advantages and disadvantages. The proposed Japanese high-speed maglev system involves the use of interconnected figure-8 ("null-flux") coils on the sidewalls. The null-flux arrangement tends to reduce the magnetic drag force and thus the propulsion power needed. A prototype of this system has been operated for many years on a 7-km-long test track in Miyazaki, Japan. The Japanese currently plan to construct a commercial version that will operate between Tokyo and Osaka, with the first part of this route, a 42-km-long test track, in the Yamanashi prefecture. A three-car test train achieved the design speed of 550 km/h on this test track in December 1997.

### Magnetic Bearings

In active magnetic bearings (AMB), the attractive force between an energized magnet and a ferromagnetic body acts to achieve levitation. Typically, the magnet is an electromagnet, and the rotor is a ferromagnet. Such a system is inherently unstable, so, depending on the position of the rotor, active feedback is used to modulate the field of the electromagnet.

For decades, AMBs have been used in industrial applications; recently they have been used in oilfree turbomolecular vacuum pumps, compressors for gas pipelines, machine spindles for cutting operations, x-ray tube rotating anodes, pumps for cryogenics, turbo expanders in air separation plants, momentum wheels for satellite applications, flywheel prototypes, and sensitive pressure gauges. The three principal advantages of an AMB are: (1) absence of lubrication oil (required in most mechanical bearings); (2) the ability to attain higher speeds without the heating caused by mechanical friction; and (3) the ability to change in real time the stiffness and damping of the bearing to meet operating requirements.

In the basic setup of an AMB, position signals from gap sensors are used by a controller/power amplifier to set the appropriate currents and voltages of the electromagnets in such a way that stable levitation is achieved. In the example of an iron rotor below a ferromagnet, if the rotor starts to fall so that the gap becomes too large, the current in the magnet is increased to increase the attractive force. Similarly, if the gap becomes too small, the current is decreased. Control becomes effective when the bearing is activated and does not require any motion to levitate.

An advantage of the AMB is that it can adapt to operating conditions and communicate with its environment. Within certain limits, the AMB can be set to operate with arbitrary stiffness and damping, and the gap between the rotor and electromagnet can be changed. A rigid rotor can be made to rotate about its principal axis without transmitting vibrations to the

foundation. By changing the stiffness and damping, the rotor is capable of easily crossing the critical speeds of the bearings. In addition, bending vibrations of an elastic rotor can be significantly reduced by an appropriately designed active control loop.

Analog controllers were used in early AMBs; however, the tendency now is toward digital controllers. The design of early AMB systems was usually based on proportional-integral-differential (PID) control algorithms. Generally, the feedback law is chosen to be linear and is developed on the basis of the system linearized about the static equilibrium position. Recent investigations have shown that there is performance merit in the use of certain nonlinear feedback laws. Typically, switching amplifiers are used to power the coils. These amplifiers are basically transistor switches that connect the coils to a dc voltage. The transistors are either on or off and produce a square wave voltage on the coil, with the result that a sawtooth current is produced in the inductive load. By controlling the on and off times, most current wave forms can be produced in the coils. The disadvantage of the switching amplifiers is the oscillation in the current, which causes remagnetization loss in the magnetic bearing. However, the shorter the switching period, the weaker these oscillations.

Typically, a high-frequency carrier signal is used in the position sensor, which provides a linear signal versus the rotor location. Optical sensors, inductive displacement sensors, capacitance sensors, Hall effect sensors, and eddy current sensors are commonly used. The sensors must be contact free and capable of measuring a rotating surface. The surface quality of the rotor and the homogeneity of the material at the sensor will influence the measurements. A bad surface will produce noise disturbances, and geometry errors may cause disturbances in the rotational frequency or in multiples thereof. Training of the sensors while in place may alleviate some of these noise problems. In this technique, the rotor is rotated at some fixed speed, the surface is measured over many revolutions, and the average of the measurements for a given set of angles is stored in a computer memory so the sensor remembers the surface. The set of data serves as a background, which is subtracted from the measurement during regular operation.

A new type of AMB, currently under development, does not involve the use of sensors. Instead of the usual current amplifier, a voltage amplifier is used. The current in the amplifier is measured and a control algorithm can be developed so that the rotor can be stabilized from this measurement.

One of the fundamental limitations of an AMB is that the bearing force slew rate is limited because the magnet coils are inductive and the power supply to the driving amplifier has a maximum voltage. The slew rate limitation causes the bearing force to change more slowly than the control signal demands and thus introduces a phase lag. Another way to state this is that the inductance of the coils and the maximum voltage provide an upper limit to the stiffness of the system.

In principle, the current requirements of the coils of an AMB can be made arbitrarily small if the sensors have a high enough resolution and the sensor and power electronics are fast enough to respond. That is, less force is needed (and therefore less current) to return a rotor from a 1  $\mu\text{m}$  deviation from equilibrium than it does to return it from a 1 mm excursion. As sensors become faster and achieve greater resolution, and as power electronics become faster, one may expect that the losses in AMBs will decrease. At some point, the homogeneity of the rotor and

the magnet system will then provide the determining limit of these losses.

For a complete rotor, five degrees of freedom must be controlled. If magnetic biasing is used, the system can often be designed so just one degree of freedom need be controlled. Because an AMB typically provides stable stiffness in only one direction, it may be appropriate to design the magnetic bias with a rather large stiffness that provides stability in all directions but one. Then, only one AMB controller is required to stabilize against movement in the unstable direction.

In any of the bearing systems discussed in this section, considerable advantage in terms of reduced losses can be achieved by using a magnetic bias to take up most of the bearing load. (For example, see discussion of Fig. 1.) The rotational losses of a permanent magnet pair generally consist of eddy currents induced in one of the magnets by the azimuthal field inhomogeneity of the other. These losses are typically smaller than those associated with the bearing that makes up the stabilizing part of the system. Low-loss magnetic biasing can also be accomplished with a combination of permanent magnet and ferromagnet.

## BIBLIOGRAPHY

1. T. B. Jones, M. Washizu, and R. Gans, Simple theory of the Levitron, *J. Appl. Phys.*, **82**: 883–888, 1997.
2. T. D. Rossing and J. R. Hull, Magnetic levitation, *Phys. Teacher*, **29** (9): 552–562, 1991.
3. S. Earnshaw, On the nature of molecular forces which regulate the constitution of the luminiferous ether, *Trans. Cambridge Philos. Soc.*, **7**: 97–114, 1842.
4. W. Braunbek, Freischwebende korper im elektrischen und magnetischen Feld, *Zeit. Physik*, **112**: 753–763, 1939.
5. A. H. Boerdijk, Technical aspects of levitation, *Phillips Res. Rep.*, **11**: 45–56, 1956.
6. V. Arkadiev, Hovering of a magnet over a superconductor, *J. Phys. USSR*, **9**: 148, 1945. V. Arkadiev, A floating magnet, *Nature*, **160**: 330, 1947.
7. S. Evershed, A frictionless motor meter, *J. Inst. Electr. Eng.*, **29**: 743–796, 1900.
8. J. C. Maxwell, *A Treatise on Electricity and Magnetism*, Vol. 2, Oxford: Clarendon Press, 1891. Reprinted by Dover, New York 1954.
9. J. C. Maxwell, On the induction of electric currents in an infinite plane sheet of uniform conductivity, *Proc. Roy. Soc. London A*, **20**: 160–168, 1872.
10. W. M. Saslow, How a superconductor supports a magnet, how magnetically 'soft' iron attracts a magnet, and eddy currents for the uninitiated, *Am. J. Phys.*, **59**: 16–25, 1991.
11. W. M. Saslow, On Maxwell's theory of eddy currents in thin conducting sheets and applications to electromagnetic shielding and MAGLEV, *Am. J. Phys.*, **60**: 693–711, 1992.
12. J. R. Reitz, Forces on moving magnets due to eddy currents, *J. Appl. Phys.*, **41**: 2067–2071, 1970.
13. J. R. Powell and G. T. Danby, Magnetic suspension for levitated tracked vehicles, *Cryogenics*, **11**: 192–204, 1971.
14. R. H. Frazier, P. J. Gilinson, Jr., and G. A. Oberbeck, *Magnetic and Electric Suspensions*, Cambridge: MIT Press, 1974.
15. M. V. Berry, The Levitron<sup>TM</sup>: An adiabatic trap for spins, *Proc. Roy. Soc. London A*, **452**: 1207–1220, 1996.
16. M. D. Simon, L. O. Heflinger, and S. L. Ridgway, Spin stabilized magnetic levitation, *Am. J. Phys.*, **65**: 286–292, 1997.

17. M. V. Berry and A. K. Geim, Of flying frogs and levitrons, *Eur. J. Phys.*, **18**: 307–313, 1997.
18. J. R. Hull, Efficiency of passive magnetic-confinement techniques for rapidly rotating rings, *J. Appl. Phys.*, **58**: 3594–3600, 1985.
19. H. Tadano et al., Levitational melting of several kilograms of metal with a cold crucible, *IEEE Trans. Magn.*, **30**: 4740–4743, 1994.
20. J. R. Hull, D. M. Rote, and T. Wiencek, Magneto-hydrodynamic stability in the electromagnetic levitation of horizontal molten metal sheets, *Phys. Fluids A*, **1**: 1069–1076, 1989.
21. H. T. Coffey et al., Dynamic performance of the SRI maglev vehicle, *IEEE Trans. Magn.*, **10**: 451–457, 1974.
22. R. H. Borcherts et al., Baseline specifications for a magnetically suspended high-speed vehicle, *Proc. IEEE*, **61**: 569–578, 1973.
23. D. L. Atherton and A. R. Eastham, Flat guidance schemes for magnetically levitated high-speed ground transport, *J. Appl. Phys.*, **45**: 1398–1405, 1974.
24. D. L. Atherton and A. R. Eastham, Superconducting maglev and LSM development in Canada, *IEEE Trans. Magn.*, **11**: 627–632, 1975.
25. H. T. Coffey, U.S. Maglev: Status and Opportunity, *IEEE Trans. Appl. Supercond.*, **3**: 863–868, 1993.

### Reading List

- H. Bleuler, A survey of magnetic levitation and magnetic bearing types, *JSME Int. J. Ser. III.*, **35**: 335–342, 1992.
- B. V. Jayawant, Electromagnetic suspension and levitation, *Rep. Prog. Phys.*, **144**: (1981). Also Electromagnetic suspension and levitation techniques, *Proc. Roy. Soc. London A*, **416**: 245–320, 1988.
- E. R. Laithwaite (ed.), *Transport without Wheels*, London: Elek Science, 1977.
- F. C. Moon, *Superconducting Levitation*, New York: Wiley, 1994.
- W. A. Pfeifer, Levitation melting: A survey of the state of the art, *J. Met.*, **17**: 487–493, 1965.
- R. G. Rhodes and B. E. Mulhall, *Magnetic Levitation for Rail Transport*, Oxford: Clarendon Press, 1981.
- T. D. Rossing and J. R. Hull, Magnetic levitation comes of age, *Quantum* **5** (4): 22–27 March/April, 1995.
- G. Schweitzer, H. Bleuler, and A. Traxler, *Active Magnetic Bearings*, Hochschulverlag AG an der ETH Zurich, 1994.

JOHN R. HULL  
Argonne National Laboratory

## MAGNETIC MATERIALS

### HISTORICAL BACKGROUND

Magnetic materials have been known since ancient times—for example, in 380 B.C.E. Plato wrote (1) of the “stone which Euripides calls a magnet,” which we infer was  $\text{Fe}_3\text{O}_4$ , now known as magnetite. The scientific quality of magnetism studies abruptly and dramatically jumped with the publication in 1600 by Gilbert of the classic text *De Magnete* (2). Quantitative measurements of magnetic materials were enabled by the 1820 discovery by Oersted that an electric current creates a magnetic field. In 1846 Faraday made systematic studies of the attraction and repulsion of materials in a gradient field and classified materials as *diamagnetic* if they are repelled by a region of increased flux density and *paramagnetic* if they

are attracted. To this we add *ferromagnetic* (strongly magnetic, like iron) to form the set of three basic classes of magnetic response.

Since the early part of the twentieth century, magnetic materials have been the subject of deep and broad research and development due to their economic and scientific importance, and much of our knowledge is mature. Nevertheless, startling discoveries continue to be made, such as the discovery of Nd–Fe–B permanent magnets and the “giant magnetoresistance” effect in thin-film multilayers.

### MAGNETIC FIELDS AND THE MAGNETIC RESPONSE OF MATERIALS

The magnetic properties of matter may be viewed as a response to an applied stimulus, namely the magnetic field strength  $\mathbf{H}$ . The macroscopic response of a material is given by its magnetization,  $\mathbf{M}$ , and the overall field is the sum of the two, called the magnetic induction  $\mathbf{B}$ . In a vacuum the magnetization is strictly zero. For this article we adopt SI units, so we have  $\mathbf{B} = \mu_0 \mathbf{H}$  in a vacuum, where  $\mathbf{B}$  is measured in tesla ( $\text{Wb}/\text{m}^2$ ),  $\mathbf{H}$  is measured in amperes per meter, and by definition  $\mu_0 = 4\pi \times 10^{-7} \text{ H/m}$ . The magnetic response adds directly to the applied field, giving  $\mathbf{B} = \mu_0(\mathbf{H} + \mathbf{M})$ .

The issue of units in magnetism is perennially vexing. In the past, cgs (Gaussian) units have been commonly used by scientists working with magnetic materials. In that system,  $\mathbf{B}$  is measured in gauss,  $\mathbf{H}$  in Oersteds, and  $\mathbf{M}$  in  $\text{emu}/\text{cm}^3$ , where emu is short for the uninformative name *electromagnetic unit*. The constitutive relation in Gaussian units is  $\mathbf{B} = \mathbf{H} + 4\pi \mathbf{M}$ . Important conversion factors to keep in mind are  $10^4 \text{ Gauss} = 1 \text{ T}$  and  $12.5 \text{ Oe} = 1 \text{ kA/m}$ . A definitive discussion of units and dimensions is given in the Appendix of Jackson's *Classical Electrodynamics* (3).

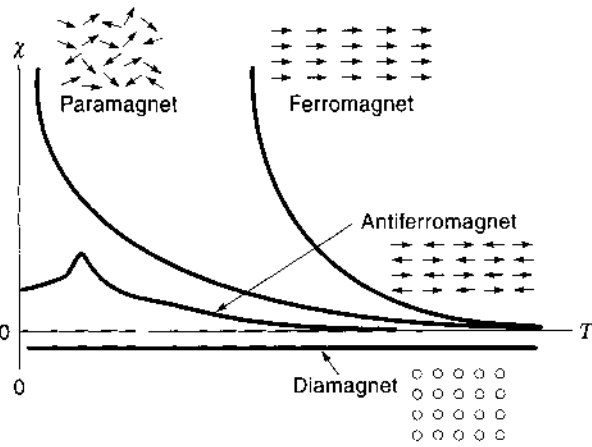
### TYPES OF MAGNETIC MATERIALS: TAXONOMY

#### Basic Families

Two of the basic families of magnetic materials involve a highly linear response (i.e.,  $\mathbf{M} = \chi \mathbf{H}$ , where  $\chi$  is defined as the magnetic susceptibility). The main magnetic response of all materials is due to the magnetic moment of individual electrons, a property directly connected to their spin. The moment of a single electron is 1 Bohr magneton,  $\mu_B = 1.165 \times 10^{-29} \text{ Wb}\cdot\text{m}$ . Due to the Pauli principle, in many cases the electrons in an atom are precisely paired with oppositely directed spins, leading to an overall cancellation. Nevertheless, a magnetic response can be discerned in all materials, as observed by Faraday.

#### Diamagnetism

Diamagnets have a negative value for  $\chi$ , that is, the induced moment is opposite to the applied field. The susceptibility is temperature independent and typically small (see Fig. 1). Diamagnetism is due to the effect of a magnetic field on orbital motion of paired electrons about the nucleus (superficially comparable to Lenz's law). The diamagnetic susceptibility of most materials is very small—in the vicinity of  $-1 \times 10^{-5}$ . A tabulation of diamagnetic susceptibilities of various atoms, ions, and molecules is given by Carlin (4).



**Figure 1.** Schematic temperature dependence of the susceptibility of a diamagnet, paramagnet, ferromagnet, and antiferromagnet.

A large negative magnetic susceptibility is characteristic of only one class of materials (namely, superconductors). A Type I superconductor in the Meissner state exhibits complete exclusion of magnetic flux from the interior of the sample,  $\mathbf{M} = -\mathbf{H}$ , or  $\mathbf{B} = 0$ . Superconductors can also exhibit partial flux penetration,  $0 < B < \mu_0 H$ . In both cases the spectacular observation of stable levitation is possible, something that cannot be achieved using only materials with  $\chi > 0$  (as proven by Earnshaw's theorem). Note that stable levitation is possible even for bodies that are only weakly diamagnetic given a sufficiently large magnetic field gradient (5).

**Paramagnetism**

Paramagnets have a positive value for  $\chi$ , that is, the induced moment is in the same direction as the applied field. Paramagnetism is chiefly due to the presence of unpaired electrons—either an overall odd number of electrons or an unfilled inner shell. Nuclei can also show paramagnetism, though typically of an extremely small magnitude. The electron gas of a metal is also usually slightly paramagnetic, though exchange coupling can sometimes lead to ordering (e.g., ferromagnetism). Independent unpaired electrons give each atom or molecule a small permanent dipole moment, which tends to be aligned by an external magnetic field. Langevin showed that thermal energy disrupts this alignment, leading to a susceptibility  $\chi = Nm^2/3k_B T$ , where  $N$  is the density of dipoles,  $m$  is the moment of each dipole,  $k_B$  is the Boltzmann constant, and  $T$  is the absolute temperature. Curie and Weiss found that the temperature in this formula should be replaced by  $T \rightarrow (T - T_c)$  for materials with an ordering temperature  $T_c$  (the “Curie temperature.”) The paramagnetic susceptibility of a material can give important insights into its chemistry and physics, but it is an effect of limited engineering significance at present.

**Ferromagnetism**

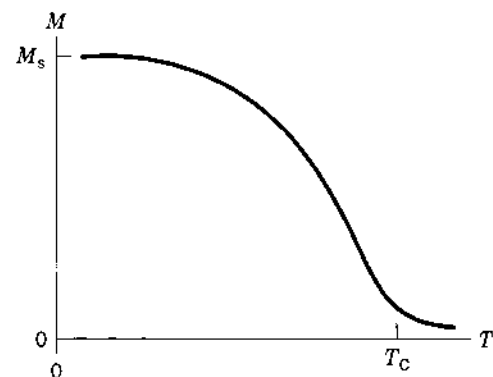
Ferromagnetism is the spontaneous magnetic ordering of the magnetic moments of a material in the absence of an applied magnetic field. Nearly all technologically important magnetic materials exhibit some form of ferromagnetism. In such materials, the magnetic moments of electrons couple together, so that they respond collectively. In this manner it is possible for all of the magnetic moments in an entire sample to point in

the same direction, potentially giving a very strong effect. The details of how the individual moments couple with each other can be understood in terms of quantum mechanics. There are three types of “exchange” interaction generally found:

- The first is direct exchange, in which an unpaired electron on one atom interacts with other unpaired electrons on atoms immediately adjacent via the Coulomb interaction. This is the strong mechanism that dominates in most metallic magnetic materials, such as Fe, Ni, Co, and their alloys. It results in a positive exchange energy, so the spins on adjacent atoms tend to align parallel.
- The second is indirect exchange, or superexchange, in which the moment of an unpaired electron on one atom polarizes the (paired) electron cloud of a second atom, which in turn interacts with the unpaired electron on a third atom. This is the mechanism that dominates in most oxide materials, such as ferrites. For example, in  $\text{Fe}_3\text{O}_4$  the Fe ions (with unpaired electrons) interact through O ions (which have only paired electrons). Superexchange creates a negative exchange energy.
- Finally, there is the possibility of interaction between electrons that are not localized but can move freely as in a metal. This interaction, known as the RKKY interaction after its discoverers (Ruderman, Kittel, Kasuya, and Yoshida), is usually weaker than direct exchange. It plays an important role in the behavior known as giant magnetoresistance and can result in either a positive or negative exchange energy.

The main properties that characterize ferromagnetic materials are the Curie temperature,  $T_c$ , the saturation magnetization,  $M_s$ , the magnetic anisotropy energy,  $K$ , and the coercive field,  $H_c$  (see Fig. 2). The first two are intrinsic to a material. The third has both intrinsic and extrinsic factors. The last is extrinsic and depends on the form (microstructure, overall shape, etc.) of the material and will be discussed later.

- The exchange interaction that leads to ferromagnetism can be disrupted by thermal energy. At temperatures above  $T_c$ , the disruption is so great that the ferromagnetism ceases, and the material exhibits only paramagnetism. Thus  $T_c$  measures the magnitude of the exchange coupling energy. For example, the  $T_c$  of Fe is  $770^\circ\text{C}$  while for Co,  $T_c = 1115^\circ\text{C}$ ,



**Figure 2.** Schematic temperature dependence of the saturation magnetization,  $M_s$ , for a ferromagnet.

and for Ni,  $T_c = 354^\circ\text{C}$ . The ferromagnetic transition is a second-order phase transition, which means that the order parameter (magnetization) increases continuously from zero as the temperature is lowered below  $T_c$ .

- The saturation magnetization is the macroscopic magnetic moment of all of the spins averaged over the volume of the sample. Thus, in a material with many unpaired electrons per atom,  $M_s$  will be large (e.g., Fe with  $\mu_0 M_s = 2.16\text{ T}$  at room temperature). Conversely,  $M_s$  will be much smaller in materials that also contain nonmagnetic atoms or ions (e.g.,  $\text{Fe}_3\text{O}_4$  with  $\mu_0 M_s = 0.60\text{ T}$  at room temperature).
- The electron spins couple weakly to their orbital motion in a process known as spin-orbit coupling, a relativistic effect. As a result, the energy of the system depends on the orientation of the spins (i.e., the magnetization) with respect to the orbitals of the atoms (i.e., the orientation of the magnetization to the crystal lattice). It leads to magnetic anisotropy—that is, the energy of the system depends on the orientation of the magnetization with respect to the sample. The direction along which the magnetic moment tends to lie is known as the easy axis. The magnitude of the anisotropy may be large, as in  $\text{SmCo}_5$  permanent magnets that strongly resist demagnetization with  $K \sim 10^7\text{ J/m}^3$ , or it may be quite small, as in the high-permeability materials  $\text{Ni}_{0.8}\text{Fe}_{0.2}$  (Permalloy) or  $\alpha\text{-Fe}_{0.80}\text{P}_{0.13}\text{C}_{0.07}$  (an amorphous alloy).
- Another source of anisotropy can arise from the shape of the specimen, or from the shape of individual grains within the specimen. This is a local magnetostatic effect, rather than an intrinsic effect, and is called shape anisotropy (see Fig. 3). It is an extremely important factor in any real application. Two extremes are illustrative: A long thin needle (i.e., an acicular particle) can be readily magnetized along its long axis but will require a large field to force the magnetization to be across a short axis. The magnitude of field required is  $H_a = M_s/2$  (i.e.,  $H_a = 8.5 \times 10^5\text{ A/m}$  for the case

of an Fe needle). A flat plate, on the other hand, will require twice that field, i.e.,  $H_a = M_s$  to magnetize it parallel to the normal.

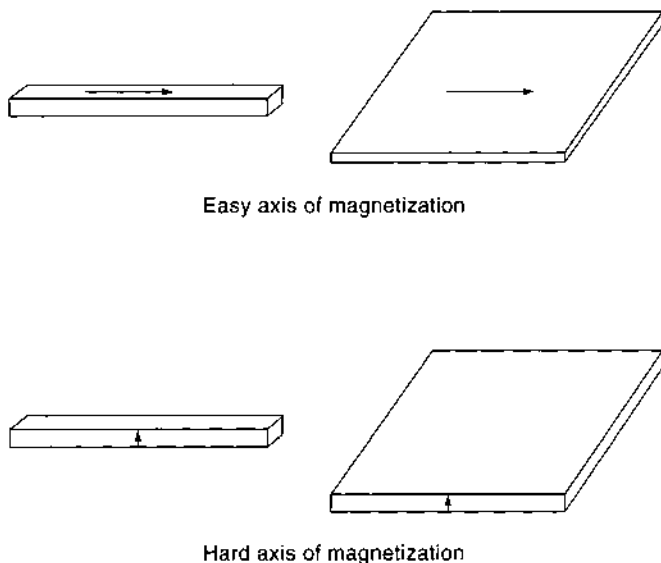
- A third source of anisotropy is due to the magnetostriction of magnetic materials, coupled with stresses in the material. Magnetostriction is the change in dimensions of a sample when the magnetization is aligned along various crystallographic directions; it occurs as a response that minimizes the magnetocrystalline energy. Conversely, when a sample is strained along some crystallographic direction, this contributes to the magnetic anisotropy. This is called stress anisotropy. It can be an important effect in low-anisotropy materials that are highly strained, such as almost all thin films.
- The various magnetic anisotropies that may exist in a material all act simultaneously. The best way to analyze their cumulative effect is in terms of the anisotropy energy, which is the sum of all of the energies arising from individual anisotropies. The details of this analysis can be complex; see Bozorth or Brailsford, listed in the Reading List, for examples and guidance.

Useful magnetic materials almost inevitably consist mostly of Fe, Co, or Ni or a combination of these three elements, because these are the elements that are ferromagnetic at room temperature and above. A great variety of other elements may be added to form alloys or compounds with specific useful properties, but inevitably a large fraction of Fe, Co, or Ni will be present. When a nonmagnetic metal is alloyed with these elements,  $M_s$  and  $T_c$  generally decrease rapidly due to dilution. For example, Fig. 4(a) (Bozorth pp. 308–309) shows the effect of alloying Ni with Cu (which together form a continuous solid solution), showing the monotonic decrease in  $M_s$  with increasing Cu content. Other effects may occur, such as band-structure effects or the formation of compounds, which will alter the trends with alloying [e.g., formation of  $\text{Fe}_3\text{Al}$ , as shown schematically in Fig. 4(b)].

Alloying with rare-earth metals is often used in cases where a high intrinsic anisotropy is desired, such as in permanent magnets. The lanthanide rare-earth metals are all highly magnetic due to unpaired electrons in the  $4f$ -shell (inner) orbitals. The  $T_c$  of these materials is below room temperature because the exchange interaction between inner orbitals of adjacent atoms is small, but the intrinsic anisotropy is generally large because the spin-orbit interaction is largest in atoms with high atomic number (and therefore highly relativistic orbitals). Adding a small amount of a rare earth can dramatically increase the magnetocrystalline anisotropy of an Fe-, Co-, or Ni-based compound, often with only a modest decrease in  $T_c$  and  $M_s$ . The modern “rare-earth” permanent magnetic materials use this effect, as in  $\text{SmCo}_5$  and  $\text{Nd}_2\text{Fe}_{14}\text{B}$ .

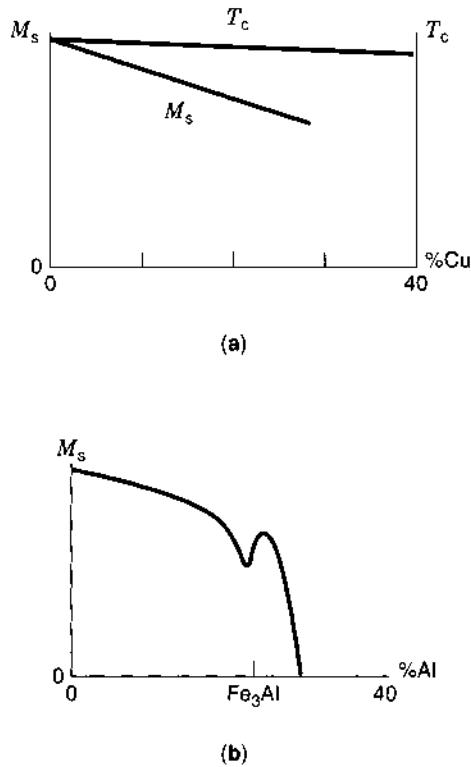
### Domains and $M$ - $H$ Loops

While positive exchange coupling tends to align all of the spins in the same direction, real materials generally exhibit this uniformly oriented state only if they are very small ( $<100\text{ nm}$ ). Larger samples “demagnetize” by breaking up into magnetic domains. In each domain the local magnetic moments are uniformly aligned, usually along an easy axis. The directions of magnetization of the various domains can balance such that the overall magnetization is zero and the magnetostatic energy



**Figure 3.** Shape anisotropy quantitatively describes the observation that needles and plates are most easily magnetized along a long dimension.



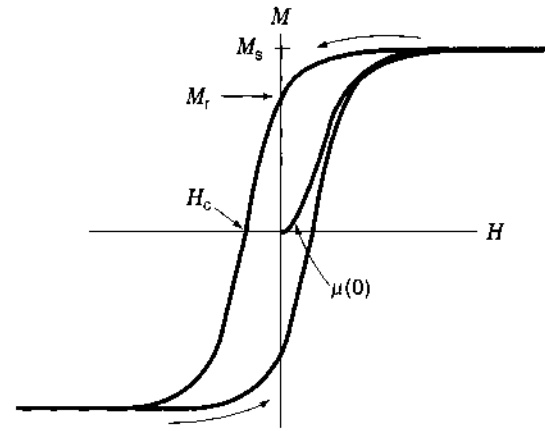


**Figure 4.** (a) Saturation magnetization and  $T_c$  for Fe–Cu alloys, normalized to the values for pure Fe. The monotonic decrease is typical of systems that form a continuous solid solution. (b) Saturation magnetization of Fe–Al compositions, normalized to the value for pure Fe. The anomalous behavior near the composition  $\text{Fe}_3\text{Al}$  (25% Al) is due to the formation of the  $\text{Fe}_3\text{Al}$  phase.

is small. When an external field is applied along an easy axis, domains aligned with the field tend to grow, while those antialigned tend to shrink. At high enough field the sample will be forced into a single-domain state, and the saturation magnetization will be observed. The formation of domains implies the presence of domain walls—boundaries between adjacent domains—that have increased exchange and anisotropy energies due to misalignment of neighboring spins. The density and orientation of domains in a sample is partly determined by energy balance between the domain wall and magnetostatic terms, but is also strongly affected by nonequilibrium considerations such as domain wall nucleation and pinning. In general, the growth and shrinking of domains (i.e., the motion of domain walls) dissipates energy, so the  $M$ – $H$  curve is hysteretic, as shown schematically in Fig. 5.

This hysteretic, sigmoidally shaped  $M$ – $H$  curve is very typical of ferromagnetic materials. Four important parameters are immediately evident from examination of the  $M$ – $H$  curve.

- First, the limiting magnetization is just  $M_s$ , the single most important measure of a ferromagnetic material.
- Second, the slope of the  $M$ – $H$  curve at  $M = 0$  is the small-signal permeability  $\mu(0)$ , which measures the responsiveness of the magnetic material to an external field when it is close to its demagnetized state. This parameter is particularly important for soft magnetic materials, which use the magnetic material to obtain a flux multiplication by the factor  $\mu(0)$ . This parameter is determined partly by



**Figure 5.** Schematic  $M$ – $H$  curve, showing saturation magnetization,  $M_s$ , remanent magnetization,  $M_r$ , coercive force,  $H_c$ , and initial permeability,  $\mu(0)$  (defined for an initially demagnetized sample, i.e., with  $H = 0$  and  $M = 0$ ).

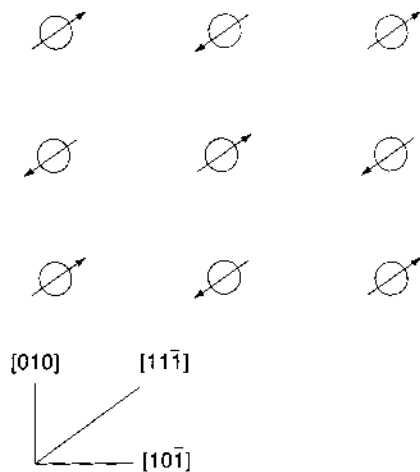
the magnetic anisotropy that is characteristic of the material but is also affected by factors that impede domain-wall motion, such as physical grain structure, microscopic inclusions, dislocations, or magnitude of the magnetocrystalline anisotropy.

- Third, the magnetization observed at zero field (after the sample has been fully magnetized) is called the remanence,  $M_r$ . This is an important parameter for permanent magnets, as it measures the magnitude of  $M$  available when the material is isolated. Note that the “squareness ratio,”  $M_r/M_s$ , is dominated by extrinsic aspects of the material, such as grain structure and defect, along with underlying anisotropies including the shape of the specimen.
- Fourth, the field required to reduce the external magnetization to zero (again, defined only after the sample has first been fully magnetized) is called the “intrinsic coercivity” or coercive field,  $H_c$ . At this field, the sample is in a multidomain state and the magnetizations from all of the various domains exactly cancel out. The coercive field is an important property for permanent magnets, as it measures the ability of a material to withstand the action of an external magnetic field, whether applied or self-generated. It is also determined mainly by extrinsic aspects of the material such as grain structure.

The interpretation of  $M$ – $H$  loops can often involve subtle aspects of the loop, including directional properties, the approach to saturation, possible nonsigmoidal curving, discrete jumps (known as Barkhausen jumps), and so on. These may reflect coherent rotation of spins in a domain when the external field is not aligned with an easy axis or may be due to subtleties of domain wall motion. Development of superior magnetic materials often involves intensive research into these issues, but usually the designers of devices need only focus on a few properties.

### Negative Exchange Interaction

The exchange interaction, as mentioned previously, need not be positive, inducing alignment of adjacent spins. When it is negative, adjacent spins will tend to align antiparallel. This can lead to a variety of behaviors depending on the structure of the material.



**Figure 6.** Antiferromagnetic structure of NiO, showing Ni atoms in the  $(\bar{1}01)$  plane. The spins are aligned along  $[1\bar{1}\bar{1}]$  directions as shown. The magnetic unit cell is twice the length of the crystallographic unit cell.

### Antiferromagnetism

The simplest configuration that can be obtained with a negative exchange energy is antiferromagnetism, in which the spins on adjacent sites in a unit cell cancel to give no net magnetic moment. A simple example is NiO, which forms in the rock salt (NaCl) structure (see Fig. 6). The ordering temperature for antiferromagnetic materials is called the Néel temperature,  $T_N$ , after the discoverer of antiferromagnetism, and is analogous to the Curie temperature of a ferromagnet. Above  $T_N = 250^\circ\text{C}$ , NiO is, of course, paramagnetic. In the antiferromagnetic state the susceptibility is not negative, as in the case of a diamagnet (which has no permanent dipoles) but is positive, small, and depends on the direction of the external field due to intrinsic magnetocrystalline anisotropy. The details of spin configurations and other properties of antiferromagnets can be very complicated. Antiferromagnetism is difficult to detect by conventional magnetic measurements. Neutron scattering mea-

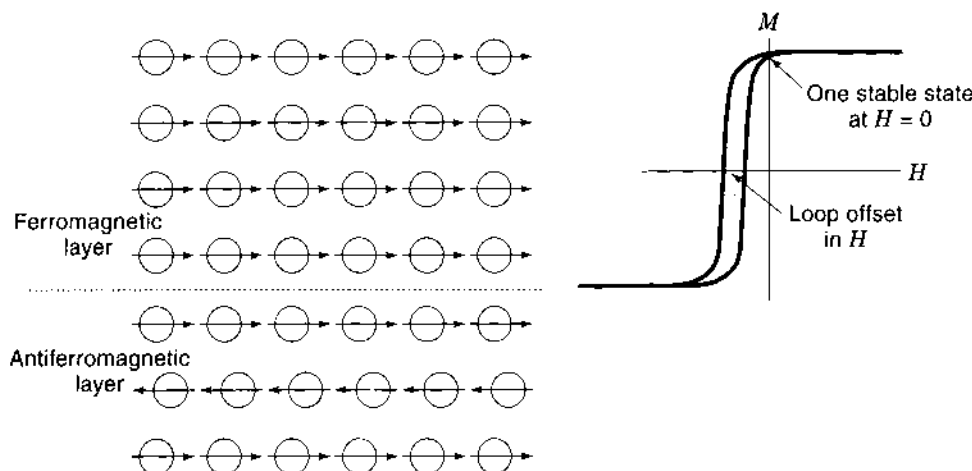
surements are typically required to confirm the existence of antiferromagnetism.

Antiferromagnetic materials have been known and understood since the work of Néel beginning in 1932, but there are presently no important applications of bulk antiferromagnetic materials. Thin films ( $\sim 1$  nm to 100 nm thick) of antiferromagnetic materials now play an important role in state-of-the-art magnetic recording, specifically in magnetoresistive read heads. The antiferromagnetic thin films are used to magnetically bias the magnetoresistive sensor using a phenomenon called exchange anisotropy: the surface interaction between a ferromagnetic and antiferromagnetic material in intimate contact (see Fig. 7). Since this is an interfacial phenomenon, its magnitude is only significant when the surface/volume ratio is high, as in a very thin film.

### Ferrimagnetism

In a compound with two magnetic sublattices and antiferromagnetic coupling, the magnetic moments of each sublattice will generally not cancel exactly. Then the material will exhibit an overall magnetization that in many regards will appear exactly like that of a ferromagnet, with a hysteretic  $M$ - $H$  loop, a coercivity, and a remanence. Such materials are called ferrimagnets, because the prototypical examples are ferrites. Some properties, such as the temperature dependence of the magnetization, can be radically different from those of ferromagnets. For example, the different temperature dependencies of the magnetization on two sublattices can sometimes lead to exact cancellation of the net magnetization at a particular temperature, called the compensation temperature,  $T_{\text{comp}}$  (often denoted  $T_c$ , which leads to confusion with the Curie temperature). At that temperature the material behaves as if it were an antiferromagnet.

While ferrimagnets behave in many ways like ferromagnets, the highest saturation flux density in ferrimagnets is typically only about 0.6 T, and they cost significantly more than iron or silicon iron. Their crucial advantage is that they are usually good insulators and therefore are useful at high frequencies



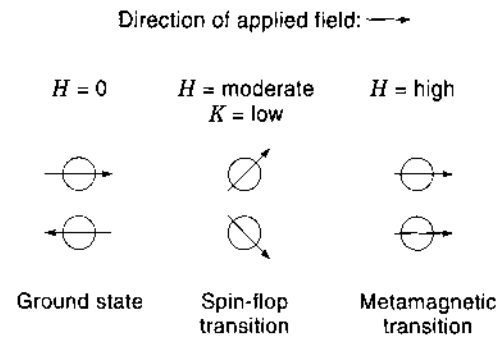
**Figure 7.** Schematic illustration of exchange anisotropy arising from interface coupling between an antiferromagnetic and ferromagnetic material. The schematic  $M$ - $H$  loop indicates that the loop is offset in  $H$  and that with no external field there is only one stable state (namely, the saturated state). That is, at  $H = 0$  there can be no domain structure.

due to low eddy-current losses. Three classes of ferrimagnetic materials are predominant in applications:

- Garnets have a generic formula of  $R_3\text{Fe}_5\text{O}_{12}$ , where  $R$  represents a lanthanoid (Sc, Y, or lanthanide rare earth). These compounds have a  $T_c$  around  $275^\circ\text{C}$  and a rather low saturation flux density at room temperature,  $B_s = 0.18\text{ T}$ . They have proven useful for bubble memories because high-quality single-crystal garnets can be prepared, and they continue to be used for UHF applications because they have particularly low losses in that frequency regime.
- Spinel ferrites are an especially large class of materials with a wide range of properties. The generic formula unit is  $AB_2O_4$ , where  $A$  is a divalent ion and  $B$  is a trivalent ion, usually  $\text{Fe}^{3+}$ . Most of the useful spinel ferrites are magnetically soft (that is, they have a low anisotropy energy and a high permeability). The prototypical spinel ferrite is  $\text{Fe}_3\text{O}_4$ , but Zn-substituted  $\text{MnFe}_2\text{O}_4$  and  $\text{NiFe}_2\text{O}_4$  are the soft ferrites used in most applications. Another extremely important ferrite is commonly used as a magnetic recording medium—namely,  $\gamma\text{-Fe}_2\text{O}_3$ , which is a modified spinel in which one in nine Fe sites is systematically vacant.
- Hexagonal ferrites are a much smaller class of materials, but this class includes the important ceramic permanent magnet materials. A typical formula unit for a hard hexagonal ferrite is  $\text{BaFe}_9\text{O}_{12}$ . These materials have a platelet-type growth habit with a very high uniaxial anisotropy and an easy axis normal to the platelet. This makes it difficult for the magnetization of a platelet to change, which accounts for the hard magnetic properties. The fact that these materials are insulating is often not an important issue since they are used to create a dc magnetic field.

When a magnetic dipole is aligned (e.g., by intrinsic anisotropy) along an axis and a radiofrequency (RF) field is applied perpendicular to that axis, the dipole does not respond simply by oscillating in the direction of the RF field, but it precesses around its axis in accordance with classical mechanics. The precession frequency is commonly expressed as  $\omega = \gamma H_{\text{an}}$ , where  $\gamma$  is the gyromagnetic constant [ $\gamma = 35\text{ kHz}/(\text{A}/\text{m})$  for most materials] and  $H_{\text{an}}$  is the anisotropy field. If the RF field is at exactly this frequency, the dipole can readily absorb energy from the field (and convert it into heat via coupling to the lattice). This phenomenon is known as ferromagnetic resonance, FMR, though it is most important in insulating ferrimagnets where eddy currents do not already dominate the losses. Above the FMR frequency the magnetic material has a nonmagnetic response.

At very high frequencies the response of ferrimagnets is not dominated by domain-wall motion, which is sluggish, but by coherent rotation of the spins in the sample. Then the permeability is given simply by  $\mu_c = B_s/H_{\text{an}}$ , so  $\omega_{\text{FMR}}\mu_c = \gamma B_s$ . This equation, known as Snoek's law, says that for a given material, a higher FMR frequency can only be obtained at the cost of a correspondingly smaller permeability. It is a basic limitation to the use of ferrimagnetic materials at frequencies above about 10 MHz. Other issues, such as domain-wall resonances, may reduce the maximum frequency even further.



**Figure 8.** The spin-flop transition and metamagnetic transition in antiferromagnetic materials. The spin-flop transition only occurs in materials with relatively low magnetic anisotropy energy,  $K$ , binding the spins to the easy axis (in this case, the horizontal axis).

### Metamagnetism

If a large enough magnetic field is applied to an antiferromagnet along an easy axis, the spins that are antialigned with the field will suddenly flip their orientation to achieve a lower energy state. That is, for a sufficiently high magnetic field,  $H$ , the magnetostatic energy  $m \cdot H$  (where  $m$  is the dipole moment of an individual atom) will inevitably outweigh the exchange energy. In some antiferromagnetic materials this flipping can be observed with achievable magnetic fields; it is then called metamagnetism. Note that in principle all antiferromagnets will exhibit this behavior at high enough field—the distinction is only in whether the required field can be produced in the laboratory.

Antiferromagnets with a relatively low anisotropy energy can exhibit an intermediate state between the antiferromagnetic and metamagnetic states as the field is increased. In this case, application of the field along the easy axis will cause the spins to reorient perpendicular to the magnetic field, and still in an approximately antiferromagnetic configuration, as shown schematically in Fig. 8. This transition is called spin flopping.

Neither metamagnetism nor the spin-flop transition are of practical significance in bulk applications of magnetism. However, the metamagnetic transition is an essential feature of the phenomenon of giant magnetoresistance (GMR), which is observed in metallic thin-film ferromagnet/paramagnet multilayers (vide infra).

### Spin Glass State

When a magnetic material has structural disorder, it is sometimes not possible for the exchange interaction among various neighbors to be satisfied, and no long-range orientational order (either ferromagnetic or antiferromagnetic) can be achieved. At low enough temperature such a "frustrated" material will achieve a quasi-ordered configuration in which the spins are static but aligned in random directions. This is the "spin glass" state. In a given sample, the magnetic properties are found to be history dependent: For example, the saturation magnetization depends on whether the sample was cooled in a magnetic field or in zero field. So-called spin glass materials should not be confused with the metallic glasses discussed later. The nature of the spin glass state has been a productive area of study for physicists for many years, but the phenomenon has no current engineering significance (6).

A related concept is that of “geometrical frustration,” which occurs in materials that have triangular site coordination and that therefore are frustrated even in a perfectly ordered material. These materials are presently being explored intensively by physicists, but also have no current engineering significance (7).

### Double-Exchange Materials

Along with the exchange mechanisms listed previously, there is a fourth mechanism, double exchange, that is relevant only to a small class of materials. The prototype material is  $(\text{La,Ca})\text{MnO}_3$ , a perovskite oxide, in which the Ca substituted for La acts as an electron donor. Electrons hopping from one Mn atom to another do not change their spin orientation during the hop, so the Mn atoms orient ferromagnetically. The most interesting property of these materials is observed in the paramagnetic state: In zero field the Mn spins are randomly oriented, inhibiting electron hopping and yielding a high resistivity. When the Mn spins are partially aligned by applying a large magnetic field, the hopping probability is enhanced, and the resistivity decreases dramatically. At an optimum temperature (near  $T_c$ ), a resistivity decrease by a factor of  $10^5$  to  $10^6$  has been observed at a field of 4 MA/m, a result that has led the effect to be known as colossal magnetoresistance (CMR). The effect is colossal at low temperatures and for only a small temperature range, and it requires very large fields to be applied, so it has not proven useful for engineering applications at present.

### Superparamagnetism

In small single-domain particles of ferromagnetic material, the magnetization tends to align along the easy axis, and the energy barrier required to move the magnetization is on the order of  $KV$ , where  $K$  is the anisotropy energy density and  $V$  is the volume of the particle. For the simple case of uniaxial anisotropy,  $\Delta E = KV$ . If the thermal energy  $k_B T$  is greater than this energy, then fluctuations in the orientation of the magnetization will lead to zero spontaneous magnetization and a response to an applied field that is analogous to the local moments of a normal paramagnetic material. The effect is seen, for example, in Co particles smaller than about 7.5 nm. The moment of a Co sphere of diameter 7.5 nm is roughly 20,000 times larger than the moment of single Co atom, so the effect is called superparamagnetism. While superparamagnetism is not technologically useful itself, it does represent an important limitation to the particle size of future magnetic recording media, and is therefore being intensively investigated.

### Thin Films

Thin films of magnetic materials behave exactly like bulk magnetic materials in most respects, albeit in profoundly different regimes for some parameters. For example, thin films have a high demagnetization factor normal to the film and essentially zero demagnetization factor in the plane of the film. Thus it is generally very hard to magnetize a thin film normal to the plane but fairly easy to move the magnetization in the plane. A simple consideration such as this can have wide-ranging implications, from an increased FMR frequency to gross effects on the structure of domain walls (i.e., the transition from conventional Bloch walls to Néel walls as the film thickness is decreased below about 100 nm). Thin films also tend to

have stresses that are extremely high compared to stresses in bulk materials—500 MPa values are not uncommon. These stresses couple to the magnetostriction of the material to create a stress anisotropy that can strongly influence the magnetic behavior.

Thin films are used in a wide variety of applications, the most important of which are as media in hard disks and magneto-optic disks, miniature electromagnets in hard disk write heads, and magnetoresistive sensors in hard disk read heads.

Thin films inherently possess a unique direction, the growth direction. This is usually the normal, although it can be oblique if the incident atomic flux used to grow the film arrives from an oblique angle. For some materials the growth direction directly leads to a large intrinsic uniaxial anisotropy. For example, in amorphous Tb-Fe one might expect that there would be no anisotropy at all. Instead, films grown with the incident Tb and Fe atoms arriving essentially perpendicular to the substrate exhibit a large intrinsic uniaxial anisotropy oriented along the normal, and with a sense that leads to a perpendicular easy axis. The anisotropy is sufficient to overcome demagnetization, so domains form in which the magnetization is oriented perpendicular to the film. Such materials are used in magneto-optic recording (vide infra) (8).

Some behaviors seen in thin films are either absent or not commonly observed in bulk magnetic materials. An important example is the phenomenon of so-called giant magnetoresistance in thin-film multilayers. These multilayers are typically formed by sequentially depositing metallic ferromagnetic and paramagnetic layers, each  $\sim 1$  nm to 3 nm thick, using sputtering or evaporation in a high-vacuum chamber. Between two and a hundred layers might be built up in this way. The ferromagnetic layers couple with each other by the RKKY interaction through the paramagnetic metal, so depending on the thickness of the paramagnetic layer, the interlayer coupling may be antiferromagnetic or ferromagnetic. For example, a film consisting of 100 repeats of 1.0 nm-thick Co adjacent to 0.6 nm-thick Cu exhibits an antiferromagnetic state at zero field: Even though each Co layer is individually ferromagnetic, alternate layers have oppositely directed magnetizations. When a moderate field,  $H$ , is applied, the magnetizations of all of the layers align with the external field, producing a metamagnetic transition.

Baibich et al. (9) discovered the most interesting aspect of the metamagnetic transition in metallic multilayers: the effect it has on the resistivity of the sample. The aligned state has a greatly reduced resistivity compared to the antialigned state. The magnetoresistance ratio  $(R(H=0) - R(H=H_s))/R(H=H_s)$  can be as high as  $\sim 100\%$  depending on the choice of materials. This is far greater than the highest normal magnetoresistance observed in any material at room temperature; hence the name giant magnetoresistance. It is attributed to spin-dependent scattering of electrons, which is enhanced when magnetizations of adjacent layers are antialigned. This effect is crucial for the highest-performance magnetic disk read heads being currently designed (vide infra) (10).

## MAGNETIC MATERIALS USED IN APPLICATIONS

Useful magnetic materials are often divided into three categories:

- Soft magnets, in which the magnetization is readily changed with an external field, thereby providing a flux-multiplying effect
- Hard magnets (permanent magnets), which have high coercive fields and which therefore resist demagnetization by stray fields including their own
- Magnet recording media, which combine aspects of softness and hardness

### Soft Magnetic Materials

Soft magnetic materials are used in applications such as transformers and inductors. An obvious example would be the iron, known as electrical steel, used in transformers for inexpensive power supplies. At frequencies above about 10 kHz, eddy currents limit the use of metallic magnetic materials, so high-resistivity ferrites such as  $(\text{Mn,Zn})\text{Fe}_2\text{O}_4$  are used.

**Permeability.** An important property of soft magnetic materials is their relative permeability, nominally defined by  $\mu_r = B/\mu_0 H$ . Actually, since the  $B(H)$  curve is neither linear nor single valued, a large number of useful permeability parameters can be defined, such as the initial permeability, the maximum permeability, and the anhysteretic permeability. For simplicity, we will consider only the initial small-signal permeability, defined as  $\mu(0) \equiv \partial B/\partial H|_{H=0}$ . Values from  $\mu(0) = 10$  (high-frequency ferrites) to  $\mu(0) = 10,000$  (low-frequency inductors) are typically encountered in applications.

**Power Devices.** Probably the main use of magnetic materials at present, on a weight basis, is in power transformers. These can range from huge transformers used in substations, to miniature transformers used to convert line voltage to a level suitable for small consumer devices, to small ferrite transformers used in switching power supplies.

Low-frequency power applications almost inevitably employ low Si percentage Fe–Si alloys known as “silicon iron” or “electrical steel.” This is because Fe is by far the cheapest magnetic material available. The coercive field of pure Fe is typically about 80 A/m. The addition of a small amount of Si to Fe lowers the anisotropy, resulting in lower losses and a coercive force of about 40 A/m. It also substantially increases the resistivity, which decreases eddy currents. Eddy currents in transformers are usually further reduced by lamination (i.e., using a stack of Fe–Si plates, each electrically insulated by a coating layer, rather than a single thick piece). The laminations are arranged so that eddy currents are interrupted by the presence of the insulator and forced to circulate only within each lamination (i.e., so that any ac magnetic flux is perpendicular to the normal). Properly designed, the thickness of each plate should be smaller than a skin depth,  $\delta$ , given by  $\delta = (2\rho/\omega\mu)^{1/2}$ , where  $\rho$  is the resistivity,  $\omega$  is the angular frequency of the ac magnetic field, and  $\mu$  is the permeability. Note that the permeability can also be frequency dependent.

Improved varieties of silicon iron are prepared using specific sequences of forging (rolling) and annealing to obtain a grain-oriented microstructure. In such materials the crystallographic orientation of individual grains is forced to be aligned over the entire piece. The advantage is that materials with lower losses are obtained, though the material is also somewhat more expensive. At present, most power transformers utilize grain-oriented silicon iron.

The highest-performance materials for low-frequency transformer applications are the metallic glasses, amorphous alloys of Fe and Co with one or more metalloid additions (usually B, C, Si, and P). These materials can have losses that are 10 times smaller than silicon iron and coercive fields below 0.5 A/m. But the saturation flux density is generally  $\sim 1.5$  T, and the materials are substantially more expensive than silicon iron. These negative aspects of metallic glasses have precluded their widespread use.

Power supplies with low weight and volume are highly desirable for some electronic applications, and this can be achieved with the design called “switching” power supplies. An additional benefit is circuit versatility and flexibility, while the main tradeoffs are increased cost and design complexity. A switching power supply uses power electronics [typically Metal-Oxide-Semiconductor Field-Effect Transistors (MOSFETs)] to chop and rectify power at high frequency, a power transformer or inductor to change the voltage, and control electronics to synchronize and control the system. The frequencies used range from about 50 kHz to a present upper limit of 1 MHz.

At these frequencies, metallic ferromagnets cannot be used as the desirable lamination thickness would be prohibitively small. Fortunately at such high frequencies, the energy that must be stored in the transformer or inductor is correspondingly small for a given power capacity ( $E \sim P/\omega$ , where  $E$  is the maximum stored energy,  $P$  is the maximum power, and  $\omega$  is the angular frequency). Therefore, the cost of the magnetic part need not dominate, especially considering the cost of the electronics involved, so the use of relatively expensive but very high resistivity ferrites is feasible. In most cases an  $(\text{Mn,Zn})\text{Fe}_2\text{O}_4$  spinel ferrite is chosen as a compromise between saturation flux density, losses, resistivity, and cost. At the highest frequencies,  $(\text{Ni,Zn})\text{Fe}_2\text{O}_4$  may prove useful, as it has a higher resistivity and therefore is less susceptible to eddy current losses.

Motors and generators inevitably employ iron or silicon iron to act as a flux concentrator. The armature and stator are commonly constructed of grain-oriented silicon iron, though small motors often employ nonoriented silicon iron or metallic-glass materials. As in transformers, the ferromagnetic parts must be laminated to reduce eddy current losses.

**Inductors.** Small-signal transformers and inductors are used in a variety of circuit applications, as in impedance-matching and isolation transformers, antennas, and chokes. Signal-level devices do not have to carry substantial power, so they can be small and the cost of materials can easily be outweighed by performance considerations. Thus a wider variety of magnetic materials is used in these devices.

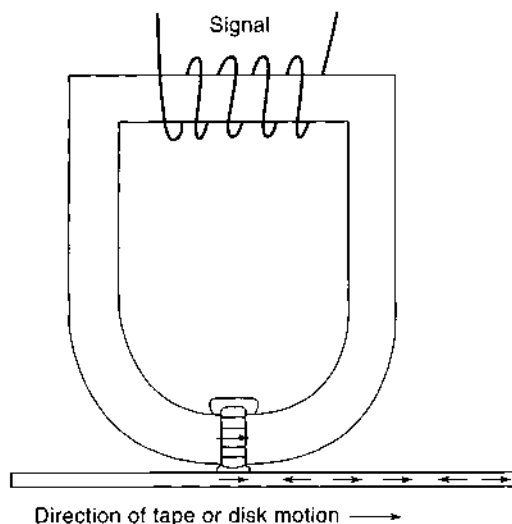
At audio frequencies and below, transformers and inductors were once commonly used for signal applications. For example, long-distance analog telephone circuits were balanced by the periodic addition of loading coils—inductors designed to match the large distributed capacitance of phone lines. A common choice for the magnetic core in those coils was Permalloy ( $\text{Ni}_{80}\text{Fe}_{20}$ ) or a related alloy, and the cores were formed by rolling a long tape into a toroidal core. However, in modern telecommunications systems the analog signals are quickly converted to digital signals at the central exchange and then transmitted by fiber-optic, satellite, or microwave relay. The need for loading coils is minimal. Similarly, the function of audiofrequency impedance-matching transformers and other inductive electronic components has largely been displaced by more elaborate

but much cheaper integrated circuit designs. Low-frequency magnetics are used in modem isolation transformers to provide dc electrical isolation with audio coupling; these are made with either laminated metallic or solid ferrite cores.

At RF frequencies (50 kHz to 50 MHz) magnetic cores are widely used; for example, in antennas, RF transformers, chokes, and resonant circuits. The usual choices for magnetic cores are the spinels  $(\text{Mn,Zn})\text{Fe}_2\text{O}_4$  (up to about 1 MHz) and  $(\text{Ni,Zn})\text{Fe}_2\text{O}_4$  (up to about 10 MHz to 50 MHz). The manganese zinc ferrites are cheaper but have a resistivity typically less than  $1000 \Omega\text{-cm}$ . Eddy current losses limit their usefulness at high frequencies. The initial permeability is roughly 1000 to 3000. Nickel zinc ferrites can have resistivities as high as  $10^5 \Omega\text{-cm}$  and initial permeabilities on the order of 1000. Their upper frequency limit is dictated by the need to avoid FMR losses. Material designed for operation above about 10 MHz typically is prepared with moderate porosity to inhibit domain-wall motion and losses associated with domain-wall resonance. The permeability mechanism then is limited to that of coherent rotation of the spins, which implies a much smaller value ( $\mu_c \sim M_s/H_k \sim 10$  to 100).

Above about 50 MHz, magnetic materials are not commonly used in transformers and inductors, as Snoek's law demands that the permeability be uselessly small in order for the FMR frequency to be sufficiently greater than the frequency of operation. Snoek's law can be circumvented by the use of materials with a large biaxial anisotropy (e.g., the hexagonal magnetoplumbite-type ferrites) or by using thin films with a high saturation magnetization. At present, however, those approaches are not commercially important.

**Write Heads.** An important application of soft magnetic materials is in the recording heads used in tape and disk systems. Write heads essentially consist of an electromagnet with a toroidal magnetic core and a very small air gap ( $\sim 100$  nm to 300 nm). The flux that extends from the air gap, called the fringing field, is used to magnetize the magnetic medium passing nearby (see Fig. 9). In tape and floppy disk systems the magnetic medium is in actual contact with the head, while in hard



**Figure 9.** Schematic of a magnetic recording head. Recording is accomplished using the magnetic fringing field (i.e., the field that leaks from the gap).

disk systems the head flies aerodynamically over the spinning medium at a height of 25 nm to 75 nm. The maximum magnetic field available for magnetizing the medium is proportional to the saturation magnetization of the head material.

Originally the magnetic recording heads were made from laminated metal alloys, such as Permalloy or Sendust (an alloy of Fe, Si, and Al notable for being magnetically soft and physically very hard), but as recording densities and frequencies increased, an inevitable move was made to ferrite materials. Ferrite heads are made from cast pieces that are carefully polished to form a precise air gap and then are assembled with the driving coil. As recording densities have increased, media with higher coercivities are necessary, so head materials with high-saturation magnetization are needed. Unfortunately this is where the ferrites are most deficient, with maximum  $B_s \sim 0.6$  T. One approach to obtain improved performance is to add a thin film of relatively high  $B_s$  material (such as Permalloy,  $B_s = 0.9$  T to 1.1 T) on the inside edge of the gap of each piece. The thin film acts as a flux concentrator and increases the fringing field significantly. This approach is called the metal-in-gap (MIG) design.

The highest-performance recording heads are constructed using thin films and photolithography. They have the advantage of great precision in layout, leading to extremely narrow pole tips, precise gap widths, and precise placement of the driving coils. Permalloy and related alloys are most commonly used at present, but soft amorphous alloys such as  $\text{Co}_{0.85}\text{Nb}_{0.08}\text{Zr}_{0.07}$  offer significant improvements with  $B_s \sim 50\%$  greater than Permalloy. The resistivity of the amorphous alloys is in the range  $\rho \sim 100 \mu\Omega\text{-cm}$ , about 5-fold higher than Permalloy, which reduces eddy current losses, allowing recording at higher frequencies. Further advantage can be obtained by using the recently developed nanocrystalline (grain size  $\sim 1$  nm to 5 nm) alloys such as  $\text{Fe}_{0.92}\text{Ta}_{0.05}\text{Nb}_{0.03}$ , which have extremely high  $B_s \sim 2.0$  T and also have resistivities in the range of  $100 \mu\Omega\text{-cm}$  to  $150 \mu\Omega\text{-cm}$ . These materials are being vigorously developed for future generations of high-performance recording, especially hard disk drives.

**Read Heads.** The same magnetic structure that is used for magnetic recording can also be used for reading the recorded signal. The passage of recorded domains across the gap of the recording head will induce a small voltage on the driving coil, which is amplified and processed. Indeed, this inductive read head is the simplest and cheapest approach to magnetic recording. However, the signal level is very low and limits the performance of hard disk drives. The present generation of hard disk drives uses a separate magnetic structure (incorporated in a single read/write head) for reading the recorded magnetic signal. The sensor in this device is a magnetoresistive thin film made of Permalloy and biased with a dc current. External flux from the recording medium couples to the Permalloy film, rotating the magnetization with respect to the fixed direction of the dc current. This changes the resistance of the Permalloy (the so-called conventional or anisotropic magnetoresistance effect, AMR) and therefore the voltage developed across the device. The AMR effect is relatively small, about 5%, but is sufficient to deliver signal superior to that of an inductive head.

Superior performance can be obtained by incorporating materials that exhibit giant magnetoresistance. These can have responses that are about 10-fold higher than AMR materials. So-called spin valves employ the same physical principle as GMR

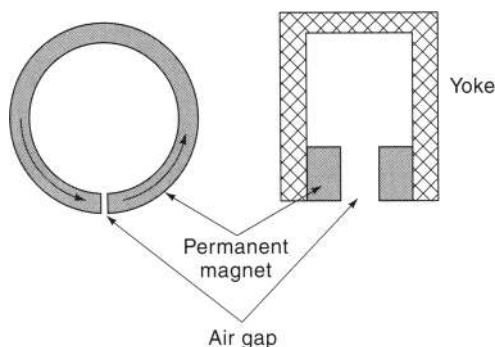
(namely, spin-dependent scattering) in a more highly controlled and responsive magnetic structure. The higher signal output of spin-valve read heads, in conjunction with careful redesign of the entire magnetic recording system (including media, head structure, electronics, signal processing, etc.), can lead to substantial increases in recording density.

**Shields.** Soft magnetic materials are also used to make magnetic shields, typically used to protect electronic components from magnetic interference or to contain the external field around a component that generates magnetic flux. An example of the former are shields for cathode-ray tubes, such as computer monitors, while an example of the latter are shields for speakers, such as “multimedia” speakers intended for placement close to a (unshielded) monitor. These shields are usually formed from Permalloy and related alloys. Often “Mumetal” is specified for these applications—this originally designated a particular Fe–Ni–Cu alloy but it is now used generically to refer to many high-permeability alloys. Note that RF shields (unlike dc magnetic shields) rely on eddy current screening and use high-conductivity paramagnetic metals such as Cu.

A related application is the use of materials that absorb RF energy. These are typically lossy ferrites and are used in the form of beads threaded on wires, where it is desired to suppress high-frequency signals. Manganese-zinc ferrites are generally used for this purpose, although nickel-zinc is used to obtain the highest cutoff frequencies. Lossy ferrites can also be used as an antiradar coating on military aircraft.

### Hard (Permanent) Magnets

Permanent magnets are used in a wide variety of applications where a static magnetic field is desired. The dominant uses are in speakers and dc motors and as holding magnets. The field available from a given permanent magnet depends on the physical configuration but is limited to  $B_s$ , the saturation flux density in the optimum case. Specifically, for a toroidal part with a small air gap (see Fig. 10), the flux density in the gap will be  $B_s$ . For Fe this is  $B_s = 2.15$  T, while for  $\text{Fe}_{0.6}\text{Co}_{0.4}$  it is  $B_s = 2.43$  T (11, p. 190), the highest value for any known bulk material. In less optimum geometries, the available flux density can be greatly reduced, as determined by magnetostatics. For arbitrary-shaped parts, the usual approach is to employ finite-element numerical calculations to infer the flux distribution.



**Figure 10.** Air gap in magnetic circuits. The gray regions represent permanent magnet material, with the direction of magnetization shown. The hatched region represents soft magnetic material, which is used to complete the magnetic circuit.

In general, where the design figure of merit is the flux available per unit volume of magnetic material, the corresponding figure of merit for the magnetic material is the energy product  $(BH)_{\max}$  (i.e., the largest value measured at any point in the second quadrant [positive  $B$ , negative  $H$ ] of the  $B$ – $H$  curve). This is only a crude way to evaluate the usefulness of a material; nevertheless it is indicative and is commonly quoted by permanent magnet manufacturers. Four main types of permanent magnet materials are commonly used:

- Alnico, a class of Fe–Al–Ni–Co–Cu alloys. The properties of this material are entirely dependent on complex metallurgical processing and microstructural control. Commonly used Alnico materials have  $(BH)_{\max} = 50 \times 10^3 \text{ J/m}^3$ .
- Barium ferrite, typically  $\text{BaO} \cdot (\text{Fe}_2\text{O}_3)_6$ , is the standard “ceramic” magnetic material. The typical value  $(BH)_{\max} = 25 \times 10^3 \text{ J/m}^3$  is smaller than that of Alnico, but the material is both lower density and cheaper to produce than Alnico and has almost entirely replaced Alnico in highly cost-sensitive applications. The high value of  $(BH)_{\max}$  in this and the following materials is due to the very high intrinsic anisotropy.
- Sm–Co is the standard high-performance “rare-earth” permanent magnet, with  $(BH)_{\max} = 160 \times 10^3 \text{ J/m}^3$ . The major disadvantage of this material is its cost.
- Nd–Fe–B, a more recent material, has an even higher value for  $(BH)_{\max} = 320 \times 10^3 \text{ J/m}^3$  than Sm–Co, and it is less expensive. The major disadvantage of this material in some applications is that the  $T_c$  is somewhat low,  $T_c \sim 150^\circ\text{C}$ .

Loudspeakers have long been a dominant application for permanent magnetic materials. The permanent magnet is used to establish a magnetic field in an annular region in which the voice coil is mounted. When a current is driven through the voice coil, an axial force is produced, in accordance with the Lorentz relation,  $\mathbf{F} = e\mathbf{v} \times \mathbf{B}$ , where  $\mathbf{F}$  is the force on an electron,  $e$  is the charge on an electron, and  $\mathbf{v}$  is the velocity of the electron. Motion of the voice coil is coupled to a speaker cone to move the air efficiently and thereby produce sound waves.

The energy product of a magnetic material is a good figure of merit for speaker applications, since for a given design, a higher-energy product will result in a higher flux density in the annular gap. Most loudspeakers are low-priced components, so the cost of the magnetic material is the other key factor. For this reason, barium-ferrite is the dominant material used. In some applications, such as high-performance earphones, the amount of magnetic material is small so materials cost is less critical. In the past, Sm–Co magnets have been used for these applications, though Nd–Fe–B magnets are now clearly the best overall choice.

The materials used for permanent magnets inevitably cost more than silicon iron, so large motors and generators always use soft magnetic materials wound with coils to create the required magnetic field. In small motors, the economics are dominated by the cost of fabrication: The small coils and fine tolerances needed for electromagnet motors outweigh the added cost of permanent magnetic materials. A vast number of small dc motors are produced for a very wide range of applications, from clock motors to the dozens of motors in modern automobiles

used to drive windows, locks, seats, windshield wipers, and so on.

An economically important use of permanent magnets is in the low-tech application of holding magnets. These range from decorative magnets for holding notes to a refrigerator door to functional magnets for holding and sealing the refrigerator door shut to strong magnetic chucks for holding ferrous materials for machine-forming operations. In almost all cases cost is paramount, and barium ferrite is used. For some applications the ferrite powder is mixed with a polymer precursor, formed into a tape, and polymerized to form a flexible magnet, albeit with reduced net flux density and therefore reduced holding power.

### Magnetic Recording Media—Intermediate Between Soft and Hard

Magnetic recording is a huge business, dominated by the hard disks and floppy disks pervasive in personal computers and by tape recording—audio, video, instrumentation, digital data storage, and so on. The media used in recording are magnetic materials that must have a relatively high coercive force so that they do not spontaneously demagnetize and lose information. But the coercive force cannot be much greater than about 100 kA/m because the leakage flux from the recording head is limited and must nevertheless be sufficient to saturate the medium.

The standard material used in tape and floppy-disk media is gamma iron oxide,  $\gamma\text{-Fe}_2\text{O}_3$ . It is moderately expensive to prepare (compared to conventional ferrites) because extremely reproducible and controllable properties are required. Acicular (needle-shaped) particles  $\sim 50$  nm in diameter are prepared in order to obtain good recording characteristics. The coercivity of most tapes is about 20 kA/m to 30 kA/m. Recent high-performance formulations use cobalt-modified  $\gamma\text{-Fe}_2\text{O}_3$ , which has a thin cobalt-rich region on the surface of the particles. This material has an increased coercivity of about 50 kA/m and is routinely used for video tape. In the past,  $\text{CrO}_2$  was used as a high-performance medium because it has a higher  $M_s$  than  $\gamma\text{-Fe}_2\text{O}_3$  and can be prepared with a coercivity as high as 80 kA/m. However, it has a low  $T_c$  (only 130°C) and is relatively expensive, so it has been displaced by cobalt-modified  $\gamma\text{-Fe}_2\text{O}_3$ .

Hard disks used in digital recording are aluminum platters coated with paramagnetic Ni-P or Cr and then a recording medium such as Co-Cr. An extremely thin layer of C is usually then deposited along with an even smaller amount of lubricating fluid, in order to avoid catastrophic head contact with the medium ("head crashes"). The coercivity of hard disk media is in the range 60 kA/m to 100 kA/m, and the saturation magnetization is about 1000 kA/m.

For many years it was expected that magnetic bubble memory might find a role for data storage in computers. However, bubble memory is slow and expensive compared to hard disks and semiconductor random access memory (RAM). Its only advantages are that it is mechanically more robust than a hard-disk system, and it is more radiation resistant than semiconductor RAM. Therefore, the only present use for bubble memory is in certain military applications.

Bubble memories store data in cylindrical domains, perpendicular to the surface, that are generated in liquid-phase epitaxy (LPE) grown garnet films. The films are designed to have a large uniaxial anisotropy perpendicular to the film plane and

are chosen to meet a number of other criteria. The bubbles are moved around using a rotating external magnetic field created by a miniature electromagnet, and they are detected using a magnetoresistance bridge.

Finally, there is magneto-optic recording. In this scheme data are stored on a plastic disk coated with an amorphous metallic thin film such as  $\text{Tb}_{0.2}\text{Fe}_{0.8}$  or similar compositions doped with Dy and Co. They have a large uniaxial anisotropy, with an easy axis perpendicular to the plane of the film, so as with the garnet films designed for magnetic bubble applications, cylindrical domains are stable. Data are written by focusing a laser on the desired spot, which heats the films above its  $T_c$ ,  $\sim 100^\circ\text{C}$  to  $150^\circ\text{C}$ . If the film is exposed to a moderate magnetic field while it cools, the heated region will magnetize in the direction of the applied field. Thus alternating regions of, say, north-up and south-up can be written. The data are read with the same laser at lower power to avoid heating, using a polarizer to detect Faraday rotation (i.e., rotation of the polarization of light when it interacts with a magnetic material). This effect is known as Kerr rotation when it occurs on reflection from a metallic magnetic surface. The Kerr rotation in a magnetic material is in the opposite sense for regions magnetized north-up versus south-up. The maximum Kerr rotation in Tb-Fe films is rather small,  $\sim 0.2^\circ$ , which gives a low signal/noise ratio and correspondingly low data rate.

The data density in magneto-optic recording is comparable to that of a conventional CD-ROM but has the great advantage of being endlessly rewriteable. Magneto-optic drives are slower for writing data and more expensive than conventional hard drives, but they confer the advantage of cheap removable media and good archiveability and have found a small but significant market niche.

### Miscellaneous

Along with the three conventional classes of applications for magnetic materials, there are a wide variety of specialized applications, too numerous to list exhaustively.

**Nonreciprocal Materials.** Some of the most interesting magnetic devices are based on the nonreciprocal propagation of UHF signals in insulating magnetic materials. The behavior is formally identical to the small optical Faraday rotation observed in some nonferromagnetic materials, but the effect in ferrites can be very large and is commonly used in microwave applications (for example in isolators and circulators).

The nonreciprocal phenomena are due to interaction of the incident radiation with the precessing electron spin (see MAGNETIC RESONANCE). A particularly straightforward case arises when the incident microwaves are circularly polarized with a propagation vector parallel to the easy axis of the magnetic material and at a frequency equal to the natural precession frequency of the electron spins. Then, if the sense of the circular polarization is the same as that of the electron spin precession, energy is readily transferred to the spins and dissipated as loss. If the sense is opposite (corresponding to propagation in the opposite direction), then there is little interaction and the loss is minimized. It is easy to imagine an isolator based on the directionality of this phenomenon, though the practical design of this and other nonreciprocal microwave devices can be very complex. In general, a moderately large external biasing magnet is required to set the FMR frequency equal to



the operating frequency (a larger field is required for a higher operating frequency).

Three classes of materials dominate the magnetic materials used for microwave applications. The figure of merit for microwave devices is usually proportional to  $1/\Delta H$ , where  $\Delta H$  is the FMR linewidth.

- First, the garnet-structure ferrites exhibit the highest performance available because they have the lowest FMR linewidths: A typical value for polycrystalline ceramic yttrium iron garnet, YIG, is  $\Delta H = 4000$  A/m. Single crystals with linewidths as low as 40 A/m have been reported. Single crystals are relatively expensive but are usually the best choice at low frequencies.
- Second, the spinel ferrites are useful in a number of cases. Nickel zinc ferrite is particularly useful at high powers because it has a higher  $T_c$  than garnets, and it is often used in the range above 10 GHz. Manganese-magnesium ferrites are used in the range of 5 GHz to 10 GHz because their lower saturation magnetization allows biasing at lower field.
- Third, the Z-type hexagonal ferrites have a hard axis normal to the basal plane, so the need for external biasing is reduced. They require expensive processing to produce oriented, high-quality ceramics. They are most useful for mm-wave applications.

In general, the key issues in developing materials for microwave applications are related to processing and microstructure control rather than exploring new compositions. Small grain size is important to minimize losses from domain-wall resonance, to maximize the resistivity, and to make stronger materials that resist the thermal stresses caused by high-power operation. Obtaining a high density is always important, but densification usually is accompanied by grain growth. The art of designing or choosing a microwave magnetic material is in balancing these conflicting requirements.

**Ferrofluids.** Ferrofluids are liquid suspensions of magnetic particles coated to avoid agglomeration. Typically the liquid is a hydrocarbon or silicone, and the magnetic particles are ferrites, but a wide range of combinations is possible. The behavior of ferrofluids is unique and scientifically interesting, and ferrofluids are commercially used for such diverse applications as gas-tight seals for rotary shafts, heat transfer agents for loud-speaker voice coils, and damping systems.

## BIBLIOGRAPHY

1. Plato, in *The Dialogs of Plato*, Trans. B. Jowett, *Great Books of the Western World*, Vol. 7, Chicago: Benton, 1952, p. 144.
2. W. Gilbert, *De Magnete*, P. F. Mottelay (trans.), Mineola, NY: Dover, 1958.
3. J. D. Jackson, *Classical Electrodynamics*, New York: Wiley, 1962.
4. R. L. Carlin, *Magnetochemistry*, Berlin: Springer-Verlag, 1986, p. 3.
5. E. H. Brandt, *Science*, **243**: 349–355, 1989; and *Physics World*, **10** (4): 23–24, 1997.
6. K. Binder and A. P. Young, Spin glasses: Experimental facts, theoretical concepts, and open questions, *Rev. Mod. Physics*, **58**: 801–976, 1986.
7. A. P. Ramirez, *Ann. Rev. Mater. Sci.*, **24**: 453–480, 1994.

8. R. B. van Dover et al., Intrinsic anisotropy of Tb—Fe films prepared by magnetron co-sputtering, *J. Appl. Phys.*, **57**: 3897–3899, 1985.
9. M. N. Baibich et al., *Phys. Rev. Lett.*, **61**: 2472–2475, 1988.
10. S. S. P. Parkin, Giant magnetoresistance in magnetic nanostructures, *Ann. Rev. Mater. Sci.*, **25**: 357–388, 1995.
11. R. M. Bozorth, *Ferromagnetism*, New York: Van Nostrand, 1951.

## Reading List

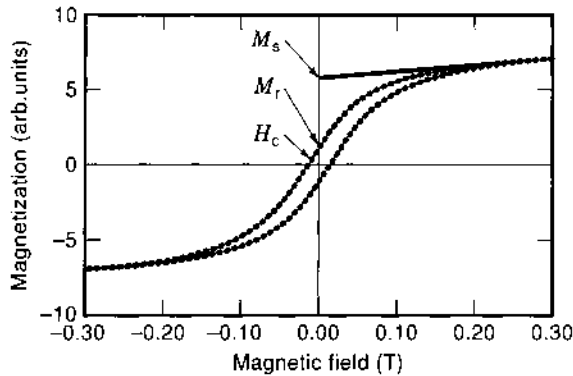
- R. M. Bozorth, *Ferromagnetism*, New York: Van Nostrand, 1951.
- F. N. Bradley, *Materials for Magnetic Functions*, New York: Hayden, 1971.
- F. Brailsford, *Physical Principles of Magnetism*, New York: Van Nostrand, 1966.
- P. Campbell, *Permanent Magnetic Materials and Their Applications*, Cambridge: Cambridge Univ. Press, 1994.
- C.-W. Chen, *Soft Magnetic Materials*, Amsterdam: North Holland, 1977, reprinted by Dover, New York, 1986.
- S. Chikazumi and S. H. Charap, *Physics of Magnetism*, Malabar, FL: Krieger, 1964.
- B. D. Cullity, *Introduction to Magnetic Materials*, Reading, MA: Addison-Wesley, 1972.
- D. Jiles, *Introduction to Magnetism and Magnetic Materials*, London: Chapman and Hall, 1991.
- M. M. Schieber, *Experimental Magnetochemistry*, New York: Wiley, 1967.
- J. Smit and H. P. J. Wijn, *Ferrites*, New York: Wiley, 1959.

ROBERT B. VAN DOVER  
Bell Labs, Lucent Technologies

## MAGNETIC NOISE, BARKHAUSEN EFFECT

The change of the magnetic state of a ferro- or ferrimagnetic body under the effect of a slowly changing externally applied magnetic field is a complex process, involving reversible and irreversible changes of the magnetization. Due to irreversible magnetization processes, any change of the magnetic state is accompanied by losses, manifested in the presence of the magnetic hysteresis. The reason for the losses is that any real material has a defect structure, responsible for the details of the hysteresis loop. The magnetization changes discontinuously, by jumps from defect to defect, giving rise to magnetic noise, which can be made audible through a headphone. This process and the corresponding noise is named after its discoverer, H. Barkhausen (1919) as *Barkhausen jumps* and *Barkhausen noise*.

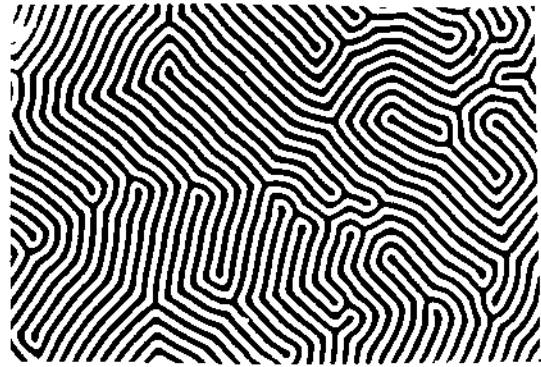
The magnetization process ( $M$ ) of a ferromagnetic (or ferrimagnetic) body is represented by a nonlinear and multivalued function of the applied magnetic field ( $H$ ), the magnetic hysteresis loop. Figure 1 shows the magnetic hysteresis loop of a magnetic composite of nanometer size iron particles embedded in a nonmagnetic ZnO matrix. At any given temperature the maximum attainable magnetization is called the saturation, or spontaneous magnetization,  $M_s$ . Reducing the magnetic field from saturation to  $H = 0$  brings the magnetic body to the remanent state with the remanent magnetization (or remanence) of  $M_r$ . Further reducing the field in the negative direction down to the coercive field (coercivity), at  $-H = -H_c$ , the magnetization will be reduced to  $M = 0$ .



**Figure 1.** Magnetic hysteresis loop of a nanocomposite of iron in ZnO.  $M_s$  saturation magnetization,  $M_r$  remanent magnetization,  $H_c$  coercivity (M. Pardavi-Horvath, unpublished data).

The equilibrium state of the magnetization corresponds to the minimum of the free energy of the magnetic material. The most important contributions to the free energy are due to (1) the quantum mechanical exchange energy, responsible for the collective ordering of the individual spin magnetic moments of the electrons; (2) the magnetocrystalline anisotropy energy, favoring some crystalline directions for the magnetization direction with respect to others, giving rise to the easy and hard directions of magnetization; (3) magnetostrictive stresses, that is, the length change of a magnetic material in the presence of mechanical stress; (4) the Zeeman energy of the interaction with the externally applied magnetic field(s); (5) the magnetostatic energy due to the creation of internal and external demagnetizing fields, depending on the shape and size of the magnetic body; and (6) the energy of interaction of the magnetization with defects in the magnetic body. Due to (6) the details of the magnetization curve of two apparently similar pieces of the same magnetic material might differ substantially. In the absence of any applied magnetic field, the equilibrium state of a magnetic body would be that of zero magnetization. The existence of the remanent magnetization, the basis of the operation of any permanent magnet or magnetic recording device, is the result of the defect structure in the material, and/or shape and size distribution of the particles or grains. These features are controlled by the technology of manufacturing the material. However, any magnetic material can be brought to the  $M = 0$  state in  $H = 0$ , by demagnetizing it. The most frequent process of demagnetization is to apply an ac magnetic field of an amplitude large enough to saturate the material, and slowly and gradually reduce the amplitude to zero.

In the demagnetized state the distribution of the magnetization inside of the magnetic body is nonuniform. The magnetic structure consists of domains, small regions of different magnetization direction (1). The vectorial sum of all of the domain magnetizations is zero in the absence of any applied magnetic field. The geometry and the size of the domain structure is governed by the interactions, delineated previously. The domain structure can be observed by a number of techniques, including magnetic force microscopy, Lorentz electron microscopy, or in polarized light, utilizing the magneto-optic Faraday or Kerr effects. In some cases the domain structure can be very simple, consisting of only two sets of domains, magnetized “up” and “down,” as in the case of a  $5 \mu\text{m}$  thin, single crystalline ferri-

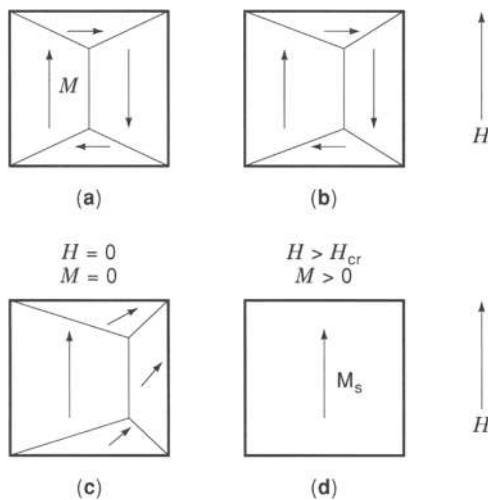


**Figure 2.** Magnetic domain structure in  $H = 0$  applied magnetic field of a defect-free  $5 \mu\text{m}$  thin single crystalline yttrium iron garnet film. The magneto-optic contrast is due to the opposite direction of the magnetization in black-and-white domains with respect to the surface normal. The periodicity of the domain structure in  $10 \mu\text{m}$  (Microphotograph in polarized light).

magnetic garnet sample, a microphotograph of which, taken in polarized light via the magneto-optic Faraday effect, is shown in Fig. 2. The areas of dark and light contrast correspond to the directions of the magnetization perpendicular to the sample surface, directed “downward” and “upward,” respectively. In  $H = 0$  the two regions have equal volumes, such that the total magnetization  $M = 0$ . The boundary line between the domains is the domain wall (DW), where the direction of the magnetization is changing gradually between the two domains whose magnetization directions differ by  $180^\circ$ . The size and shape of the domains and the position of the DWs correspond to the minimum free energy configuration. In a flawless material the domain structure is very regular by crystal symmetry, as seen in Fig. 2 for a single crystal grown along the  $[111]$  crystalline direction. The threefold symmetry of that axis is evident.

Upon applying an external magnetic field to a previously demagnetized magnetic body, its magnetization follows the initial (virgin) magnetization curve. Assuming a perfectly homogeneous, flawless material without any preferred axis, the process of magnetization change might proceed reversibly, with no losses involved (1,2). In such a case, the field will exert a force on the domain magnetizations and increase the magnetizations of the domains lying near to the field direction by increasing their volume. This magnetization change happens via the motion of the DWs to new positions. This wall motion is reversible up to a certain magnetic field value and if the field is reduced to zero, the walls move back to their original position and  $M$  returns to zero ( $M_r = 0$ ) in  $H = 0$ . In an ideally perfect material this reversible wall motion would dominate the magnetization process up to magnetic field values near to saturation. Then the magnetization vectors of the remaining domains would start turning into the direction of the field, completing the magnetization process by rotation. This magnetization process is illustrated in Fig. 3. For a thin slab of a magnetic material, having an easy axis of the magnetization along the side of slab, the magnetic field is applied along this easy axis.

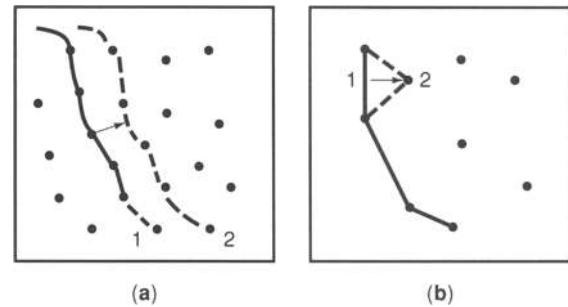
In any real material there are always microstructural defects: inclusions of foreign phases, voids, dislocations, inhomogeneous stresses, irregular surfaces and interfaces, grain boundaries, cracks, and so on, such that the free energy of the material is not uniform over the body, and the conditions for



**Figure 3.** Magnetization processes by wall motion and rotation. (a) Demagnetized state:  $H = 0$ ,  $M = 0$ ; (b) Irreversible wall motion:  $H > 0$ , domains having favorable magnetization direction with respect to the applied magnetic field are growing; (c) The magnetization in domains with unfavorable direction is rotating toward the magnetic field direction; (d) Saturation: domains disappeared, magnetization aligned with the applied field.

magnetization change will vary locally. The free energy might have local minima (or maxima) at some defects, where the DWs will be pinned. They can only move and increase (decrease) the magnetization when the magnetic field changes enough to provide a force sufficient to overcome the energy barriers between the defects. The domain wall will then discontinuously jump from one pinned position to another. The irreversible processes of discontinuous DW displacement are called Barkhausen jumps. In a macroscopic body, the defects are distributed more or less randomly in space and energy. When the magnetic field reaches the critical value for the irreversible DW motion of the weakest defect,  $H = H_{cr}$ , the wall will move until it reaches the next defect, where it will be pinned again. Upon increasing the field there might be another DW pinned at another defect, having a critical field equal to the applied field, and now that wall will jump to the next available energy minimum, and it will be pinned at the next defect. Depending on the density of the defects and the strength of the interaction between defects and DWs, it might happen that a wall will be pinned by many similar defects and it will move, breaking away from many defects simultaneously, giving rise to a large magnetization jump. Figure 4 shows these two different scenarios for Barkhausen jumps. Figure 5 shows strong DW pinning at inclusions in a garnet crystal, near magnetic saturation. The applied field is high enough to reduce the black domains to a very small area, but it is not enough to overcome  $H_{cr}$  of the inclusions.

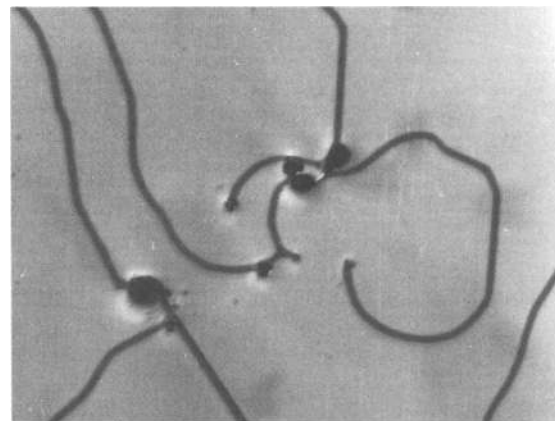
A finite size ferromagnetic material is usually broken into many magnetic domains. Any material, including the demagnetized state, is not, in general, a unique state, as there is a very large number of possible domain arrangements summing to a particular  $M$  value. Consequently, the sequence of the Barkhausen jumps is probabilistic. Moreover, the critical field for a jump depends not only on the strength of the actual defect-DW interaction, but on the magnetic state of the



**Figure 4.** Magnetization change by irreversible wall motion: Barkhausen jumps of domain walls (lines) pinned at defects (dots). (a) Wall pinned at many defects, simultaneous breakaway with a large magnetization jump. (b) Domain wall motion by individual jumps from defect to defect. 1 is the original DW position in  $H = 0$ ; 2 is the DW position in  $H > H_{cr}$ .

neighboring domains too. These domains, like small magnets, create their own magnetic fields at the site of their neighbors (magnetostatic, dipole fields). Depending on the direction of the magnetization of a given domain, it can help or prevent the externally applied magnetic field in supplying the energy for the jump. This energy, needed to overcome the barriers caused by the defects, results in hysteresis losses.

The discrete and irregular nature of the magnetization process causes hysteresis loss and noise in magnetic devices. To reduce these losses, the material should be uniform. The losses can be reduced, even in the presence of a large defect density, if the defects are uniformly distributed and/or they interact very weakly with the moving DWs. In this case the DWs are moving smoothly across the material, with no sudden jumps in the magnetization. The other limiting case is when the material has a very narrow distribution of defects yielding a rectangular hysteresis loop. In contrast, for a good permanent magnet very active and strong pinning centers have to be created to keep the magnetization from changing, that is, avoid spontaneous demagnetization.



**Figure 5.** Strong domain wall pinning at inclusions in an epitaxial garnet crystal in applied magnetic field near to saturation, where the magnetization  $M = M_{up} - M_{down} \leq M_s$ . The black stripe domains correspond to  $M_{down}$ . Upon applying a magnetic field  $H \gg H_c$ , most of the walls disappear, only those pinned by the strongest defects (black dots) are left.

## EQUILIBRIUM MAGNETIZATION DISTRIBUTION

### Domains and Domain Walls

The equilibrium state of the macroscopic magnetization of any magnetic body is the state with minimum total free energy (1–5). The term in the free energy responsible for the existence of ordered magnetic states is the exchange interaction between the electrons' magnetic moments. The magnitude of the atomic magnetic moments is constant at a given temperature. The atomic magnetic moments are rendered parallel (antiparallel) to each other due to the exchange interaction, giving rise to ferromagnetic (antiferromagnetic, ferrimagnetic) order. The thermal energy tends to randomize the direction of the atomic magnetic moments, and at the Curie temperature,  $T_c$ , it overcomes the exchange energy, and there is no magnetic order above  $T_c$ . The magnitude of  $T_c$  is proportional to  $A$ , the exchange constant, characterizing the strength of the exchange interaction. For iron,  $A = 1.49 \times 10^{-11}$  J/m.

The origin of magnetocrystalline anisotropy energy is the interaction between crystalline electric fields and atomic magnetic moments (spin-orbit coupling), rendering these moments parallel to certain, easy crystallographic directions. The simplest form of the magnetocrystalline anisotropy is the uniaxial anisotropy, with one easy axis. The anisotropy energy density can be characterized by a single anisotropy constant  $K_u$ . For the hexagonal, uniaxial cobalt at room temperature  $K_u = 4.1 \times 10^5$  J/m<sup>3</sup>. For cubic crystals the anisotropy is weaker and characterized by two constants,  $K_1$  and  $K_2$ , where usually  $K_2 \ll K_1$ . For iron  $K_1 = 4.8 \times 10^4$  J/m<sup>3</sup>.

In the absence of an applied magnetic field, the equilibrium state of the distribution of the atomic moments can be found from the minimalization of the total energy, that is, it is determined by the competition of the anisotropy and exchange energies, giving rise to the magnetic domain structure. The change of the magnetization direction between adjacent domains is not abrupt, because it would require a large investment of energy against the exchange energy. However, if the transition is very broad then many magnetic moments would be along unfavorable directions from the point of view of anisotropy energy. The result is the minimum energy situation with a formation of a DW of width:

$$\delta_w = \pi \sqrt{A/K}$$

To create a DW, a new, magnetic surface has to be formed. This surface energy of the DW energy is given by:

$$\gamma_w = 4\sqrt{AK}$$

There are other contributions to the free energy of a magnetized body, which have to be taken into account when calculating the distribution of the magnetization. These are the magnetostrictive stresses, the Zeeman energy of the interaction with applied magnetic field, and the magnetostatic energy due to production of demagnetizing fields, depending on the shape and size of the magnetic body.

The total magnetization of the body is the vector sum of the domain magnetizations,  $\mathbf{M} = \sum \mathbf{m}_i v_i$ , where  $\mathbf{m}_i$  is the magnetization vector in each domain directed along  $\phi_i$  with respect to field direction, and  $v_i$  is the volume of the  $i$ th domain. A magnetic material is saturated when all the atomic magnetization

vectors are parallel to the applied magnetic field ( $e_i = \cos\phi_i = 0$ ), that is, all the domain magnetizations are turned into the field direction  $\mathbf{M} = \mathbf{M}_s$ . (See Fig. 3.) Before saturation any change of the total magnetization can be described by:

$$\delta\mathbf{M} = \delta\sum \mathbf{m}_i v_i = \sum \mathbf{e}_i \delta|m_i| + \sum \mathbf{e}_i \delta v_i + \sum v_i |m_i| \delta \mathbf{e}$$

The first term is zero because the magnitude of the magnetic moments is constant; the second term describes the change of the volume of the domains via DW motion, and the third term is the contribution of the rotation of the magnetization angle  $\phi_i$ , inside the domains to the change the magnetization.

## DOMAIN WALL—DEFECT INTERACTION

### Domain Wall Pinning Coercivity

The equilibrium magnetization distribution in any  $\mathbf{H}$  field is expected to be determined by the minimum energy configuration, when all the contributing energies are taken into account. In a perfectly uniform material the DW energy is the same everywhere and the position and the distance between the DWs is uniform. However, real materials are not uniform. Material properties differ from point to point, and as a consequence, the material constants  $A$  and  $K$  are not constants, but they are functions of the position  $r$ :  $A(r)$  and  $K(r)$ . The result is that at some sites the DW energy is locally lowered (or increased) and extra energy is needed to move the wall from that position, that is, to change the magnetization. In a one-dimensional case, the energy needed to move a DW of volume  $v$  ( $v = hL\delta_w$ , height  $h$ , length  $L$ , width  $\delta_w$ , specific DW energy  $\gamma_w$ ) by  $dx$  has to be compensated by the change of the total DW energy (5,6):

$$\begin{aligned} 2HMdx &= d(v\gamma_w) \\ &= \gamma_w dv + vd(\gamma_w) \simeq \gamma_w d(hL\delta_w) + vd\sqrt{A(x)K(x)} \end{aligned}$$

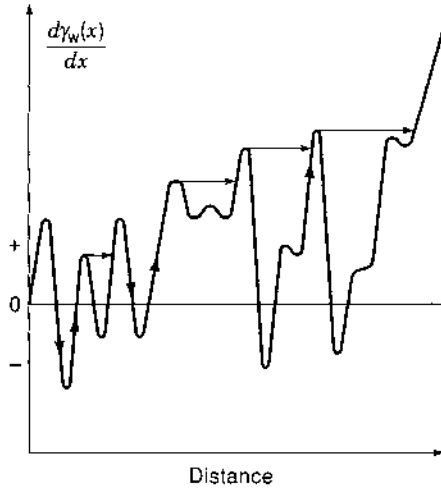
The critical field for the start of the motion of a given DW, that is, the first Barkhausen jump, is:

$$H_{cr} = \frac{1}{2M_s} \frac{d(v\gamma_w)}{dx}$$

For the whole material, the average critical field, that is, the coercive force of the start of the DW motion, for the case when the DW volume remains constant, is given by the rms value of the critical fields of all defects:

$$H_c = \overline{H_{cr}^2}^{1/2} = \frac{1}{2hLM_s} \overline{(d\gamma_w/dx)^2}^{1/2}$$

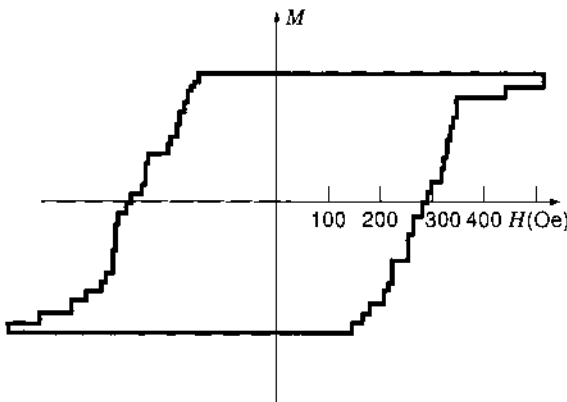
Thus the necessary condition for the occurrence of Barkhausen jumps is the presence of a gradient of the DW energy. Figure 6 shows a one-dimensional sketch of the energy landscape of a magnetic body, due to the gradient of the DW energy. The DW is located in an energy minimum, pinned by the defect. Upon reaching the critical field of interaction with a defect, the DW breaks away, jumping to a new metastable equilibrium position. The DW can freely move and the magnetization changes continuously if the defects along its path are weaker than the previous maximum in the gradient. The variation of the DW volume across the surface might have a significant contribution



**Figure 6.** One-dimensional variations of the DW energy landscape of a magnetic body due to the position dependence of anisotropy and exchange energies, caused by localized crystalline defects. Arrows denote the jump of the domain walls to the next position of equal energy.

to hysteresis losses due to surface roughness effects, as in the case of soft magnetic thin films in recording heads (6).

The change of the DW energy due to local variations of anisotropy and exchange energy causes DW pinning at localized defects. These defects are responsible for Barkhausen jumps and hysteresis losses. During the magnetization process the energy from the applied magnetic field has to overcome the energy of interaction of the moving domain wall with defects in the magnetic body. Due to the statistical nature of the defect distribution, the magnetization curve of two apparently similar pieces of the same magnetic material might differ substantially, depending on the technology of production. Figure 7 is a measured hysteresis loop of 25 quasi-identical magnetic garnet particles. Each particle is seen to be magnetized at different fields, illustrating the distribution of critical fields (7). The mean value of the critical fields for DW motion is scaled with the coercivity of the material. However, for the macroscopic critical field, the coercivity is a statistical parameter, and it has a certain distribution and standard deviation. The standard deviation of the critical field distribution characterizes the spectrum of the



**Figure 7.** Measured hysteresis loop of 25 quasi-identical magnetic garnet particles, illustrating the sequence of Barkhausen jumps corresponding to the distribution of critical fields of individual particles. (From Ref. 8.)

defects based on the strength of DW-defect interaction. More accurate, statistical models, given in Ref. 5, take into account that the process of pinning/depinning takes place only over a distance of one domain width  $w_d$ , so there will be a  $\ln(w_d/\delta_w)$  factor in all formulas for DW pinning. If the defect size is very small ( $10^{-7}$  m) compared to the DW width, then the moving DW averages out the defect potential, and there is no irreversible magnetization and hysteresis loss associated with a very fine microstructure. On the other hand, when the defects are very large compared to  $\delta_w$ , then the energy associated with the inclusions will be reduced due to the presence of secondary domain structure and the associated discontinuous DW motion will be negligibly small and the irreversible magnetization losses will be reduced.

### Inhomogeneous Stress Fields

One of the main contributors to Barkhausen noise is related to the stress sensitivity of magnetic materials. The magnetoelastic behavior is characterized by the magnetostrictive strain  $\lambda = dl/l$ , the relative shape and size change of a magnetized body upon change of the magnetization, or vice versa the change of magnetization upon deformation (1, 5). For single crystals the strain might be different in different crystallographic  $\{hkl\}$  directions, with different  $\lambda_{hkl}$ . For polycrystalline materials the average value of  $\lambda$  is used. For most of the magnetic materials  $\lambda \cong 10^{-5}$  to  $10^{-6}$ . Elastic stresses  $\sigma$  couple to magnetostriction causing an effective stress-induced anisotropy of the material:

$$K_\sigma = -\frac{3}{2}\lambda\sigma \cos^2 \phi$$

where  $\phi$  is the direction of the stress relative to the magnetization. The total anisotropy of the magnetic body is the sum of the magneto-crystalline and stress-induced parts:  $K = K_u + K_\sigma$ . Localized fluctuations of the internal stresses  $\sigma(\mathbf{r})$  cause fluctuations in the DW energy through  $K(\mathbf{r})$  and contribute to the coercivity, that is, to the critical fields for Barkhausen jumps.

If there are long range one-dimensional stress fluctuations  $\sigma(x)$  in the material, with an average spatial periodicity (wavelength)  $\Lambda$ , large compared to the DW width, then the DW energy will fluctuate as

$$\gamma_w = 4\sqrt{A[K_u + (3/2)\lambda\sigma(x)]}$$

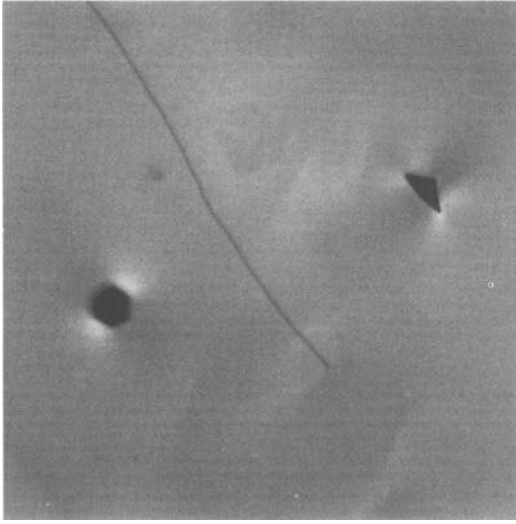
causing a critical field for the DW motion of:

$$H_{cr}^\sigma = \frac{3\lambda\sigma_w(x)}{4M_s} \frac{d\sigma(x)}{dx}$$

This is the basic equation of the stress theory of Barkhausen jumps and coercivity. If  $\sigma(x)$  is known then  $H_{cr}$  can be calculated.

Assuming a quasi-sinusoidal stress field with an amplitude of  $(\sigma_0 \pm \Delta\sigma/2)$  in the form of

$$\sigma(x) = \sigma_0 + \frac{\Delta\sigma}{2} \sin(2\pi x/\Lambda)$$



**Figure 8.** Repulsion of domain walls in a magnetic garnet material by the stress fields of triangular and hexagonal Ir inclusion, originating in the nonmagnetic substrate. The inclusion's stress field penetrates the magnetic layer and interacts with the domain walls. The sample is near saturation, only the thin black domain remains due to the very strong interaction with the stress field.

the critical field for the start of the DW motion is given by

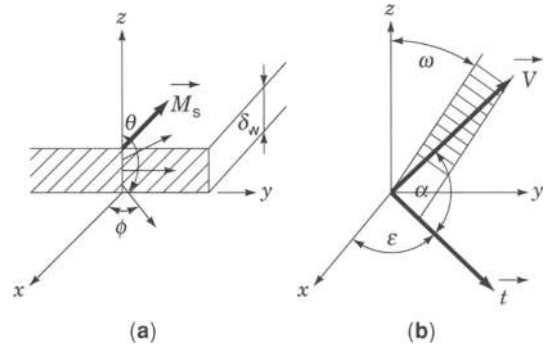
$$H_{cr}^r = p \frac{\lambda |\sigma_i|}{2M_s}$$

where  $p$  depends on the ratio of  $\sigma_w/\Lambda$ . Although the usual effect of the stress fields of the defects is that the anisotropy of the material is locally reduced, thus creating an effective energy well for the DW, it was demonstrated that it is possible to have anisotropy barriers with locally higher DW energy, preventing the change of the magnetization very efficiently (8). Figure 8 shows such a case, where the stress field from an inclusion in the nonmagnetic substrate penetrates a magnetic crystal and the DWs are repulsed by both stress fields, preventing saturation of the magnetization at fields much higher than the theoretical saturation field.

### Domain Wall Interaction with Dislocations

A special case of the stress effects on magnetization processes is due to the presence of dislocations (9,10,11,12,13,14). Plastic deformation produces defects, such as dislocations, stacking faults, and point defects, each associated with an internal stress field, which will affect the Barkhausen noise. For low deformation the number of dislocations increases linearly with shear stress  $\tau$ , and depends on the material's shear modulus  $G$  and the magnitude of the Burgers vector  $b$  characterizing the dislocation. The moving DW interacts with the stress field of the dislocation, depending on the direction and the distance between them. The dislocation exerts a force  $\mathbf{f}$  upon the DW according to its line element  $d\mathbf{l}$  and the magnetostrictive stress tensor  $\sigma$  of the DW:  $\mathbf{f} = -d\mathbf{l} \times \sigma \mathbf{b}$ . The calculation of this interaction for a general case is not trivial.

For a simple case of a  $180^\circ$  DW, lying in the  $x$ - $y$  plane and a dislocation running parallel to the plane of the DW, the geometry shown in Fig. 9, the critical field is calculated in Ref. 10.



**Figure 9.** Geometry of the DW dislocation interaction.  $M_s$  saturation magnetization,  $\delta_w$  DW width,  $V$  Burger's vector of the dislocation.

For the simple case of the change of the magnetization across the DW  $\phi = \pi/\delta_w$ ,  $\theta = 0$ . The edge and screw components of the dislocations are  $b \sin \alpha$  and  $b \cos \alpha$ . For an edge dislocation  $\alpha = \omega = 90^\circ$ , for a screw dislocation  $\alpha = 0$ . For a dislocation running parallel to the  $x$  axis  $\epsilon = 0$ . The angle between  $b$  and the  $z$  axis is  $\omega$ . The component of the force  $f_z$  is given as

$$f_z = (3/2)Gb\lambda(\sin 2\phi \cos \alpha + 2 \sin^2 \phi \sin \omega \sin \alpha)$$

The DW interacts with the dislocation only as long as the dislocation is inside of the wall, because  $f_z = 0$  for  $\phi = 0$ , consequently,  $f_{max} = (3/2)Gb\lambda$ . To move the DW by  $dz$ ,  $dW = f dz$  work should be performed. The energy necessary to get the whole DW over the single dislocation barrier is  $W = (3/2)Gb\sigma_w\lambda$ . The critical field for the DW motion from a single straight edge dislocation is:

$$H_{cr}^d = (3/2M_s)Gb\lambda$$

For a dislocation density of  $N \text{ m}^{-2}$   $H_{cr}$  will be proportional to  $N^{1/2}$ , that is, the linear defect density.

In some cases the measured  $H_{cr}$  is much higher than the expected from the above equation. This is due to the fact that, depending on material parameters, the DW might be elastically deformed, that is, stretched between two dislocations before breaking free (1,5,15). Assuming that the amplitude of DW bulging is one-quarter of the distance between dislocations, the critical field for the DW jump is:

$$H_{cr}^d = \sqrt{2N} \gamma_w / M_s$$

### Domain Wall Pinning at Inclusions and Voids

In developing new magnetic materials the control of the microstructure is of primary importance. Nonmagnetic grain boundaries are essential for magnetic recording media in decreasing the exchange coupling between grains; the properties of permanent magnets are defined through inclusions of phases with magnetic properties different from that of the matrix, while soft magnetic materials should be as homogenous as possible. In case of nonmagnetic inclusions or voids, the DW is pinned because the local DW energy is reduced by the missing volume when the DW includes the inclusion or void. At the same time, depending on the size of the defects, there will be

free magnetic poles on the surface of the inclusions and voids, producing demagnetizing (stray) fields with a significant energy contribution (10,16,17,18,19).

**Small Inclusions,  $d \ll \delta_w$ .** When a spherical inclusion of size  $d$ , volume  $v = \pi d^3/6$ , is located outside of the DW it behaves like a dipole, having a dipole energy of  $W_D \approx (1/3) M_s^2 v$ . If it is inside of the DW, then it behaves as a quadrupole, and  $W_Q = W_D/2$ . If the interaction depends only on the distance between the inclusion (or void) and DW, then  $d\gamma_w/dx \approx d^3\gamma_w/\delta_w^2$ , and the critical field for the start of the DW motion, due to the volume effect for  $d \ll \delta_w$  is given by (10):

$$H_{cr}^i = \frac{2.8\gamma_w}{M_s w} \left[ \frac{d}{\delta_w} \right]^{3/2} v^{1/2} \left[ \ln \frac{2w}{\delta_w} \right]^{1/2}$$

The contribution due to the demagnetizing field effects:

$$H_{cr}^D = \frac{2.8M_s d^{7/2}}{w \delta_w^{5/2}} v^{1/2} [\ln(2w/\delta_w)]^{1/2}$$

The critical fields  $H_{cr}^i$  and  $H_{cr}^D$  are comparable when  $d \approx \delta_w/4$ .

**Large Inclusions,  $d \gg \delta_w$**

For large inclusions there would be a very large demagnetizing field contribution due to the strong dependence on the size of the inclusions, which is disadvantageous from the point of view of energy minimalization. The consequence is that a wedge-shaped secondary domain structure is developed around large inclusions to dilute the density of free poles over a larger surface of the new DWs. The irreversible motion of the DW starts before it would be completely torn off the inclusion, at a field:

$$H_{cr}^I = \gamma_w v^{2/3} / (2M_s d)$$

The theory of inclusions was originally developed for spherical inclusions. The case of elliptical inclusions was treated in Refs. 18 and 19, showing that the strongest demagnetizing field effects are due to flat inclusions lying perpendicular to the plane of the DW. The critical field is for the case of extended planar inclusions, width  $d$ , including magnetic inclusions,  $H_{cr}$  depends on the relative magnitude of the anisotropy and exchange constants of the two materials ( $A$  and  $A'$ , and  $K$  and  $K'$ ), and it is given in Ref. 20:

$$H_{cr}^i = \frac{2Kd}{3^{3/2} M_s \delta_w} [A/A' - K'/K], \quad \text{for } d \ll \delta_w$$

$$H_{cr}^i = \frac{KA'}{2M_s A} [A/A' - K'/K], \quad \text{for } d \gg \delta_w$$

The calculation of critical fields can be solved analytically for simple cases only (21,22,23). Numerical methods of micromagnetism are very powerful in calculating the critical fields for DW pinning for any geometry and combination of defects. The most cumbersome part of the calculation, the magnetostatic fields due to the nonzero divergence of magnetization at the defects, can be handled easily by numerical methods. Details of the pinning process can be revealed by these methods, like the repulsion of the approaching DW by the dipole field of small inclusions in high magnetization materials (24,25). Figure 5 shows the domain structure of a garnet single crystal in  $H \gg$

$H_{cr}$ , near to saturation, where the strongest defects keep the last DWs pinned (26,27,28).

## IRREVERSIBLE MAGNETIZATION PROCESSES

In a ferromagnet with a domain structure irreversible magnetization takes place by DW motion, due to the sudden changes of the magnetization as the applied magnetic field reaches the critical field of depinning a DW from a defect, as given in the previous section. On reducing the size of the magnetic body down to the typical size of a single domain, no DW can be formed and, consequently, there is no DW motion, and no irreversible losses associated with it. However, there are still irreversible magnetization changes due to the sudden rotation of the magnetic moment of the single-domain particles at the switching field of the particle. The sequence of individual rotation events of particle magnetizations gives rise to the hysteresis loops of particulate materials, usually described by Preisach models (1,7,29). The energy barrier for magnetization switching is related to the effective anisotropy field of the particles:  $H_{cr}^{rot} = 2 K_{eff}/M_s$ , a fairly large value. These irreversible wall motion or rotational magnetization processes are the cause for the hysteresis losses. The measure of the energy loss associated with the hysteresis can be determined from the area under the hysteresis loop:

$$W_H = \oint_{\text{loop area}} H dM$$

The loss associated with an individual Barkhausen jump  $dW_B$  is determined in a similar way, by the area of the minor loop of one jump. Integrating over all individual jumps gives the total loss,  $W_H$ .

The shape of the hysteresis loop is expected to be determined by the sequential events of depinning the DWs from the defects, or the sequence of the individual switching events of particles by rotation. The sequence would depend only on the strength of the interaction between a defect and a DW. In this case the shape of the hysteresis loop would be reproducible and no random noise would be associated with the magnetization process. At the absolute zero of temperature  $T = 0$  K, with no interaction between the particles, this could be a valid assumption. However, the shape of the hysteresis loop and the associated losses are influenced by several other irreversible processes. First of all, at *finite temperatures* temperature fluctuations add a random noise to the magnetization process, randomly raising or decreasing the energy barriers what the moving DW should overcome in order to change the magnetization (30,31). Moreover, the height of the barrier, that is, the strength of the interaction between the DW and the defect, might depend on the temperature itself.

At any given temperature if one waits long enough there will be a finite probability that a temperature fluctuation appears, large enough to push the DW over the next barriers. This characteristic time delay between an instantaneous change in the applied field and the subsequent change of the magnetization is the magnetic aftereffect (4,30). The higher the temperature the more effective is the assistance of temperature fluctuations. The slower the change of the field, the more chance the DW has to jump over the energy barrier, waiting for the moment of the temperature fluctuation assisted decrease in the

energy barrier. The critical field of the start of the DW motion at one given defect depends on the temperature and time:  $H_{cr} = H_{cr}(T, t)$ . As a result, the faster the rate of change of the field, the larger the field needed to change the magnetization, that is, the coercivity  $H_c$  is increasing with frequency. The defect-DW interaction is influenced by the aftereffect, leading to an apparent frequency dependence.

Another effect that can't be neglected is the shape and size dependence of the magnetization curve. If a magnetic body of finite size is magnetized, free magnetic poles occur on the surface, producing demagnetizing field,  $H_d$ , which opposes the direction of the magnetization. The demagnetizing field is proportional to the magnetization,  $H_d = -NM$ , where  $N$  is the demagnetizing factor (1,2,32). When the magnetization changes in a small region of the body, the demagnetizing field changes too, and the slope of the hysteresis loops changes, depending on the demagnetizing factor and the permeability  $\mu = dB/dH = d(\mu_0 \mathbf{H} + \mathbf{M})/dH$ . This change causes an apparent change in the magnetization at a given applied field. A more serious problem is that the magnetic body is not homogeneously magnetized. The direction and the magnitude of the domain magnetizations vary, and the permeability is not a well-defined quantity. The effect of the sample's shape and size makes the comparison of experimental data very difficult.

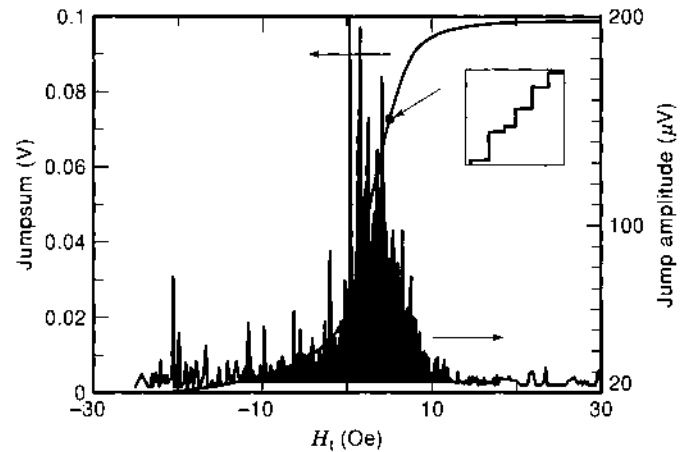
Eddy current losses can't be neglected in the case of ac magnetization processes of metallic magnets. When the magnetization is changed in a conductor, eddy currents and, associated with it, a time-dependent magnetic induction  $B(t)$  is produced. Integrating  $B(t)$  over the sample surface, the change of the flux in time, that is, the voltage induced by the eddy current is obtained. Due to interference from this eddy-current induced signal, the measurement of the Barkhausen noise in metallic magnets is not a simple task (33).

## BARKHAUSEN NOISE

Any discontinuous, sudden change of the magnetization in a slowly changing magnetic field, due to the irreversible break-away of a DW from a pinning center or an irreversible act of rotation, induces a voltage spike in a coil surrounding the specimen. The sequence of these pulses is the Barkhausen noise. The pulses are characterized by their number in the applied field interval, magnitude of the induced voltage, width of the pulse, and the integrated voltage, the *jumpsum*. The change of the magnetization during one jump is related to the integral under the area of the resultant pulse. The steepest part of the hysteresis loop, around the coercivity  $H_c$  contains the largest number and highest amplitude pulses. The defect structure influences both mechanical and magnetic properties of the materials. The distribution of jump size, shape, duration, and the power spectrum of the Barkhausen noise is directly related to the microstructure of the magnetic material, and carries very important information for nondestructive testing. Figure 10 shows the applied magnetic field dependence of the Barkhausen noise data for a steel specimen in terms of pulse heights of the induced voltage, and the integrated voltage (jumpsum), which is related to the magnetization (33).

### Stress Effects

The defect-DW interaction is very sensitive to the stress due to stress-induced anisotropy in materials with finite



**Figure 10.** Barkhausen jump statistics of a low carbon steel specimen, and the integrated jumpsum signal. The inset is the measured signal at higher magnification. (Courtesy of L. J. Swartzendruber, NIST, unpublished data.)

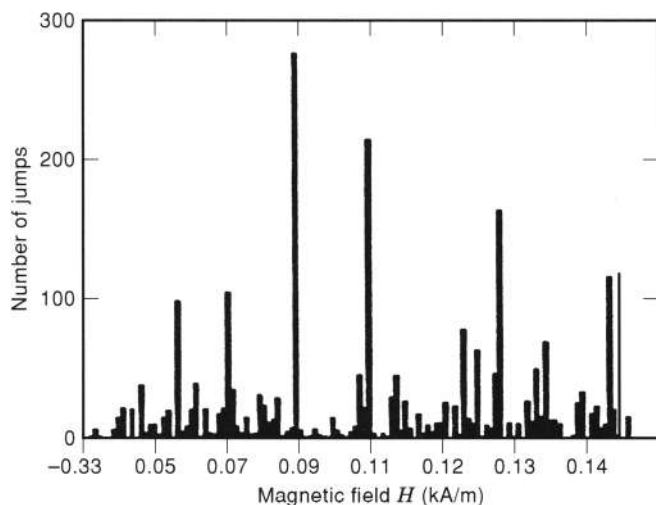
magnetostriction. Upon an application of stress the DW energy landscape and the critical field for individual DW jumps change and, as a result, the shape of the hysteresis loop and the permeability change with stress as well. After proper calibration, and over a wide range of stress, the Barkhausen noise can often be used as a measure of the strain state of the material (34,35).

In materials with strong uniaxial anisotropy, the equilibrium domain structure is relatively simple. In a stressed long wire or ribbon, domain magnetizations will be aligned along the long axis and the magnetization process is dominated by one huge Barkhausen jump. The hysteresis loop of such a material is nearly square. For a wire with large magnetostriction,  $\lambda > 0$ , under tensile stress  $\sigma$  along the axis, there is a stress-induced uniaxial anisotropy:  $K_s = (3/2)\sigma\lambda$ . The large change in magnetic induction upon changing the magnetic field will induce a large, narrow pulse of voltage in a coil wound around the sample, used in many sensors and devices. Stress can be induced in amorphous magnetostrictive wires during heat treatment, and can be used as inductors. Figures 11 and 12 show the effect of magnetostriction and mechanical stress on Barkhausen noise statistics upon magnetization of a highly magnetostrictive amorphous metallic ribbon in a free state and when the same ribbon is slightly stretched along its axis (36). The change of the Barkhausen jump spectrum from numerous small jumps to a very few huge jumps is evident. The integral of the jumps over the magnetic field gives a curve similar to the magnetization curve (37).

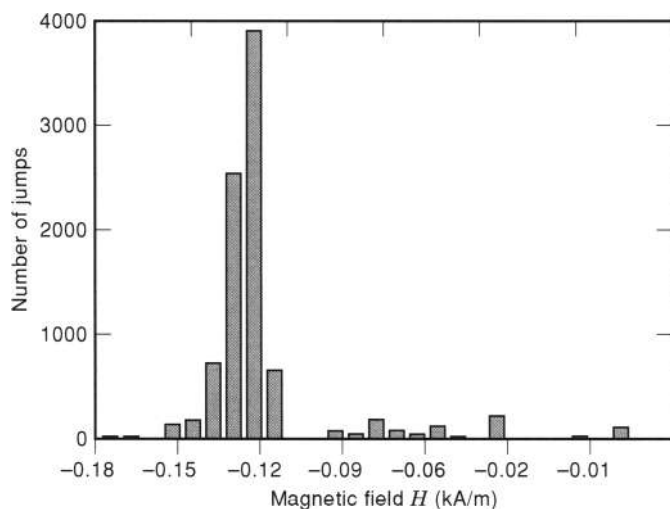
### Temperature Dependence

The Barkhausen effect has a strong temperature dependence. The DW energy itself depends on the magnetic anisotropy, a strong function of temperature, leading to the increase of the DW energy and the coercivity on lowering the temperature. The energy of temperature fluctuations play a lesser role at low temperatures, and the characteristic time constant of the magnetic aftereffect is increasing with decreasing temperature. As the probability for a jump decreases, the DW has to wait longer for a kick from the thermal fluctuations, and the number of jumps in a given field and time interval decreases (33). On





(a)



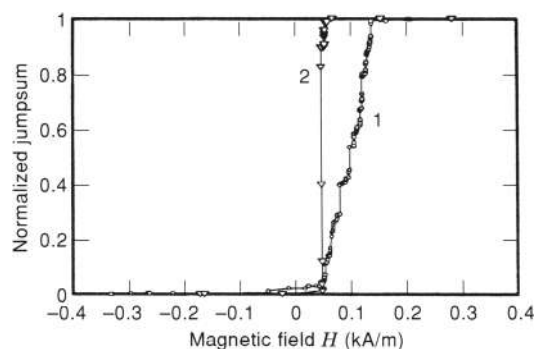
(b)

**Figure 11.** Effect of magnetostriction and mechanical stress on Barkhausen noise statistics upon magnetization of a highly magnetostrictive amorphous metallic ribbon (a) Free ribbon; (b) Ribbon under weak tensile stress (M. Pardavi-Horvath, unpublished data).

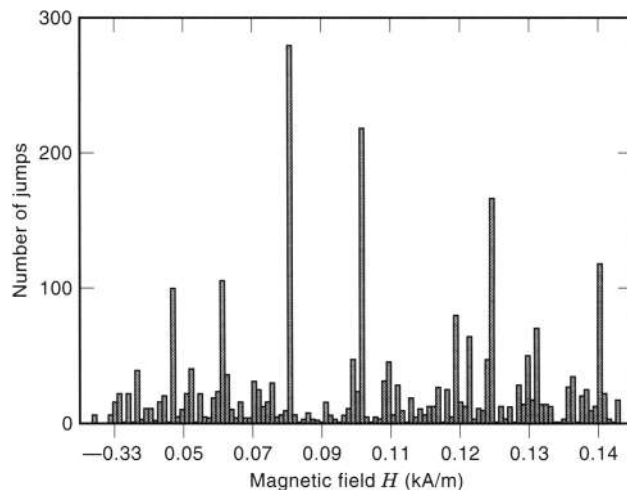
the contrary, upon increasing the temperature the coercivity decreases, and due to strong thermal activation processes, it's easier for a DW to make a jump. At the same time the thermal energy becomes larger than some of the DW-defect interaction energies, many of the low energy defects are "turned off" and the width of the defect-DW interaction spectrum decreases too. The magnetic material becomes magnetically softer.

### Statistics of Barkhausen Jumps

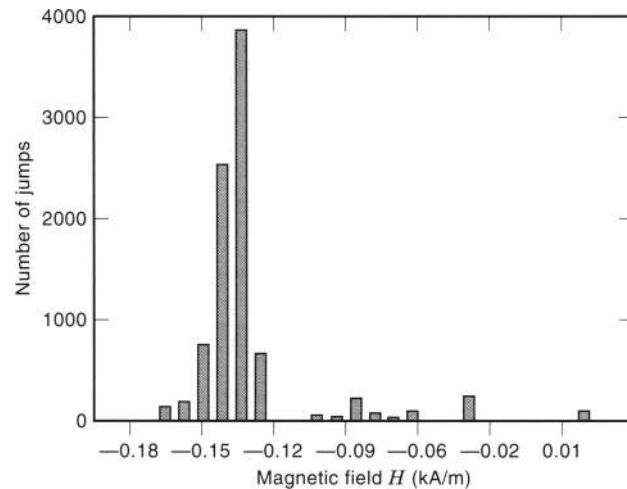
The statistics of Barkhausen jumps has been the topic of wide ranging investigations for many years. The static interaction of *one* defect with *one* domain wall is just the beginning of the theoretical treatment of the magnetization process of a magnetic material. In a real material, a very large number of statistically distributed defects interact with a large number of DWs, resulting in a complex Barkhausen noise behavior. The process depends on temperature, time, sample size, and shape (38,39). For the interpretation of the Barkhausen signal the



(a)



(b)



(c)

**Figure 12.** (a) Integrated signal of individual jumps for the sample in Fig. 12. 1 is the free ribbon; 2 is the ribbon under tensile stress. (b) Number of jumps for 1. (c) Number of jumps for 2. (M. Pardavi-Horvath, unpublished data.)

functional form of the energy landscape is needed. All these unknown parameters can be taken into account as fluctuations in the DW energy, and described by an ensemble of stochastic Langevin functions. The model generates both hysteresis loops and Barkhausen jump distributions, showing a power-law behavior for small jumps and a rapid cut-off at large jump sizes, in agreement with the experimental data (40,37,41).

Unfortunately, along the hysteresis loop different types of magnetization processes take place. Therefore, the treatment is usually restricted to the constant permeability region around the coercivity. The assumption used in predicting the noise power spectrum (42) is that both  $dB/dH$  and  $dH/dt$  are constant. This approach was extended to the whole hysteresis loop assuming that the Barkhausen activity in terms of the jumpsum is proportional to the differential susceptibility, which can be determined from hysteresis models (43).

An important question is whether the individual jumps are correlated or not. For statistically independent jumps the theory predicts that the power density is constant at low frequencies and decays as  $1/f^2$  at high frequencies. Some experimental data agree with this prediction, however, in some cases there is a maximum at low frequencies. The analysis of the Barkhausen noise signal characteristics, as autocorrelation, jump amplitude correlation, jump time and amplitude correlation, and power density show no evident deviation from random noise behavior. Evidently, a different mechanism is responsible for the low frequency losses (44).

The progress in nonlinear dynamics and chaos, relatively easily applicable to the behavior of domains and domain walls, leads to simple nonlinear models for DW dynamics, taking into account the viscous damping of the DW moving under the effect of a harmonic external force in the field of defects. It was shown that Barkhausen noise can be characterized by a low fractal dimension (45,46). Barkhausen jumps show the attributes of self-organized criticality, with the distribution of lifetimes and areas of discrete Barkhausen jumps following a power-law behavior (47,48).

### Experimental Techniques

Modern Barkhausen spectrometers are based on the same principle of Faraday's law of induced electromotive force (emf), as used by Barkhausen in 1919. According to Faraday's law, whenever there is a magnetic flux change  $dB/dt$  there will be an induced voltage  $V$ , proportional to  $dB/dt$  and the number of turns of a measuring pickup coil. The difference is that the original amplifier is now substituted by a digital computer controlled data acquisition system and digital storage oscilloscope (33,44,49). Usually the size and the duration of Barkhausen pulses is measured and statistically registered. The results are frequency distributions. The entire sequence of Barkhausen pulses can be considered as a stochastic process and then the spectral density of the noise energy, related to the total irreversible magnetization reversal is measured (33,39). Figure 13 shows the block diagram of a typical apparatus. The equipment consists of a magnetizing system used to sweep the sample through the full hysteresis loop from negative saturation to positive saturation, at a frequency of typically 0.05 Hz; a Hall probe to measure the magnetic field; a surface coil to measure the induced voltage and/or another pick-up coil, surrounding the sample, for power spectrum measurements. The measurement is controlled and the data are analyzed by a digital computer. The bandwidth of the system is from 100 Hz up to 100 kHz. A discriminator is used to reject signals below a certain noise level. At each instant of time,  $t_n$ , the value of the magnetic field,  $H_n$  and the induced voltage is  $V_n$  measured. So the amplitude of jumps versus field, the integrated signal, the jumpsum, and the rate of the jumpsum can be determined. The jumpsum rate is qualitatively similar to the rms noise. Power density

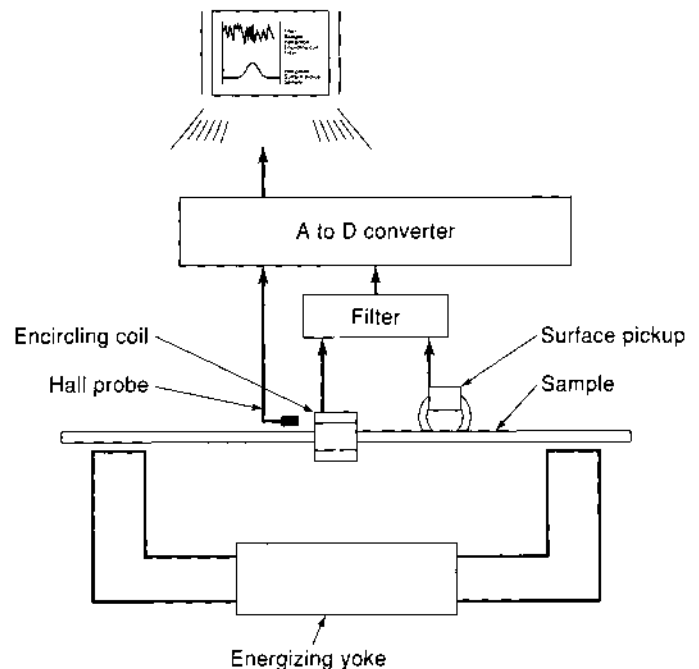


Figure 13. Block diagram of Barkhausen noise measurement. (Courtesy of L. J. Swartzendruber, NIST.)

spectra and autocorrelation functions can also be obtained from the data. For a stationary random process, the power density spectrum is the Fourier transform of the autocorrelation function,  $\Phi(f) = \langle V(t) \times V(t + \tau) \rangle$ , so that statistical properties of the Barkhausen noise can be easily investigated by this technique (49).

A widely used method in studying the Barkhausen effect is to measure the power as a function of the frequency of cycling the field. This technique yields information about the intervals between jumps (33). The frequency of the Barkhausen jumps is related to the average DW velocity, and it contains information about the average time it takes for a moving DW to reach the next barrier through the average distance between defects.

Utilizing the optical activity of magnetic materials it is possible to identify individual defects, pinning mechanisms, and measure actual DW velocities. In transparent magnetic materials individual DW-defect interactions can be investigated via the magneto-optical Faraday-effect (26,27,28), as illustrated in Figs. 2, 6, 7, and 9. In metals, the Kerr effect offers a possibility to visualize the defect-DW interaction, as it was done for several metallic film magneto-optical recording media, in order to study the cause of media noise in readout (50). It was shown that the distance between pinning centers is about  $0.4 \mu\text{m}$  and the DW "waits" up to several seconds before the next jump. The pinning time decreases exponentially with increasing driving field.

### Nondestructive Testing

When applicable the measurement of magnetic Barkhausen noise is a fast, reliable, and simple technique for nondestructive material evaluation, as compared to x-ray diffraction, ultrasound, and other more sophisticated measurements (49,51). The Barkhausen effect is especially well-suited for the study of steel, one of the most important structural materials. Commercial instruments to characterize materials based on

Barkhausen noise are now available. The Barkhausen spectrum can be, for example, related to the grain size of the material, so it can be used for grain size measurement after cold-rolling (52). The stress sensitivity of the DW motion makes it very convenient to study residual stresses, using the Barkhausen emission under stress. Depending on the bandwidth of the measurement, either the surface or near subsurface layers can be tested (34). The Barkhausen emission test is used in the grinding industry to check the residual stress due to thermal damage and microhardness (53). Plastic deformation increases the dislocation density, an effect easily detected by the DW-dislocation interaction induced critical field increase. There is a possibility to use deformed thin wires of iron and iron-nickel alloys as strain gages in the  $10^{-3}$  range with a sensitivity of  $\pm 5\%$  (35).

Barkhausen noise plays an important role in magnetic recording. The head-to-medium velocities in hard drives may range from 1 to 50 m/s, with data frequencies from 100 kHz to 50 MHz; for tape systems the velocity might be as low as 25 mm/s at 1 Hz frequency. These parameters are very near to the range of Barkhausen noise. The main source of the medium and head noise is the Barkhausen noise. The nonuniformity of media leads to a localized, regular noise, superimposed on the random noise from magnetization switching processes. For particulate materials the noise is related to the shape, size, and orientation dispersion of the particles; for multigrain thin film media the signal-to-noise ratio (SNR) depends on the microstructure, texture, and grain cluster size. In soft magnetic write-read head materials the DW pinning and the related hysteresis losses are the most important source of noise (50,54).

## FLUX PINNING AND LOSSES IN SUPERCONDUCTORS

The basic parameters defining the transition from the superconducting to the normal state are the critical temperature  $T_c$ , critical magnetic field  $H_c$ , and critical current density  $J_c$ . Type II superconductors possess the highest possible critical parameters. Destruction of their bulk superconductivity occurs at the upper critical field  $H_{c2}$ . If no applied magnetic field is present, then  $J_c$  up to  $10^{11}$  Am $^{-2}$  can be reached. High current density and small energy dissipation is the main reason for industrial applications of superconductors in magnets, electrical machinery, and power transmission lines. Ideal superconductors are lossless, however, using Type II superconductors for practical applications dissipative processes are of great importance. In Type I superconductors dissipation is negligible up to very high frequencies, where the electromagnetic field destroys the Cooper pairs, responsible for superconductivity. Losses in ideal Type II superconductors are associated with viscous motion of vortex lattices. Losses in the most important class of superconductors, the Type II nonideal superconductors, are due to magnetic hysteresis, similar to the case of ferromagnetic materials. In superconductors the irreversible motion of magnetic flux lines through pinning centers causes the hysteresis. The problem is more complicated in technical materials: superconducting cables are multifilament, multicore, twisted, stabilized composites where the coupling losses to the normal matrix play a significant role in dissipation (55).

### Vortex Structure

Superconductors (SC) are characterized by the penetration depth of the weak magnetic field  $\lambda$ , the coherence length  $\xi$ ,

and the Ginzburg-Landau parameter  $\kappa = \lambda/\xi$ . Superconductors can be classified according to the magnitude of  $\kappa$ . Type I SCs have  $\kappa < 1/\sqrt{2}$ , and the external field is totally screened by diamagnetic supercurrents flowing in the penetration depth distance from the surface (Meissner effect). For Type II SCs  $\kappa > 1/\sqrt{2}$ . The N/S boundary is in equilibrium only in high applied magnetic field. A type II SC below the first critical field  $H_{c1}$  behaves as a Type I material. In a field of  $H > H_{c1}$  it is energetically more favorable to break into the mixed state of alternating N and SC phases, consisting of thin, normal core vortex lines, parallel to the external field, with circulating paramagnetic supercurrents. Each vortex carries a quantum of flux  $\phi_0 = hc/2e$ , where  $h$  is the Planck constant,  $c$  is the light velocity and  $e$  is the electron charge. The size of a vortex is typically a few hundred nanometer. The vortex lattice is in some way analogous to the magnetic domain structure of ferromagnetic materials, as discussed previously. Upon increasing the applied magnetic field  $H$ , the number of vortices increases as the vortex lattice period becomes smaller. At  $H = H_{c2} = \sqrt{2} \kappa H_c$  the vortex cores overlap and the material becomes a normal metal.  $H_{c2}$  of SC for practical applications is on the order of  $10^2$  T.

### Vortex Motion and Flux Jumps

Many technical applications of SC require high current densities and low losses, which can be accomplished by immobilizing the vortices by pinning them at defects. The vortices have a complex magnetic field around them, decaying exponentially within the distance  $\lambda$ . If the SC is placed in an applied field  $H$ , parallel to its surface, then supercurrents  $J_m$  will be generated by a vortex located at  $x_0$ :

$$J_m = \lambda^{-1} H \exp(-x_0/\lambda)$$

and a repulsive Lorentz force will act on the vortex,  $F_m = J_m \phi_0$ . This force changes with the distance, thus changing the free energy of the vortex system. If there is an energy gradient, that is, a local change in the potential for the vortex motion, then the vortex can be pinned at the defect lowering the free energy of the system. There is an energy barrier to the penetration of vortices in and out of the SC at the boundary, and vortex motion into the SC can be prevented by surface treatment. Ideal Type II SCs are thermodynamically reversible if there are no defects in the material to act as pinning centers. But even in the absence of bulk pinning, there is always the surface, itself an irregularity, so there is no ideal SC. In the absence of pinning centers, the vortex lattice's motion is like a hydrodynamic flow in a viscous medium. Dissipation is due to the viscous damping process. The power necessary to move the vortices is  $W_{SC} = \rho_n J_s^2$ , where  $\rho_n$  is the specific resistance. Because there are always thermal fluctuations, and  $J_c$  is decreasing with increasing  $T$ , any slight increase in the temperature decreases the critical current and, by preserving the total current, the current distribution has to change, and as a consequence, the magnetic field distribution generated by the current distribution is changes too. According to Maxwell's law whenever there is a flux change there will be an electric field, and this field does the work to move the vortices, that is, a flux jump is observed. Flux jump instabilities are very dangerous because they can transform the SC into the N state completely. Thermally activated flux flow is one of the main reasons for the low current density in new high- $T_c$  SCs.

Another case of flux jumping occurs due to magnetic field fluctuations. As the applied field increases, critical shielding currents are induced to prevent the penetration of flux into the SC until a certain field  $H$  (or magnetic flux density associated with that field  $B$ ) value, where flux jumps occur. For thin filaments the flux completely penetrates the material, and there is no flux jumps even for applied fields much greater than  $H$ . The constraint on the thickness  $d$  of the SC material is that:

$$\mu_0 J_c d < B^* = [3\mu_0 S(T_c - T)]^{1/2}$$

where  $S$  is the volume heat capacity of the material (56). The distribution of the induction is usually described by the Bean-London model of critical state, where the magnitude of the critical current is directly related to the gradient of the local induction  $J_c = dB/dx$ . A change in the penetration occurs only when the change in the external field exceeds the surface shielding fields created by shielding currents, according to Maxwell's equations. This picture of SC hysteresis losses is valid up to about  $10^3$  Hz. In the Bean-London model the contribution of the vortex structure is not included. A complete analytical solution for the distribution of current and induction in hard SCs is still missing.

#### Nonideal Type II SCs

Real materials always have spatial inhomogeneities acting as pinning centers that prevent the penetration of flux into the SC upon increasing the applied field; or preventing the exclusion of flux upon decreasing field, giving rise to hysteresis and associated irreversible losses. There is a remanence flux frozen in the SC even in the absence of external field. In analogy with hard magnets, these nonideal Type II SCs are called *hard SCs*.

The pinning centers are similar to those in magnetic materials, preventing the motion of domain walls, such as: point defects of inclusions, voids, precipitates of second phase, line defects like dislocations, grain boundaries, twin boundaries, and so on. The bigger the difference between the properties of the defect and the SC the larger is the pinning effect, again similar to the case of the strong DW pinning for high DW energy gradient in ferromagnets.

There is no complete, exact theory of pinning in SCs, and the methods used to describe the vortex-defect interactions are very similar to those dealing with DW pinning. Strong pinning and high current density is achieved mostly by microstructural developments through technology of preparation, compositional modifications, and different treatments (for example, irradiation). In intermetallic SCs ( $\text{Nb}_3\text{Sn}$ ,  $\text{Nb}_3\text{Ge}$ ) the pinning centers are related to the fine grained structure on the order of a few nm, and precipitation of oxides or carbides at grain boundaries. In the high- $T_c$  cuprate SCs vortex pinning is linked to the layered structure of the material, the vortices are interacting strongly within the layers (pancake vortices) and weakly between layers. As a result the electrical properties are highly anisotropic. The pancake vortices are very mobile. Increasing the interlayer coupling would make a more effective pinning with higher (and isotropic) current densities (57).

In an ideal SC the vortex structure is a regular two-dimensional network. Due to the very inhomogeneous structure of the hard SCs, the vortex lattice is distorted, following the

microstructural features, that is, the energy landscape of the material. For high defect densities and high current densities the vortex lattice becomes "amorphous", the lattice is "melted".

#### BIBLIOGRAPHY

1. Soshin Chikazumi and Stanley H. Charap, *Physics of Magnetism*, New York: Wiley, 1964.
2. Richard M. Bozorth, *Ferromagnetism*, New York: IEEE Press, 1993.
3. Ami E. Berkowitz and Eckart Kneller (ed.), *Magnetism and Metallurgy*, New York: Academic Press, 1969.
4. H. Kronmüller, *Magnetisierungskurve der Ferromagnetika I*. In Alfred Seeger (ed.), *Moderne Probleme der Metallphysik*, Vol. 2, Ch. 8, Berlin: Springer Verlag, 1966.
5. H. Träuble, *Magnetisierungskurve der Ferromagnetika II*. In Alfred Seeger (ed.), *Moderne Probleme der Metallphysik*, Vol. 2, Ch. 9, Berlin: Springer Verlag, 1966.
6. M. Pardavi-Horvath and Hyunkyu Kim, Surface roughness effects on the coercivity of thin film heads. *J. Korean Magn. Soc.*, **5**: 663–666, 1995.
7. M. Pardavi-Horvath, A simple experimental Preisach model system. In G. Hadjipanayis (ed.) *Magnetic Hysteresis in Novel Magnetic Materials*, Amsterdam: Kluwer Publ., 1997.
8. J. A. Jatau, M. Pardavi-Horvath, and E. Della-Torre, Enhanced coercivity due to local anisotropy increase. *J. Appl. Phys.*, **75**: 6106–08, 1994.
9. H. Träuble, in Ref. 3, pp. 622–685.
10. H. Träuble, in Ref. 5, pp. 257–279.
11. Martin Kersten, Über die Bedeutung der Versetzungsdichte für die Theorie der Koerzitivkraft rekristallisierter Werkstoffe, *Z. Angew. Phys.*, **8**: 496–502, 1956.
12. K.-H. Pfeffer, Zur Theorie der Koerzitivfeldstärke und Anfangsuszeptibilität, *Phys. Stat. Sol.*, **21**: 857–872, 1967.
13. K.-H. Pfeffer, Wechselwirkung zwischen Versetzungen und ebenen Blochwänden mit starrem Magnetisierungsverlauf, *Phys. Stat. Sol.*, **19**: 735–750, 1967.
14. K.-H. Pfeffer, Mikromagnetische Behandlung der Wechselwirkung zwischen versetzungen und ebenen Blochwänden, *Phys. Stat. Sol.*, **21**: 837–856, 1967.
15. Horst-Dietrich Dietze, Theorie der Blochwandwölbung mit Streufeldeinfluss. *Z. Phys.*, **149**: 276–298, 1957.
16. L. J. Dijkstra and C. Wert, Effect of inclusions on coercive force of iron, *Phys. Rev.*, **79**: 979–985, 1950.
17. R. S. Tebble, The Barkhausen effect. *Proc. Phys. Soc., London*, **B86**: 1017–1032, 1995.
18. E. Schwabe, Theoretische Betrachtungen über die Beeinflussung der ferromagnetischen Koerzitivkraft durch Einschlüsse mit rotationellelliptischer Form, für den Fall, dass deren Abmessungen klein gegen die Dicke Blochwand sind, *Ann. Physik*, **11**: 99–112, 1952.
19. K. Schröder, Magnetisierung in der Umgebung unmagnetischer Einschlüsse in Ferromagnetika, *Phys. Stat. Sol.*, **33**: 819–830, 1969.
20. D. I. Paul, Extended theory of the coercive force due to domain wall pinning, *J. Appl. Phys.*, **53**: 2362–2364, 1982.
21. P. Gaunt, Ferromagnetic domain wall pinning by a random array of inhomogeneities, *Phil. Mag. B*, **48**: 261–276, 1983.
22. Xinh Chien and P. Gaunt, The pinning force between a Bloch wall and a planar pinning site in MnAlC, *J. Appl. Phys.*, **67**: 2540–2542, 1990.
23. Wolfgang Prause, Energy and coercive field of a porous ferromagnetic sample with Bloch walls. *J. Magn. Magn. Mater.*, **10**: 94–96, 1979.

24. M. A. Golbazi et al., A study of coercivity in Ca-Ge substituted epitaxial garnets, *IEEE Trans. Magn.*, **MAG-23**: 1945, 1987.
25. E. Della Torre, C. M. Perlov, and M. Pardavi-Horvath, Comparison of coercivity calculations of anisotropy and exchange wells in magneto-optic media, *J. Magn. Magn. Mater.*, **104-107**: 303-304, 1992.
26. M. Pardavi-Horvath, Defects and their avoidance in LPE of garnets, *Progress in Crystal Growth and Characterization*, **5**: 175-220, 1982.
27. M. Pardavi-Horvath, Coercivity of epitaxial magnetic garnet crystals, *IEEE Trans. Magn.*, **MAG-21**: 1694, 1985.
28. M. Pardavi-Horvath and P. E. Wigen, Defect and impurity related effects in substituted epitaxial YIG crystals, *Advances in Magneto-Optics, J. Magn. Soc. Jpn.*, **11**: S1, 161, 1987.
29. Franz Preisach, Untersuchungen über den Barkhauseneffekt, *Ann. Physik*, **5**: 737-799, 1929.
30. Isaak D. Mayergoyz and Can E. Korman, Preisach model with stochastic input as a model for magnetic viscosity, *J. Appl. Phys.*, **69**: 2128-2134, 1991.
31. I. D. Mayergoyz and G. Friedman, The Preisach model and hysteretic energy losses, *J. Appl. Phys.*, **61**: 3910-3912, 1987.
32. Xiaohua Huang and M. Pardavi-Horvath, Local demagnetizing tensor calculation for rectangular and cylindrical shapes, *IEEE Trans. Magn.*, **32**: 4180-4182, 1996.
33. John C. McClure, Jr. and Klaus Schröder, The magnetic Barkhausen effect, *CRC Critical Rev. Solid State Sci.*, **6**: 45-83, 1976.
34. D. C. Jiles, P. Garikepati, and D. D. Palmer, Evaluation of residual stress in 300M steels using magnetization, Barkhausen effect and X-ray diffraction techniques. In Donald O. Thompson and Dale E. Chimenti (Eds.) *Rev. Progr. Quantitative Nondestructive Evaluation*, **8B**: 2081-2087, Plenum Press: New York, 1989.
35. A. H. Wafik, Effect of deformation on Barkhausen jumps of fine wires of iron, nickel and iron-nickel alloy, *J. Magn. Magn. Mater.*, **42**: 23-28, 1984.
36. M. Pardavi-Horvath, unpublished data.
37. R. D. McMichael, L. J. Swartzendruber, and L. H. Bennett, Langevin approach to hysteresis and Barkhausen jump modeling in steel, *J. Appl. Phys.*, **73**: 5848-5850, 1993.
38. Richard M. Bozorth, Barkhausen effect in iron, nickel and permalloy. I. Measurement of discontinuous change in magnetization, *Phys. Rev.*, **34**: 772-784, 1929.
39. U. Lieneweg and W. Grosse-Nobis, Distribution of size and duration of Barkhausen pulses and energy spectrum of Barkhausen noise investigated on 81% nickel-iron after heat treatment, *Int. J. Magn.*, **3**: 11-16, 1972.
40. B. Alessandro et al., Domain-wall dynamics and Barkhausen effect in metallic ferromagnetic materials, I. Theory, *J. Appl. Phys.*, **68**: 2901-2907, 1990.
41. F. Y. Hunt and R. D. McMichael, Analytical expression for Barkhausen jump size distributions, *IEEE Trans. Magn.*, **30**: 4356-4358, 1994.
42. G. Bertotti, G. Durin, and A. Magni, Scaling aspects of domain wall dynamics and Barkhausen effect in ferromagnetic materials, *J. Appl. Phys.*, **75**: 5490-5492, 1994.
43. D. C. Jiles, L. B. Sipahi, and G. Williams, Modeling of micromagnetic Barkhausen activity using a stochastic process extension to the theory of hysteresis, *J. Appl. Phys.*, **73**: 5830-5832, 1993.
44. L. J. Swartzendruber et al., Barkhausen jump correlations in thin foils of Fe and Ni, *J. Appl. Phys.*, **67**: 5469-5471, 1990.
45. B. Alessandro, G. Bertotti, and A. Montorsi, Phenomenology of Barkhausen effect in soft ferromagnetic materials, *J. Physique*, **49** (C8): 1907-1908.
46. H. Yamazaki, Y. Iwamoto, and H. Maruyama, Fractal dimension analysis of the Barkhausen noise in Fe-Si and permalloy, *J. Physique*, **49** (C8): 1929-1930.
47. P. J. Cote and L. V. Meisel, Self-organized criticality and the Barkhausen effect, *Phys. Rev. Lett.*, **67**: 1334-1337, 1991.
48. J. S. Urbach, R. C. Madison, and J. T. Markert, Reproducibility of magnetic avalanches in an Fe-Ni-Co magnet, *Phys. Rev. Lett.*, **75**: 4964-4967, 1995.
49. L. J. Swartzendruber and G. E. Hicho, Effect of sensor configuration on magnetic Barkhausen observations, *Res. Nondestr. Eval.*, **5**: 41-50, 1993.
50. S. Gadetsky and M. Mansuripur, Barkhausen jumps during domain wall motion in thin magneto-optical films, *J. Appl. Phys.*, **79**: 5667-5669, 1996.
51. G. V. Lomaev, V. S. Malyshey, and A. P. Degterev, Review of the application of the Barkhausen effect in nondestructive inspection, *Sov. J. Nondestructive Testing*, **20**: 189-203, 1984.
52. S. Titto, M. Ojala, and S. Säynäjäkangas, Non-destructive magnetic measurement of steel grain size. *Non-Destructive Testing*, **9**: 117-120, 1976.
53. H. Gupta, M. Zhang, and A. P. Parakka, Barkhausen effect in ground steel, *Acta Mater.*, **45**: 1917-1921, 1997.
54. C. Denis Mee and Eric D. Daniel (eds.), *Magnetic Recording*, New York: McGraw-Hill, 1987.
55. V. Kovachev, *Energy Dissipation in Superconducting Materials*, Clarendon Press: Oxford, 1991.
56. Lawrence Dresner, *Stability of Superconductors*, Plenum Press: New York, 1995.
57. George W. Crabtree and David R. Nelson, Vortex physics in high-temperature superconductors, *Physics Today*, **38-45**: April 1997.

#### READING LIST

A rich source of up-to-date information on Barkhausen noise, magnetic domain wall pinning, and hysteresis losses are the issues of the *J. Appl. Phys.*, and *IEEE Trans. Magn.*, publishing the material of the annual conferences on magnetism. The *IEEE Trans. Appl. Supercond.* is suggested as a source of current information on flux pinning in superconductors.

MARTHA PARDAVI-HORVATH  
The George Washington University

#### MAGNETIC REFRIGERATION

The application and subsequent removal of a magnetic field causes cooling in certain materials—the magnetocaloric effect (or magnetic refrigeration). In the laboratory it is possible to pump over liquid helium and so provide considerable cooling power at ~1 K. For this reason magnetic refrigeration has usually been used to reach temperatures below 1 K. It is also the case that the high magnetic fields required are usually provided through superconducting magnets which must operate below ~5 K. There are two types of magnetic refrigerator, adiabatic demagnetization refrigerators (ADR) and nuclear demagnetization refrigerators. The ADR uses the interaction of electrons with an applied magnetic field whereas nuclear demagnetization uses the interactions of the nuclei. The temperature region of operation for an ADR depends on the magnetic material used. This can range from tens of Kelvin down to mK (1); whereas for a nuclear demagnetization refrigerator the region is 10 mK to

several microkelvin. Cooling power (the amount of energy that can be absorbed) decreases strongly with decreasing temperature. For an ADR cooling powers in the region of microwatts at low temperature can be obtained whereas for nuclear demagnetization it is nanowatts or less. Such low cooling powers from nuclear demagnetization refrigerators leads to extreme measures in order to avoid unwanted heat entering the system, for example, seismic isolation. The use of nuclear demagnetization refrigerators is therefore restricted to specialized low temperature research laboratories. For ADRs with their higher cooling power such extreme measures are not required, making such refrigerators easier to use in a more typical laboratory environment, so that they are the more common of the two.

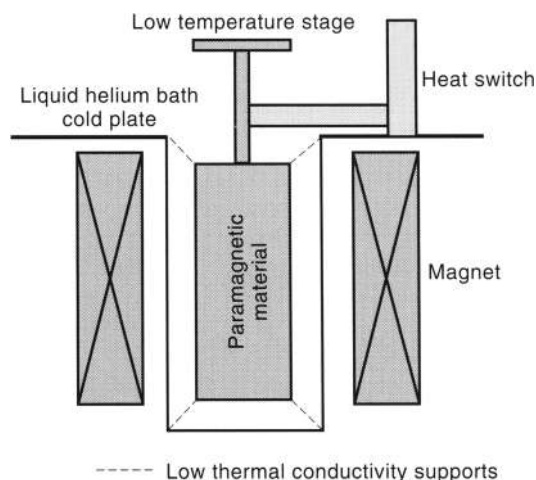
### Principle of Magnetic Refrigeration

A cooling process may be regarded to be an entropy reducing process. Since entropy (or degree of disorder) of a system at constant volume or constant pressure decreases with decreased temperature, cooling can be achieved within a medium via any process which results in the decrease of entropy of that medium. For example, the liquefaction of gases is achieved by the isothermal reduction of entropy through compression of a volume  $V_1$  at temperature  $T_1$  to a smaller volume  $V_2$  generating heat, which is extracted by contact with a cold reservoir, followed by adiabatic or isentropic expansion which results in cooling of the gas to below  $T_1$ . In the magnetic cooling process the disordered collection of magnetic dipoles associated with a particular ion within a medium (paramagnetic material) constitutes such a system described here. For such a material the application of a magnetic field causes alignment of the dipoles with the magnetic field and thus a reduction in entropy. The dipoles used are either electronic (electron cooling due to the electron spins) or nuclear (nuclear cooling due to the nuclear spins) depending on the required final temperature, millikelvin temperatures for electronic (ADRs) and microkelvin temperatures for nuclear demagnetization refrigerators. The principle of operation is the same the main difference being the starting temperature. For electron cooling starting temperatures of up to 20 K or more can be used whereas for nuclear cooling a starting temperature in the region of 0.01 to 0.02 K is needed in order for the magnetic interaction to dominate over the thermal energy.

### Components of a Magnetic Refrigerator

An ADR is essentially composed of the following three items which need to be housed in a cryostat (2) to provide a bath temperature:

1. A paramagnetic material. This is suspended via low thermal conductivity supports within an enclosure at the bath temperature, usually pumped, liquid helium in laboratory systems. The paramagnetic material is integrated with a stage or platform, upon which the experimental items under investigation may be mounted.
2. A magnet. This may either be a permanent or superconducting magnet. The latter is usually used due to their ease of operation and compactness. It is housed within the liquid helium vessel and so sufficiently cooled for superconductivity.
3. A heat switch. This is used to make and break a high thermal conductivity path between the paramagnetic



**Figure 1.** Conventional ADR schematic. A paramagnetic material is shown in a pill arrangement with a stage where samples can be attached. The pill is housed in the bore of a superconducting magnet via low thermally conducting supports. The magnet is housed in a liquid helium container and the heat switch is used to cool the pill to the temperature of the liquid helium.

material and the cold bath (e.g., liquid helium). This is used in order to cool the paramagnetic material to the starting temperature and to extract the heat of magnetization. The form of this switch may be mechanical, gaseous, or superconducting (2).

A schematic of a "classical" ADR is shown in Fig. 1. Variations on the classical form arise due to requirements to (1) increase the low temperature hold time and (2) increase the bath temperature. Increasing the low temperature hold time can be achieved by reducing the parasitic heat leak through the supports by using a second higher temperature intermediate paramagnetic material. This intercepts the heat flow from the bath to the low temperature paramagnetic material. The intermediate material comprises a higher temperature paramagnetic material which has a high heat capacity, for example, gadolinium gallium garnet. The operation of the ADR is the same as the conventional form except the intermediate paramagnetic material will demagnetize to a temperature between that of the bath and the low temperature material. Such a refrigerator is commonly called a two stage ADR (3). Increasing the bath temperature without decreasing the hold time can be accomplished by using a second ADR to cool a classical or two stage ADR. It is essentially two ADRs in series, one configured for high temperatures and the other for low temperatures. The high temperature ADR is used to cool the low temperature ADR prior to the demagnetization of that stage. This effectively simulates a lower temperature bath for a conventional or two stage ADR. This kind of refrigerator has been called a double ADR (4).

A further variation of the classical ADR is the "hybrid". This consists of a conventional ADR coupled to a low temperature cryogenic stage provided by  $\text{He}^3$ . The  $\text{He}^3$  stage is used to provide a temperature of 0.3 K reducing the parasitic heat load and providing a very low starting temperature for the paramagnetic material.

The composition of a nuclear demagnetization refrigerator is essentially the same as a conventional ADR except that its bath temperature must be in the region of 0.01 K. The bath is

provided by either an ADR or, more typically, a  $\text{He}^3\text{-He}^4$  dilution refrigerator (2,5).

## HISTORY AND CURRENT STATUS OF MAGNETIC REFRIGERATORS

Cooling by affecting the electron spins was proposed by Debye (6) in 1926 and Giauque (7) in 1927. The first practical demonstration was by De Haas, Wiersma, and Kramers (8), Giauque and MacDougall (9) in 1933, and Kurti and Simon (10) in 1934. The nuclear demagnetization refrigerator was suggested by Gorter in 1934 and Kurti and Simon in 1935. However it was not until 1956 when Kurti et al. successfully obtained cooling from 12 mK to 20  $\mu\text{K}$  via the demagnetization of nuclear spins. Since those pioneering days numerous texts on magnetic refrigeration have been published and most low temperature physics books contain a chapter describing it. For further information see Refs. 2, 5, and 11.

The advent of dilution refrigerators in the 1960s saw the demise of ADRs because of the higher cooling power and continuous operation offered by dilution refrigeration. In recent years ADRs have become more popular due to the increased use of low temperatures (0.1 to 0.01 K) for the operation of detectors for astronomy and particle physics and the development of mechanical coolers to replace the use of liquid helium (12). In astronomy the need for better signal to noise leads to low temperature detectors (10 to 100 mK) and exotic telescope locations (Antarctica, on top of high mountains, and in space). X-ray astronomy can only be conducted by space-borne instrumentation due to the absorption of X rays by the earth's atmosphere. The cost and practicalities involved require low mass systems which are gravity independent and highly reliable. The use of a consumable cryogen in space—helium—limits the duration of missions. For infrared and optical astronomy space also has enormous advantages in the form of "seeing" and sky background.

The first demonstration of an ADR in space occurred with a sounding rocket flight in 1996 (13). This was a conventional ADR comprising a liquid helium cooled magnet and an ferric ammonium alum (FAA) salt pill housed in a pumped liquid helium cryostat operating at 2 K. A similar ADR is under development (14) at the Goddard Space Flight Center for the X-ray spectrometer (XRS) experiment due to fly on the Japanese ASTRO-E mission in 2000. Advances in miniaturizing ADRs and increasing their operating temperature range further are ongoing and brought about by real-use practicalities. Such development is being achieved through new materials (especially for thermal isolation), new magnet technology (low current), and high temperature paramagnetic materials. Until recently the superconducting magnet used to generate the magnetic field would have to be housed in the liquid helium bath. The advent of conduction cooled superconducting magnets, in which the magnet is cooled via conduction through the magnet housing, enables the ADR to be a unit which is simply attached to a cold plate. This enables the ADR to be connected to a mechanical cooler thereby eliminating the need for a liquid helium bath to cool the magnet and provide the bath temperature for the ADR. Cryogen-free operation could enable the ADR to become a general purpose instrument capable of autonomous computer control without the need for the user to have either cryogenic experience or special helium handling equipment.

## THEORY

Certain paramagnetic materials are suitable for use as magnetic refrigerants. The magnetic ions of these materials have an interaction energy,  $\epsilon$  with their crystalline environment and each other which is smaller than the average thermal energy  $kT$ . In such a situation each magnetic ion is relatively "free" resulting in a distribution of randomly oriented dipoles with  $2J + 1$  degeneracy, where  $J$  is the angular momentum quantum number, that is, there are  $2J + 1$  possible orientations of the ions. This gives an  $R \ln(2J + 1)$  per mole contribution to the entropy of the material from the magnetic dipoles, where  $R$  is the gas constant. The entropy ( $S$ ) of a paramagnetic material can be thought of comprising two components. One arises from the magnetic ion (e.g., chrome in chromium potassium alum, CPA) and is given the subscript  $m$  ( $S_m$ ). The other component arises from the rest of the molecule and is refer to as the lattice ( $S_{\text{lattice}}$ ). The total entropy is given by

$$S = S_m + S_{\text{lattice}}$$

As the temperature of the paramagnetic material is reduced the lattice contribution to the entropy of the material reduces and a point is reached where the magnetic entropy given by  $R \ln(2J + 1)$  dominates. As the temperature decreases further the entropy will remain at the value given by  $R \ln(2J + 1)$  until the thermal energy approaches the interaction energy  $\epsilon$  at which point spontaneous ordering of the dipoles occurs, due to their own weak magnetic fields, and the entropy falls. When  $\epsilon \sim k\theta$ , where  $\theta$  is the magnetic ordering temperature of the material (or Néel temperature), the entropy drops rapidly. At very low temperature the internal interactions between ions removes the degeneracy and the system resides in a singlet ground state of zero entropy. At a temperature greater than  $\theta$  the entropy of the spin system of the magnetic ions can be reduced significantly if the interaction of the dipoles and the applied magnetic field is greater than the thermal excitation given by  $kT$ .

If the internal interaction energy is very low the magnet ions can be considered as free and the entropy ( $S$ ) of a collection of magnetic ions can be given by Eq. 1. Figure 2 shows the typical form of the entropy curve as a function of temperature and applied magnetic field:

$$\frac{S(B, T)}{R} = \ln \left[ \frac{\sinh(2J + 1)x/2}{\sinh(x/2)} \right] + \frac{x}{2} \coth \left( \frac{x}{2} \right) - \frac{(2J + 1)}{2} x \coth((2J + 1)x/2) \quad (1)$$

where

$$x = g\beta B/kT \quad (2)$$

and

$g$  = spectroscopic splitting factor

$\beta$  = Bohr magnetron

$B$  = magnetic field (Gauss)

$k$  = Boltzmann constant

$T$  = temperature (K)

This equation is only valid if the ions are free and for high (>1 T) magnetic fields. For low magnetic fields additional terms which take into account the weak internal field have to be added. The zero magnetic field entropy curve is determined from

$$S = \int C/T dT$$

where

$S$  = total entropy  
 $C$  = heat capacity  
 $T$  = temperature

A detailed account of paramagnetic theory as applied to magnetic refrigeration can be found in Ref. 11.

### PARAMAGNETIC MATERIALS

The paramagnetic material "the refrigerant" is the core of the refrigerator. It determines the temperature to which cooling can be achieved and how much energy can be absorbed. Cooling over a wide temperature range is possible due to the varied materials available. These can be divided roughly into four temperature ranges (~10 to 40 mK, 40 to >100 mK, >0.3 K to 1 K, and >1 K). The substances widely used in these temperature ranges are detailed in Table 1. A detailed review of the first three temperature ranges can be found in Ref. 11 and for the fourth, high temperature range (15).

#### Refrigerant Construction (Salt Pill)

Traditionally, the refrigerant stages in an ADR have been called salt pills because in the early days all refrigerants were made from a "salt". Although other forms of paramagnetic refrigerant are now available the term salt pill has remained in common use and now, incorrectly, refers to the whole refrigerant assembly used in the ADR. The refrigerant materials commonly used and listed in Table 1 are hydrated salts (CMN, CPA, CCA, FAA, and MAS), garnets ( $O_{12}$  compounds) or perovskites ( $AlO_3$  compounds). Salt pills comprised of the following four components, (1) a stage, to which items can be attached for cooling, (2) A thermal bus connecting the stage to the paramagnetic material, (3) the paramagnetic material, and (4) a container for the paramagnetic material. The stage and thermal bus are made

from high thermally conducting material, usually copper, and the container is either a high electrical resistivity material in order to minimize eddy current heating, for example, stainless steel, or fibre glass. Garnets and perovskites, which are very large dense crystals, can be bonded via an epoxy or low temperature cement to the thermal bus. In the case of hydrated salts the thermal bus comprises many hundreds of high thermally conducting (copper) wires attached to the stage and spreading into the salt enclosure. The hydrated salts are produced via an aqueous solution and the slow evaporation of the water. They fall into two categories, those that can be readily grown around copper wires and those that cannot. In the case where the salt will not grow around wires (CPA) small crystals have to be grown in a vessel (beaker) and then compressed along with the wires to form a solid mass which is then sealed in the pill enclosure. For the salts which will readily grow around wires (CMN, CCA, FAA, and MAS) the aqueous solution is allowed to evaporate from the enclosure which holds the wires and thermal bus. For CMN, where the orientation of the crystals to the magnetic field is important, seed crystals have to be cemented in place first with the correct orientation. The salts listed in Table 1 are all corrosive to copper and so if used this metal has to be plated with gold. A thickness of 30 to 50  $\mu\text{m}$  is usually sufficient. The hydrated salts have to be hermetically sealed within the enclosure in order to ensure that water of crystallization is not lost when exposed to vacuum. Such sealing can be achieved by the use of epoxy, for example, Stycast 2850.

### MAGNETIC REFRIGERATOR OPERATION

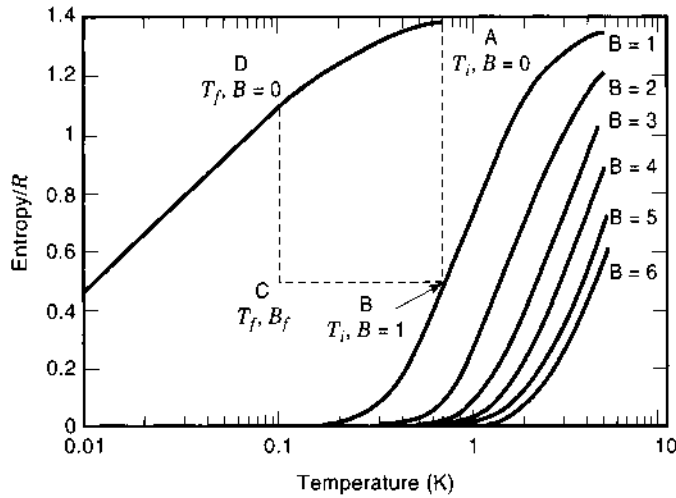
The operation of a magnetic refrigerator requires the system to be at a temperature in which the lattice entropy does not dominate the paramagnetic material in order that the applied magnetic field can reduce the entropy of the material. The process of cooling can be separated into three stages. The first is isothermal magnetization of the paramagnetic material at a temperature  $T_1$ , transferring the paramagnetic material from point A (Fig. 2) to point B. This process generates heat (magnetization energy) which has to be extracted to a heat sink at a temperature of  $T_1$ , via a heat switch. The magnetization energy ( $Q$  Joules per mole) is given by

$$Q = \int T dS \quad (3)$$

**Table 1. Magnetic Details of Some Paramagnetic Materials**

| Temperature Range | Material                         | Formation   | $J$              | $g$ | $T_n$ (K) |      |
|-------------------|----------------------------------|---|------------------|-----|-----------|------|
| 10-40 mK          | Cerium magnesium nitrate (CMN)   | $\text{Ce}_2\text{Mg}_3(\text{NO}_3)_{12} \cdot 24\text{H}_2\text{O}$ | 1/2              | 2   | ~0.01     |      |
| 40->100 mK        | Chromic potassium alum (CPA)     | $\text{CrK}(\text{SO}_4)_2 \cdot 12\text{H}_2\text{O}$                | 3/2              | 2   | ~0.01     |      |
|                   | Cesium chromic alum (CCA)        | $\text{CsCr}(\text{SO}_4)_2 \cdot 12\text{H}_2\text{O}$               | 3/2              | 2   | ~0.01     |      |
|                   | Ferric ammonium alum (FAA)       | $\text{FeNH}_4(\text{SO}_4)_2 \cdot 12\text{H}_2\text{O}$             | 4/2              | 2   | ~0.03     |      |
|                   | Manganese ammonium sulfate (MAS) | $\text{MnSO}_4(\text{NH}_4)_2\text{SO}_4 \cdot 6\text{H}_2\text{O}$   | 5/2              | 2   | ~0.1      |      |
| >0.3 K            | Dysprosium gallium garnet (DGG)  | $\text{Dy}_3\text{Ga}_5\text{O}_{12}$                                 | 1/2              | 8   | ~0.4      |      |
|                   | Erbium orthoaluminate (ErOA)     | $\text{ErAlO}_3$  | 1/2              | 9   | ~0.6      |      |
|                   | Ytterbium orthoaluminate (YbOA)  | $\text{YbAlO}_3$  | 1/3              | 7   | ~0.8      |      |
|                   | Gadolinium gallium garnet (GGG)  | $\text{Gd}_3\text{Ga}_5\text{O}_{12}$                                 | 7/2              | 2   | ~0.8      |      |
|                   | Dysprosium aluminum garnet (DAG) | $\text{Dy}_3\text{Al}_5\text{O}_{12}$                                 | 1/2              | 11  | ~2.5      |      |
|                   | Dysprosium orthoaluminate (DOA)  | $\text{DyAlO}_3$  | 1/2              | 14  | ~3.5      |      |
|                   | >1 K                             | Gadolinium orthoaluminate (GOA)                                       | $\text{GdAlO}_3$ | 7/2 | 2         | ~3.8 |





**Figure 2.** Typical behavior of entropy with temperature and magnetic field, with the operational sequence of an ADR. The sequence starts at point A, proceeding to point B on application of the magnetic field (heat switch engaged). Cooling, transfer to point C, occurs with the partial removal of the magnetic field (heat switch open). The temperature is held (transferring from C to D over time) by further reduction of the magnetic field at the correct rate.

and since  $T$  is constant

$$Q = T_1[S_A - S_B] \quad (4)$$

Adiabatic demagnetization forms the second stage in which the magnetic field is reduced to a value  $B_f$  which corresponds to the desired final temperature  $T_f$ . During this stage the entropy of the paramagnetic material remains constant, resulting in cooling as given by

$$\frac{(B_i^2 + b^2)^{1/2}}{T_i} = \frac{(B_f^2 + b^2)^{1/2}}{T_f} \quad (5)$$

where

$B_i$  = Initial magnetic field

$B_f$  = Final magnetic field

$b$  = Internal magnetic field associated with each magnetic ion

$T_i$  = Initial temperature

$T_f$  = Final temperature

The third stage can be effective in two ways, the first and most thermally efficient is isothermal demagnetization. This provides stability at  $T_f$  by reducing the magnetic field  $B$  from  $B_f$  to zero at a rate which counteracts the thermal input from the surrounding environment. The total amount of energy the paramagnetic material can be absorbed is given by Eq. (3) and Eq. (6) next, since  $T$  is constant,

$$Q = T_2[S_C - S_D] \quad (6)$$

The duration in seconds of operation at  $T_f$ , called the hold time, is given by

$$\text{Hold Time} = \frac{n T_1 [S_C - S_D]}{dQ_{th}/dt} \quad (7)$$

where

$n$  = Number of moles of magnetic ion

$dQ_{th}/dt$  = Total power into the paramagnetic material (e.g., parasitic heating)

Once the magnetic field reaches zero the whole process must be repeated (recycled) from stage 1. The second approach is for the magnetic field to be reduced to zero, completely demagnetizing the paramagnetic salt to a temperature of  $T_{min}$ . A constant temperature above  $T_{min}$  can be achieved by using a stage which is heated via a resistor and which has a weak thermal link to the paramagnetic material. The heater power needs to be reduced as the paramagnetic material warms under the parasitic load. Since the temperature of the paramagnetic material is not constant the energy which can be absorbed is given by Eq. (3) only. This process is less thermodynamically efficient to the isothermal process since heat is being added, however, it does not require active control of the magnetic field and therefore is simpler in some respects.

## TEMPERATURE REGULATION

As previously stated the third stage of operation of an ADR, namely the holding of the final required temperature, can be achieved via demagnetizing at a rate to counteract the heat flowing into and thus warming up the salt pill. Temperature regulation is achieved via a computer controlled servo system in which the rate of reduction in magnetic field is controlled in order to maintain the temperature at the desired value. For superconducting magnets the magnetic field is proportional to current and thus it is the current that is controlled. The degree of stability is limited to the sensitivity of the thermometry and the step size in magnetic field/current. Temperature regulation of  $\pm 1 \mu\text{K}$  is possible with room temperature resistance bridge phase sensitive detection (PSD) electronics and germanium thermometers. The thermometry readout is limited by the Johnson noise associated with the resistance bridge resistors. Cooling of these resistors to approximately 10 K should make it possible to achieve ~a few hundred nanokelvin stability of an ADR when operated at a temperature of 0.1 K.

## Magnetic Field Servo

Constant temperature is maintained by the reduction of electrical current ( $I$ ) by a value  $dI$  every  $dt$  seconds. The value of  $dI/dt$  is determined by the servo algorithm based on how well the temperature is being maintained. With every new value of  $dI/dt$  the current is stepped down at the corresponding value of  $dI$ , ( $dt$  is kept constant) until the new value of  $dI/dt$  is calculated. Intervals of a minute or less are appropriate for most systems depending upon the application.

The current-temperature control equation (16) is:

$$dI_i/dt = (c/\Delta t)[(T_i - T_{set}) + (\Delta t/\tau) \sum (T_{i-j} - T_{set})]$$

where

$i$  = Current value

$j$  = previous ( $i-1$ ) value

$dI_i/dt$  =  $i$ th current ramp rate

- $T_i$  =  $i$ th temperature  
 $T_{i-j}$  =  $i$ - $j$ th temperature  
 $T_{set}$  = Servo temperature  
 $\Delta t$  = time interval  
 $\tau$  = system time constant  
 $\Sigma$  = Sum over  $j=0$  to  $i-1$   
 $c$  = system constant

## SUMMARY

Millikelvin refrigeration is becoming necessary for many applications (astronomy, particle physics, material science, and biophysics). Adiabatic demagnetization refrigerators are seen as one of the most attractive ways of achieving such temperatures. Their compactness, ease of operation (turn key is a possibility since the process is purely electrical and cryogen free) and gravity independence gives many advantages over helium based apparatus. Cryogenic engineering advances now mean that an ADR can be a small bench top instrument rather than requiring a well equipped cryogenic laboratory. With this simplification and miniaturization of cryogenic instrumentation the user will no longer have to be an experienced cryogenic physicist which will open millikelvin refrigeration to a much broader community of scientists and engineers. Such an expansion process has not yet occurred with nuclear demagnetization refrigeration, however, with the constant development of lower operating temperature detectors and advances in cryogenic science and engineering it is probably only a matter of time.

## BIBLIOGRAPHY

1. J. A. Barclay, *Advances in Cryogenic Engineering*, Vol. 33, New York: Plenum Press, 1988, p. 719.
2. G. K. White, *Experimental Techniques in Low-Temperature Physics*, Oxford, UK: Oxford University Press, 1989.
3. C. Hagmann and P. L. Richards, Two-stage magnetic refrigerator for astronomical applications with reservoir temperatures above 4 K, *Cryogenics*, **34**: 213–226, 1994.
4. I. D. Hepburn et al., Submillimeter and Far-Infrared Space Instrumentation, *Proc. 30th ESLAB*, ESA SP-388, 1996.
5. O. V. Lounasmaa, *Experimental Principles and Methods below 1K*, New York: Academic Press, 1974.
6. P. Debye, *Ann. Phys.*, **81**: 1154, 1926.
7. W. F. Giaque, A thermodynamic treatment of certain magnetic effects. A proposed method of producing temperatures considerably below  $1^\circ$  absolute, *J. Am. Chem. Soc.*, **49**: 1864, 1927.
8. W. J. De Haas, E. C. Wiersma, and H. A. Kramers, *Physica*, **1**: 1, 1933.
9. W. F. Giaque and D. P. MacDougall, Attainment of temperatures below  $1^\circ$  absolute by demagnetization of  $Gd_2(SO_4)_2 \cdot 8H_2O$ , *Phys. Rev.*, **43**: 768, 1933.
10. N. Kurti and F. E. Simon, Production of very low temperatures by the magnetic method: supraconductivity of cadmium, *Nature (London)*, **133**: 907, 1934.
11. R. P. Hudson, *Principles and Applications of Magnetic Cooling*, Amsterdam: North-Holland, series in Low Temperature Physics Vol. 2, 1972.
12. S. F. Kral and J. A. Barclay, *Applications of Cryogenic Technology*, vol. 10, edited by J. P. Kelly, New York: Plenum Press, 1991.
13. D. McCammon et al., A sounding rocket payload for X-ray astron-

omy employing high resolution microcalorimeters, *Nucl. Instrum. Meth. A*, **370**: 266–268, 1996.

14. C. K. Stahle et al., Microcalorimetry arrays for high resolution soft X-ray spectroscopy, *Nucl. Instrum. Meth. A*, **370**: 173–176, 1996.
15. M. D. Kuz'min and A. M. Tishin, Magnetic refrigerants for the 4.2K–20K region: garnets or perovskites, *J. Phys. D: Appl. Phys.*, **24**: 2039–2044, 1991.
16. G. Bernstein et al., Automated temperature regulation system for adiabatic demagnetization refrigerators, *Cryogenics*, **31**: 99–101, 1991.

I. D. HEPBURN  
 University College London  
 A. SMITH  
 University College London

## MAGNETIC RESONANCE

Consider a toy top on a table. If it is not rotating, it will immediately fall down because of gravity. If it is spun, however, it will rotate about the  $z$  direction keeping the angle  $\theta$  constant as shown in Fig. 1. Now let a rod-shaped magnet (a nail with a small flywheel) in a magnetic field between the poles be the long axis  $R$  at an angle  $\theta$  from the direction  $z$  of the field, as shown in Fig. 2(a). It will tend to align in parallel like a magnetic compass with the line of magnetic force. If one spins it about the long axis  $R$ , however, it will rotate about  $z$ , keeping  $\theta$  constant as shown in Fig. 2(b). Both of these motions are called precession and are perpetual in the absence of friction. In this case, the toy top or the rod-shaped magnet has angular momentum, and the direction of force is perpendicular to the plane defined by  $R$  and  $z$ . This force is called torque. The precession occurs also on a microscopic scale. Consider a single free electron or a single proton in a uniform field  $H_0$  along the  $z$  axis. It has a magnetic dipole moment  $\mu_e$  or  $\mu_N$ , respectively, which is considered to be a tiny magnet. In this case, one does not need to spin it because it has an angular momentum a priori, and the magnetic moment is caused by the rotation of the charged particle. Because of this angular momentum, it exhibits similar motion

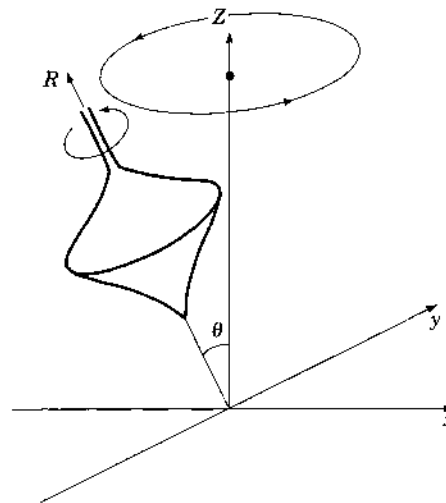
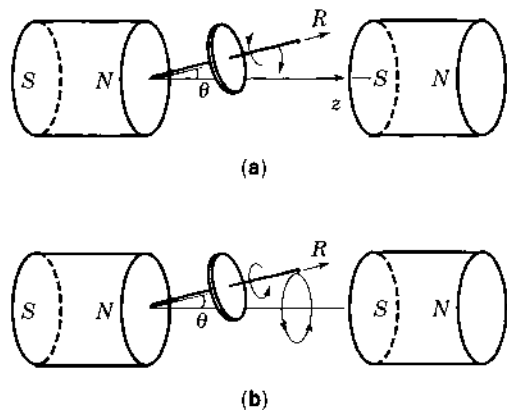


Figure 1. Precession of a top.



**Figure 2.** Small iron rod (nail) in magnetic field. To avoid the gravitational effect, the long axis is along the horizontal direction.

in a magnetic field as earlier and as shown in Fig. 3. This motion is called Larmor precession. The frequency of the Larmor precession is derived from the following torque equation, which is similar to the case of a top:

$$\frac{d\mu}{dt} = \gamma[\mu \times H] \tag{1}$$

where  $\gamma$  is a coefficient called the gyromagnetic ratio and  $\mu$  is either  $\mu_e$  or  $\mu_N$ .  $H$  is the magnetic field described as  $H = (0, 0, H_0)$ . Equation (1) can be solved easily. The  $z$  component of Eq. (1) is

$$\frac{d\mu_z}{dt} = 0 \tag{2}$$

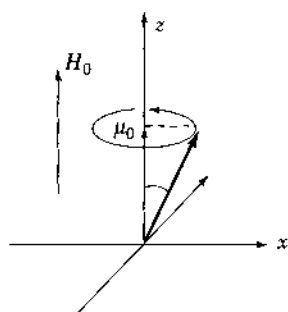
Then one finds that  $\mu_z$ , which is the  $z$  component of  $\mu$ , is constant. The magnitude of  $\mu_z$  depends on the initial condition and is unknown here.  $\mu$  is also unknown but is given by  $\cos \theta = \mu_z/|\mu|$ . This is reasonable and easily understood by Fig. 3. The  $x$  and  $y$  components of Eq. (1) are described as

$$\frac{d\mu_x}{dt} = \gamma\mu_y H_0, \quad \frac{d\mu_y}{dt} = -\gamma\mu_x H_0 \tag{3}$$

From these equations, one can obtain

$$\mu_x + i\mu_y \propto \exp(-i\gamma H_0 t) = \cos(\gamma H_0 t) - i \sin(\gamma H_0 t) \tag{4}$$

This means that the magnetic moment rotates about the  $z$  axis in the direction of a right-handed screw for  $\gamma < 0$  (in the case of an electron) or the opposite way for  $\gamma > 0$  (in the case of a



**Figure 3.** Precession of a magnetic moment  $\mu$ .

proton) with an angular frequency

$$\omega_0 = |\gamma|H_0 \tag{5}$$

where  $\omega_0$  is called the Larmor frequency. As will be explained later,  $\gamma$  is expressed as

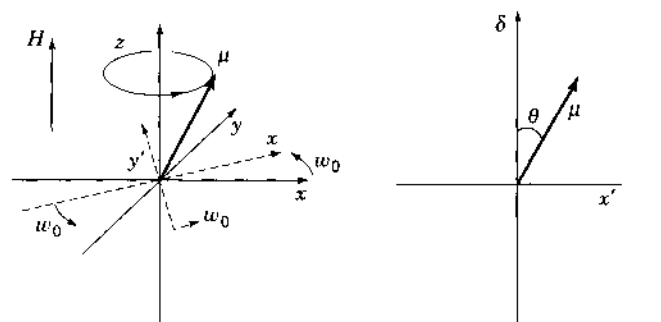
$$\gamma = g \frac{\mu_0 e}{2m} \tag{6}$$

where  $g$  is the so-called  $g$ -value and equal to  $g_e = 2.0023$  for an electron and  $g_N = 2.7896$  for a proton.  $e$  and  $m$  are the charge, which is negative for electrons and positive for protons, and the mass of the particle, respectively.  $\mu_0$  is the permeability of vacuum. The frequencies are easily calculated using  $\omega = 2\pi f$  and are approximately

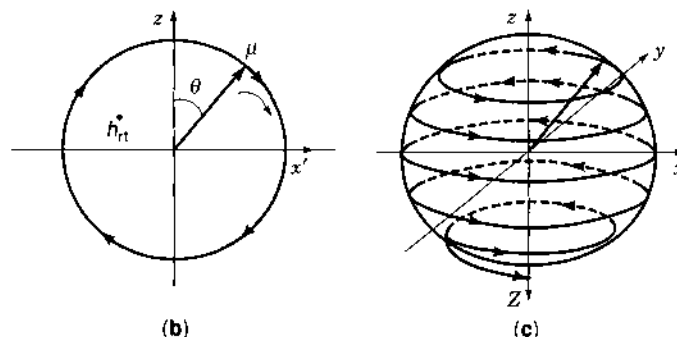
$$f_0 = 28.0 \text{ GHz} \quad \text{and} \quad f_0 = 42.6 \text{ MHz} \quad \text{at} \quad B_0 = \mu_0 H_0 = 1 \text{ T} \tag{7}$$

for moments of electrons and protons, respectively. Notice that the mass of the nucleus depends on the atom and that the  $g$  of the other nucleus is different from that of the proton.

When an ac magnetic field (electromagnetic wave at radio frequency) perpendicular to the  $z$ -axis with amplitude  $h_{rf}$  whose frequency and polarization satisfy Eqs. (4) and (5) is applied, what is its effect on the magnetic moment? First, consider a magnetic moment without an ac field in a coordinate system  $x'y'z$  rotating about the  $z$  axis with  $\omega_0$  as shown in Fig. 4(a). The magnetic moment points to the fixed direction in the  $x'z$



(a)



(b)

(c)

**Figure 4.** (a) Precession in the rotating coordinates  $x'y'z$  without  $h_{rf}$ , (b) rotation of  $\mu$  in the  $x'z$ -plane with  $h_{rf}$ , and (c) trajectory of the top of  $\mu$ .

plane keeping the angle  $\theta$  constant and never moves in this coordinate system. When an ac field is applied,  $\mathbf{H}$  in Eq. (1) must be replaced by  $\mathbf{H} = (h_{rf} \cos \omega_0 t, h_{rf} \sin \omega_0 t, H_0)$ . The magnetic moment starts to rotate about the  $y'$  axis in the  $x'z$  plane as shown in Fig. 4(b) in the rotating coordinate system.  $\theta$  is no longer constant in this case and changes as a function of time as  $\theta = 2\pi f_r t + \theta_0$ , where  $\theta_0$  is the initial angle and  $f_r$  is the repetition rate given by

$$2\pi f_r = \gamma h_{rf} \quad (8)$$

In a fixed coordinate system this means that first the precession is accelerated absorbing the electromagnetic wave power with increasing amplitude of  $\mu_x$  and  $\mu_y$  and with decreasing  $\mu_z$ , and then with decreasing amplitude of  $\mu_x$  and  $\mu_y$  and with increasing  $\mu_z$  toward the  $-z$  direction as shown in Fig. 4(c) as the trajectory of the top of the moment. As soon as  $\mu$  completely points in the  $-z$  direction, it returns until  $\mu$  points in the  $z$  direction emitting electromagnetic wave power this time. Neither absorption nor emission occurs on average. Usually  $f_r$  is much smaller than  $f_0$ , so the real precession is not like that in Fig. 4(c), (i.e.,  $f_0$  is  $10^4$  times greater than 1 cycle of  $f_r$ ). This phenomenon is called magnetic resonance in general and more directly electron spin resonance (ESR) or nuclear magnetic resonance (NMR) depending on which moment is in question. Magnetic resonance occurs when the frequency of the ac field coincides with the Larmor frequency.

As mentioned earlier, the magnetic moment of an electron is caused by the angular momentum of the electron. If one calculates  $\mu$  classically, assuming that an electron is a uniformly charged sphere with radius  $r$  rotating with angular frequency  $\omega$ , then  $\mu$  is obtained as

$$\mu = -\frac{\mu_0 e \omega r^2}{2} = -\frac{\mu_0 e}{2m_e} L \quad (9)$$

where  $e$  and  $m_e$  are the charge and mass of an electron, respectively, and  $L = m_e \omega r^2$  is the angular momentum of an electron. We now must introduce quantum mechanics. According to this theory,  $L$  must be given by  $L = s\hbar$  using Planck's constant  $\hbar$ , where  $s$  is called the spin angular momentum quantum number or simply the spin and  $s = 1/2$ . This means an angular momentum of an electron can be only  $\hbar/2$  or  $-\hbar/2$ . From the theory of quantum mechanical electrodynamics, however, it has been shown that  $g_e$  must be multiplied by Eq. (9) for the magnetic moment of an electron. Then the relation between the magnetic moment of electron  $\mu_e$  and spin  $s$  is

$$\mu_e = g_e \mu_B s \quad (10)$$

where

$$\mu_B = -\frac{\mu_0 e \hbar}{2m_e} = 1.165 \times 10^{-29} [\text{Wb} \cdot \text{m}] \quad (11)$$

is the unit of the magnetic moment of an electron and is called the Bohr magneton. The energy of the magnetic moment in a field is  $-\mu_e H$ , which is called Zeeman energy. Then the Hamiltonian described by the energy in quantum mechanics is written as

$$H = -g_e \mu_B H s \quad (12)$$

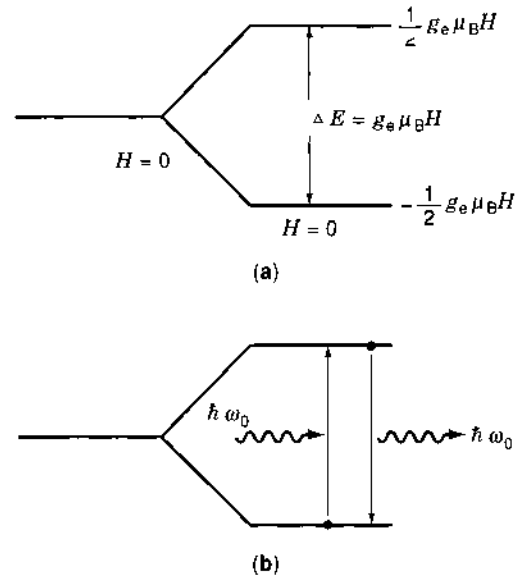


Figure 5. Schematic energy levels of a spin in magnetic field.

Because of  $s = 1/2$  for an electron, only eigenstates of  $\pm 1/2$  are allowed. The energy states are then shown in Fig. 5(a), and the energy difference between these two levels is

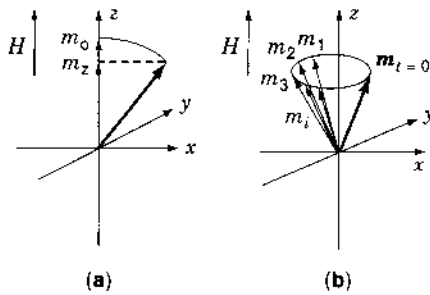
$$\Delta E = g_e \mu_B H \quad (13)$$

In an electromagnetic wave whose frequency satisfies  $\hbar\omega = \Delta E$ , the electron at the ground state is excited to the upper state absorbing the electromagnetic wave energy as shown in Fig. 5(b) and then immediately comes back to the ground state emitting an electromagnetic wave as shown in Fig. 5(b) because the transition probabilities of both transitions are identical. No energy dissipation occurs in this model. This phenomenon corresponds to that explained in Fig. 4. Equation (5) is also obtained from this argument. In the case of a proton,  $g_e$  and  $\mu_B$  must be replaced by  $g_N = 2.7896$  and

$$\mu_N = \frac{\mu_0 e \hbar}{2m_p} \quad (14)$$

where  $m_p$  is the mass of proton. In this case,  $\gamma_N = -g_N \mu_N / \hbar$  is positive because the charge of the proton is positive. In the case of an electron in an atom, it has an orbital motion around the nucleus, and it also contributes to the magnetic moment. But it is not mentioned here for simplicity.

In real materials, the magnetic resonance phenomenon is more complicated because an electron or a nucleus is located in an atom or a molecule and is no longer free. In this case, we deal with the magnetic moment  $\mathbf{m}$ , which is the average or sum of all moments on lattices in the material. Each magnetic moment in a material interacts with other moments on other lattices and with lattice vibrations (phonons). These interactions cause the relaxation phenomenon, which is another important aspect of magnetic resonance. Because of the relaxation, energy dissipation occurs in the resonance condition, and we can observe the magnetic resonance as the absorption of applied electromagnetic wave power. If a magnetic moment undergoes friction during precession, the amplitude of oscillation is supposed to damp and finally the moment becomes parallel to the



**Figure 6.** (a)  $z$  component of  $\mathbf{m}$  increases up to  $m_0$  in a time scale of  $T_1$ . (b) All moments start simultaneously at  $t = 0$ , but they diffuse in a time scale of  $T_2$ .  $m_1, m_2, \dots, m_i$  are each moment.

$z$  axis as shown in Fig. 6(a) instead of as shown in Fig. 3. In this case,  $m_z$  increases as a function of time  $t$  and reaches full length  $m_0 = |m|$  finally. Assuming that the time rate is constant and defined as  $T_1$ , one can rewrite the equation of motion given by Eq. (2) as

$$\frac{dm_z}{dt} = -\frac{m_0 - m_z}{T_1} \quad (15)$$

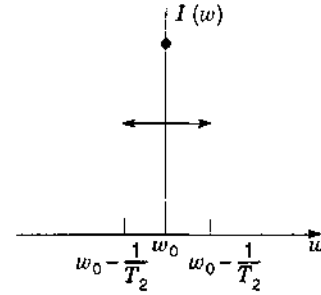
and obtain  $m_0 - m_z = \Delta m \exp(-t/T_1)$ , where  $\Delta m$  is the initial difference. Because this damping comes from the interaction between spin motion and lattice vibration and then corresponds to direct dissipation of energy to the lattice,  $T_1$  is called as "spin-lattice relaxation time" or simply pronounced "tee-one." On the other hand, the  $x, y$  components of the averaged magnetic moment  $m_x$  and  $m_y$  have a finite magnitude when each magnetic moment starts to rotate about the  $z$  axis simultaneously at  $t = 0$ . However, because of interactions between magnetic moments (mainly magnetic dipole interaction), the local magnetic field acting on each magnetic moment and consequently the Larmor frequency varies from site to site in the material. Then the phase of precession of each magnetic moment randomly distributes, as shown in Fig. 6(b). This effect results in decay of the  $x, y$  components of the averaged magnetic moment  $m_x$  and  $m_y$ , which finally becomes 0. The characteristic time of this decay is defined as  $T_2$ , and this is considered to be faster than  $T_1$  because  $T_2$  also includes the energy dissipation effect in addition to the dephasing effect of precession. This effect modifies Eq. (3) as

$$\frac{dm_x}{dt} = \gamma m_y H_0 - \frac{m_x}{T_2}, \quad \frac{dm_y}{dt} = \gamma m_x H_0 - \frac{m_y}{T_2} \quad (16)$$

Equations (15) and (16) are called the Bloch equations. From these equations, one can easily obtain

$$m_x + im_y \propto \exp\left(-i\omega_0 t - \frac{t}{T_2}\right) \\ = [\cos(\omega_0 t) - i \sin(\omega_0 t)] \exp\left(-\frac{t}{T_2}\right) \quad (17)$$

where  $\omega_0 = \gamma H_0$  is used.  $T_2$  is called the spin-spin relaxation time or simply "tee-two." This is the same as a damping oscillation. The general theory, which deals with the relaxation phenomenon more exactly, is difficult and more complicated. Because of these relaxation phenomena, the magnetic moment



**Figure 7.** Spectrum of damping oscillation is Lorentzian.

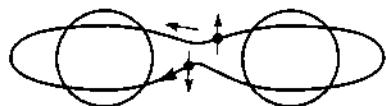
finally points to the  $z$  direction even if it starts to precess from some angle  $\theta$ , and the magnitude of  $m$  is not kept constant during the precession because of the difference of  $T_1$  and  $T_2$ . When an ac field  $h_{rf} e^{i\omega t}$  is applied perpendicularly to the  $z$  direction, the averaged magnetic moment  $\mathbf{m}$ , which first points into the  $z$  direction starts to rotate at  $\omega = \gamma H_0$  and  $\theta$  increases. If  $(\gamma h_{rf} T_1 T_2)^2 \ll 1$  (usually this condition is valid for ESR), not like Fig. 4(c), the Larmor precession becomes stationary, and the tilting angle  $\theta$  is small. If  $(\gamma h_{rf} T_1 T_2)^2 \gg 1$  (sometime this condition is valid for NMR), the Larmor precession is similar to Fig. 4(c). It is, however, complicated to solve the equation of motion precisely. It should be emphasized in this case that applied power of the ac field is absorbed at resonance condition by the spin system, and the absorbed power transmitted as lattice vibrations of the material via the relaxation mechanisms. This process results in temperature increase of the material. Resonance occurs not only exactly at  $\omega = \gamma H_0$  but also at frequencies near  $\omega = \gamma H_0$ . The distribution of the resonant frequency, namely response intensity or spectrum of the characteristic oscillation as a function of  $\omega$ , is obtained by the Fourier transformation of Eq. (17) as

$$I(\omega) \propto \frac{1}{(\omega - \omega_0)^2 + \left(\frac{1}{T_2}\right)^2} \quad (18)$$

and it shows a Lorentzian line shape with a half width of  $\Delta\omega = 1/T_2$  (or full half width  $2/T_2$ ) as shown in Fig. 7. By sweeping frequency of the electromagnetic wave and by observing the power dissipation in the material as a function of frequency, one can see that magnetic resonance with a line shape of Fig. 7 occurs. The experimental method will be discussed later.

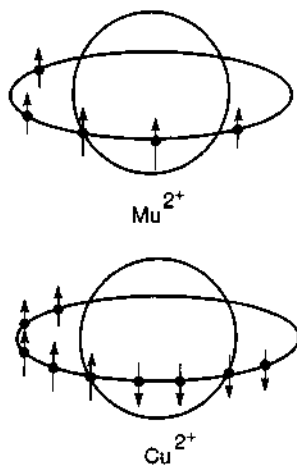
## ELECTRON PARAMAGNETIC RESONANCE

To observe electron spin resonance, there must be isolated and independent electrons in the system. Namely, the materials we are considering must be composed of atoms or molecules that have magnetic moments. Every atom or molecule, however, does not necessarily have a magnetic moment. Electrons in an atom or in a molecule strongly couple with each other, and usually spin and orbital angular momenta are compensated by making pairs of electrons. For example, a hydrogen atom has only one electron, but it becomes a molecule when coupling with another hydrogen atom, and the angular moments of two electrons in a molecule are directly opposite as shown in Fig. 8. Then the molecule has no magnetic moment.



**Figure 8.** Electrons in hydrogen molecule moving opposite in an orbit with magnetic moments caused by spins, as indicated by arrows.

The helium atom has two electrons whose spin and orbital angular moments are also compensated. In this manner, the net magnetic moment for most atoms and molecules disappears as a result of pairing by the strong intraatomic or intramolecular electron correlation that cancels spins and orbits. For example, in the case of iron group atoms (i.e., Sc, Ti, V, Cr, Mn, Fe, Co, Ni, and Cu in the periodic table), spin and orbital angular moments of electrons are not compensated because of Hund's rule, which was a result of the quantum mechanics. This means that these atoms have an unpaired electron and have magnetic moments. Figure 9 shows the example of  $Mn^{2+}$  and  $Cu^{2+}$ . For compounds that have these atoms as divalent or trivalent ions, the magnetic moments are isolated, and ESR can be observed. These ions are called paramagnetic ions and the ESR concerning these ions is called electron paramagnetic resonance (EPR). The atoms belonging to the palladium group, platinum group, and rare earth group also have similar properties. In these ions, the orbital motion of electrons caused by orbital angular momentum creates a magnetic field acting on the magnetic moment caused by spin angular momentum or vice versa. So the orbital angular momentum and spin angular momentum are not completely independent. This effect is called spin-orbit coupling, and this interaction energy is of the order of  $10^{-21}$  J. A paramagnetic ion in compounds is usually surrounded by negative ions called anions, which make a strong electric field called a crystalline field on the paramagnetic ion at the center. The energy of the crystalline field for an electron is on the order of  $10^{-19}$  J. The electrons in the paramagnetic ion suffer electric fields from both the central nucleus and surrounding anions, and the motion of the electrons are no longer simple orbital motions. This effect gives rise to a reduction of orbital angular momentum, which is called quenching. Then the contribution of the orbital angular momentum to the magnetic moment is small and the magnetic moment  $\mu$  per ion is usually



**Figure 9.** Electrons in 3d orbit in  $Mn^{2+}$  and  $Cu^{2+}$  ions.

expressed as

$$\mu = g\mu_B S \quad (19)$$

where  $S$  is the total spin. This is, for example,  $S = 5/2$  or  $S = 1/2$  for our  $Mn^{2+}$  or  $Cu^{2+}$  ion, respectively, as easily is understood by Fig. 9.  $g$  is the so-called  $g$ -value, which reflects the effects of spin-orbit coupling and the crystalline field and is different from  $g_e$  depending on the material.

Paramagnetic ions in a crystal interact with each other by dipole interaction and exchange interaction. These interactions give rise to dephasing of Larmor precession, as mentioned previously, and result in the width of the absorption line as  $1/T_2$ . If the dipole interaction is dominant, the half width is approximately given by

$$\Delta\omega = 1/T_2 \cong \omega_d \quad (20)$$

where  $\omega_d$  is the sum of dipole interaction divided by  $\hbar$ . The sum is over all magnetic moments on the crystal lattice. If the exchange interaction is larger than the dipole interaction, the half width is approximately given by

$$\Delta\omega = 1/T_2 \cong \omega_d^2/\omega_e \quad (21)$$

where  $\omega_e$  is the nearest neighbor exchange interaction divided by  $\hbar$ .

The unpaired electron is also realized in organic materials as the free radical, in semiconductors as the donor or acceptor impurity, and in color centers as some special molecules like  $O_2$  or NO. By studying EPR, one can obtain microscopic information about materials via  $g$ -value and line width.

#### HOW TO OBSERVE EPR (EXPERIMENTAL METHOD)

A simple method to observe EPR is described here. The equipment necessary for this experiment is an oscillator, a detector, a cavity, and a magnet. The Larmor frequency of electron spin is in the microwave region at an easily available magnetic field, namely  $B = 0.1$  to  $1$  T, as discussed, see Eqs. (5) and (7). The most popular way is to use X-band microwaves whose wave length is around 3 cm because the size of microwave components is moderate, and the resonance field is about 0.3 T. A Gunn oscillator with detector is now commercially available and most convenient for this purpose. This oscillator is commonly used for detection of speeding automobiles by policemen. Assemble the equipment as Fig. 10(a). The Gunn oscillator is connected with a short wave guide and terminated by a cavity. First, operate the Gunn oscillator and tune the cavity by moving a plunger so as to resonate at the oscillator frequency. A sample [A small amount of DPPH ( $\alpha, \alpha$ -diphenyl- $\beta$ -picrylhydrazyl)] may be good as a test sample] is put in advance on the bottom of the cavity where the high-frequency magnetic field is strongest. Install the cavity between the poles of the magnet as shown in Fig. 10(b). A low-frequency ac (50 Hz or high) field of a few hundred microteslas must be superposed on the main dc field by modulation coils. Then sweep a magnetic field up to about 0.3 T. In this discussion, the absorption spectrum is given as a function of frequency at constant field. In the real experiment, however, changing frequency is so difficult that usually a magnetic field is swept at constant frequency. The detected

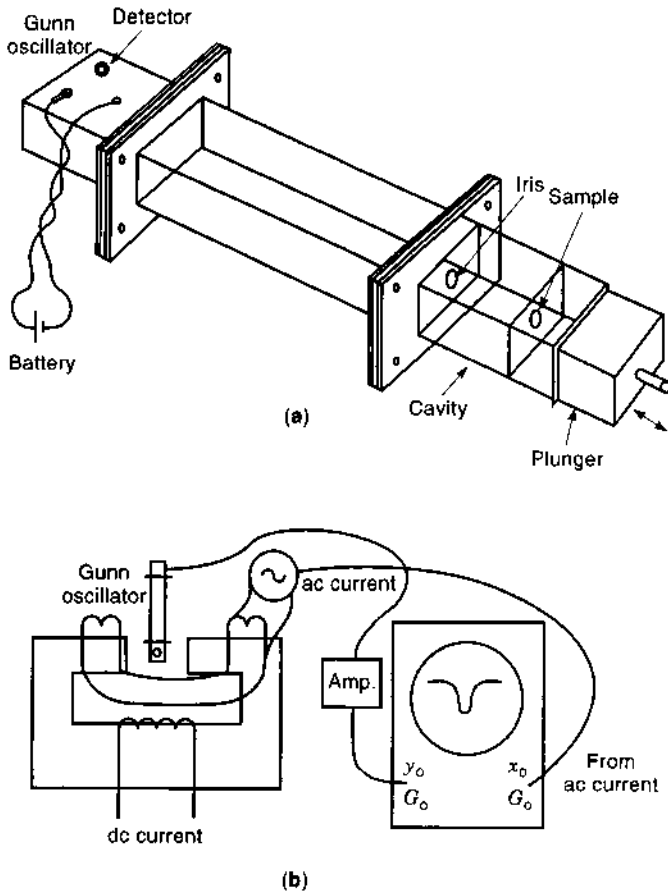


Figure 10. Experimental set-up to observe EPR.

signal through an amplifier is displayed on an oscilloscope as shown in Fig. 10(c). The resonance absorption line is obtained as a function of magnetic field and the half width of the resonance line must be converted by

$$\Delta H = \frac{\Delta\omega}{\gamma} \quad (22)$$

## FERROMAGNETIC RESONANCE

In the case of ferromagnetic materials, all magnetic moments point to same direction below a certain temperature  $T_c$  because of the strong exchange interaction among magnetic moments. The transition temperature  $T_c$  is called the Curie temperature. At a temperature sufficiently below  $T_c$ , the average magnetic moment  $m$  saturates and the total magnetic moment  $M$  of the specimen is given by

$$M = Ng\mu_B S \quad (23)$$

where  $N$  is the magnetic moment per unit volume in the specimen. The magnitude of  $M$  is comparable to the applied flux density  $B_0 = \mu_0 H_0$ , whereas in the case of paramagnetic materials, the averaged moment  $m$  is about  $10^{-3}$  of the saturation moment at room temperature and at 1 T as a result of thermal fluctuation. Then the magnetic moments experience a field produced by themselves pointing in the opposite direction to the

applied field. This field is called the demagnetizing field, and it must be taken into account when the equation of motion is solved. The field acting on the magnetic moment  $M$  including demagnetizing field is given for each direction as

$$H_x = -N_x M_x, \quad H_y = -N_y M_y, \quad H_z = H_0 - N_z M_z \quad (24)$$

where  $N_s$  are the demagnetizing factor and  $N_x + N_y + N_z = 1$  and  $H_0$  is parallel to the  $z$  axis. From the equation of motion

$$\frac{d\mathbf{M}}{dt} = \gamma[\mathbf{M} \times \mathbf{H}] \quad (25)$$

one can obtain the resonance conditions as

$$\frac{\omega}{\gamma} = \sqrt{(H_0 + (N_y - N_z)M_z)(H_0 + (N_x - N_z)M_z)} \quad (26)$$

When the sample shape is spherical,  $N_x = N_y = N_z = 1/3$ , then

$$\frac{\omega}{\gamma} = H_0 \quad (27)$$

In the case of a thin disk,  $N_x = 1, N_y = N_z = 0$  or  $N_x = N_y = 0, N_z = 1$  for an applied field parallel or perpendicular to the disk surface, respectively. In the case of a thin rod,  $N_x = N_y = 1/2, N_z = 0$  or  $N_x = 0, N_y = N_z = 1/2$  for an applied field parallel or perpendicular to the rod, respectively. To observe the absorption of ferromagnetic resonance, there must be energy dissipation caused by the relaxation mechanism. Different from the case of paramagnetic resonance, however, the magnetic moments are tightly bound to each other, and no dephasing effect in the  $xy$  plane is expected. So we cannot define  $T_2$ . Instead, the Landau-Lifshitz damping model is introduced for the ferromagnetic resonance. The equation of motion is described as

$$\frac{d\mathbf{M}}{dt} = \gamma[\mathbf{M} \times \mathbf{H}] - \lambda_L[\mathbf{M} \times [\mathbf{M} \times \mathbf{H}]] \quad (28)$$

where  $\lambda_L$  is the Landau-Lifshitz damping factor. This means that the direction of the second term is perpendicular to the direction of  $\mathbf{M}$  in the plane made by  $\mathbf{M}$  and  $\mathbf{H}_0$ . Then the motion of the total magnetic moment  $\mathbf{M}$  is a damping oscillation as shown in Fig. 11. In this case, the length of  $\mathbf{M}$  is kept constant, whereas in the case of paramagnetic resonance, the relaxation times  $T_1$  and  $T_2$  are independent, and the length of

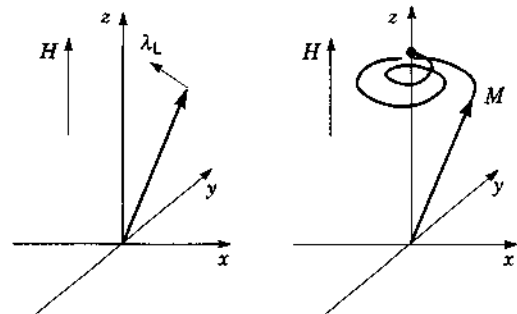


Figure 11. Direction of Landau-Lifshitz term and precession of ferromagnetic moment  $M$ .

the averaged moment  $m$  is not constant during the motion. The magnitude of  $\lambda$  is related to the interaction between the motion of magnetic moments and lattice vibration and then the energy of the Larmor precession is transferred to the lattice vibration via this mechanism. The line width of FMR is expressed using this damping factor as

$$\Delta H = \frac{MH_0}{\gamma} \lambda_L \quad (29)$$

## NUCLEAR MAGNETIC RESONANCE

In the case of nuclear spin, interactions with surrounding electrons, lattice vibration (phonon), other nuclear spins are weak, and a nuclear spin is considered to be almost isolated, which means that  $\gamma$  is almost constant and different from the case of EPR. The magnetic interactions with surrounding electrons can be replaced by the effective field  $H_{\text{eff}}$  which is included in Eqs. (1) and (5) as additional fields. Interactions also contribute to the change of relaxation time from that of free nucleus. Observing the shift of resonance frequency caused by this effective field and the change of relaxation time, one can obtain information about the microscopic behavior of materials. This is the reason why NMR is so useful as a probe to investigate properties of materials. Every nucleus does not necessarily have nuclear spin, and  $\gamma$  varies depending on the nucleus. All  $\gamma$  are listed in a standard table. NMR of copper nucleus is useful to investigate high  $T_c$  superconductors composed of copper oxide. Mn and Co nucleus are also important to study magnetic properties of materials composed of these atoms.

The most popular nucleus is the proton, which is the nucleus of hydrogen. All materials containing hydrogen show pro-

ton NMR. Water is the best example. To investigate molecular structures of organic molecules, polymers, proteins, and other biological materials, the proton NMR is useful and is now being used widely. In these cases, the absorption spectra of proton NMR have a complicated structure as a result interactions with neighboring atoms. By analyzing the structure of spectra, one can determine the molecular structure like neighboring atoms and distance. Because the resolution increases with increasing resonance frequency, high-field and high-frequency MNR is more useful, and now frequencies higher than 750 MHz are available in fields above 17 T by using high homogeneous superconducting magnets.

Magnetic resonance imaging (MRI) is well known as an important tool in finding tumors or other abnormal tissues in the human body. Every cell in organs contains hydrogen atoms and NMR is observable in any part of the body. But the shift or relaxation time varies depending on the organ. As shown in Fig. 12, a body is placed between the poles of a big magnet and  $h_{\text{rf}}$  is applied to it. The magnetic field has a gradient with respect to the position of the body, and the NMR is observed at only one point on the body. This gradient field is scanned, and the resonance point moves from head to foot. By analyzing the data by a computer, one can see the structure of the body. If the organ is abnormal, the density and relaxation times of NMR at the affected part are different from those at a normal part. This allows NMR to be used for diagnosis.

## BIBLIOGRAPHY

1. A. Abragam, *The Principle of Nuclear Magnetism*, New York: Oxford, 1961.
2. G. E. Pake, *Paramagnetic Resonance*. New York: W. A. Benjamin, 1962.
3. C. P. Slichter, *Principle of Magnetic Resonance*. New York: Harper & Row, 1963.

MITSUHIRO MOTOKAWA  
Tohoku University

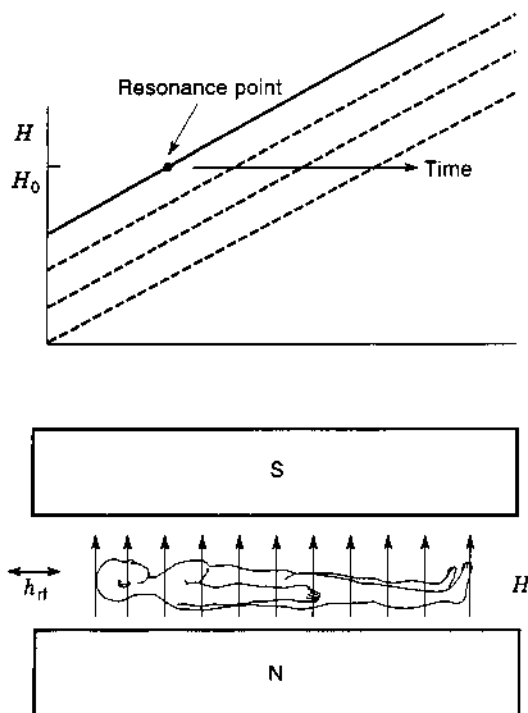


Figure 12. Schematic view of MRI.

## SUPERCONDUCTING MAGNETS, QUENCH PROTECTION

Superconducting magnets are subject to thermal instability, leading to the loss of superconductivity, known as *quench*, in which the critical values of field, temperature, and current density are exceeded and fail to recover. This phenomenon generally begins in a localized region of the coil, then spreads to the rest of the magnet or magnet system with a "quench propagation velocity." All magnets use some form of composite superconducting wire, in which superconducting filaments carry current in parallel with normal conductor, known as the stabilizer. The stabilizer has the dual purpose of preventing quench in the face of disturbances and of protecting the magnet from excessive temperatures and pressures when unwanted quenches occur. Following a quench, current rapidly transfers from the superconducting material to the stabilizer, since composites are always designed so that the resistance of the stabilizer is much less than that of the superconducting material in its normal state.



In the case of a quench, it is almost never possible to allow continued operation at constant current without unacceptable temperature rises in the magnets. The current must be reduced to zero and the stored energy in the magnet eliminated, either by absorbing the energy within the magnet or by dumping the energy externally. If the energy is dumped externally, it may be absorbed in either warm, generally room temperature, dump circuits, or in cold, generally liquid helium temperature, dump circuits. Some magnets have sufficient enthalpy in their stabilizer and magnetically coupled passive structures to absorb their own stored energy without excessive temperature or pressure rises. In fact, most magnets have sufficient energy, if the ratio of peak local heating to average global heating can be held to acceptably low levels. The peak/average ratio can be controlled by design either by activating internal resistive or inductive heaters or by rapidly dumping or heating all helium coolant in order to guarantee that large portions of the coil will heat up together. The other method of protecting magnets is to dump the magnet energy into an external resistor. This requires an absolutely reliable method of interrupting current flow from the power supply and diverting it into the dump resistor. Both methods require reliable and rapid detection of a quench.

The fundamental limit on protecting magnets against quench is the detection of that quench. This can be particularly difficult in the case of coils in an electromagnetically noisy environment, pulsed coils, and multicoil systems. It is also difficult when the coils are very conservatively stabilized, as is often the case for very large coil systems and buswork. Quench detection systems can be active or passive. Active systems usually involve some sort of balanced voltage bridge. The signal/noise ratios of voltage bridges can be improved by using cowound sensors and active cancellation. Passive systems use transformer-fed heaters to trigger superconducting switches or voltage thresholds to trigger cold diodes.

In order to size a magnet for protection, it is usually necessary to know something about the physics of quench propagation. Different physical theories are needed to predict the spread of quench in potted, pool-boiling, and cable-in-conduit superconductors (CICC). Because of the difficulties in predicting disturbances, initial quench zones, and quench propagation, a conservative design criterion is to assume that local hot spots are adiabatic and that all energy must be dumped externally. For internal quenching, the maximum time to heat a long quench zone with a cowound or surface heater has to be known.

The problem of magnet quench protection is a subset of the generic problem of magnet protection, (1) during normal operation, (2) off-normal conditions such as quench, and (3) faults, whether in the coil, bus, or power supply. The two most fatal flaws are mechanical rupture and electrical arcing. They are often preceded by excessive displacement and/or partial electrical discharges and leakage currents. Flaws that are sometime repairable can end a magnet's useful life when cracks cause leakage of helium under pressure, leakage current causes enough heat to quench the magnet, or displacements create unacceptable field errors. All of the structural design may be considered as part of the magnet protection design. This subject is too vast to be treated here, but should be discussed for specific applications in the articles on superconducting magnetic energy storage, motors/generators, fusion, and magnetic resonance imaging magnets. The design for electrical integrity

will be discussed here, since it isn't discussed elsewhere and because the internal and external voltages during a quench dump are usually significantly higher than those during normal operation.

Finally, we briefly review the actual history of failures to protect magnets. Case histories provide a cautionary tale: this article can only go so far in helping to protect magnets, since most failures are caused by mental lapses that are frequently not design errors.

## COIL PROTECTION CIRCUITS

When a superconducting magnet quenches, all of its magnetic energy is converted into heat. If a magnet has enough total mass to absorb the heat and is small enough to guarantee that a quench will propagate into a large fraction of the magnet, then no protection circuits are needed, except to disconnect the power supply, when the current isn't freely circulating. However, when the magnet is too large and stores too much energy to guarantee completely passive protection, some active measure must be taken. The main distinctions between the most commonly used coil protection circuits are whether the dump resistors are internal or external to the cryostat and whether the quench detection is triggered by active logic or passive breakdown of a switch. External dump resistors usually correspond to the design philosophy of saving the magnet by depositing almost all of the energy into a large, inexpensive structure at room temperature. Internal dump resistors are usually designed to lower the peak local to average heating of the magnet to a manageable low level.

### External Dump

Neumeier has recently reviewed the external quench protection circuits for superconducting magnets (1). He schematizes the basic external dump circuit as shown in Fig. 1. The magnet is represented by an inductance  $L$ , while the mutual inductance  $M$ , and the coupled inductance and resistance,  $L_i$  and  $R_i$ , represent the sum of all coupled magnets and passive conducting structures in the magnet system. The external dump circuit consists of a power supply, a closing switch CS, to shunt out the power supply during quench dump, an opening switch OS to interrupt magnet current, and a dump resistor  $R$  ( $\Omega$ ). The basic principle is that the dump resistor is much, much larger than the resistance of the magnet normal zone, so that almost all of the energy is deposited, at room temperature, in a resistor sized to safely absorb all of the magnet energy.

Leads to the dump resistor should generally be coaxial, in order to minimize the voltage overshoot, due to  $L_{\text{leads}} di/dt$ .

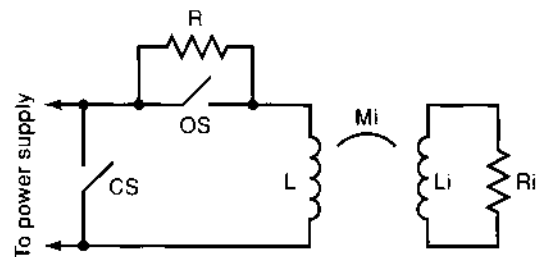


Figure 1. Simple dump circuit schematic [Neumeier et al. (1)].

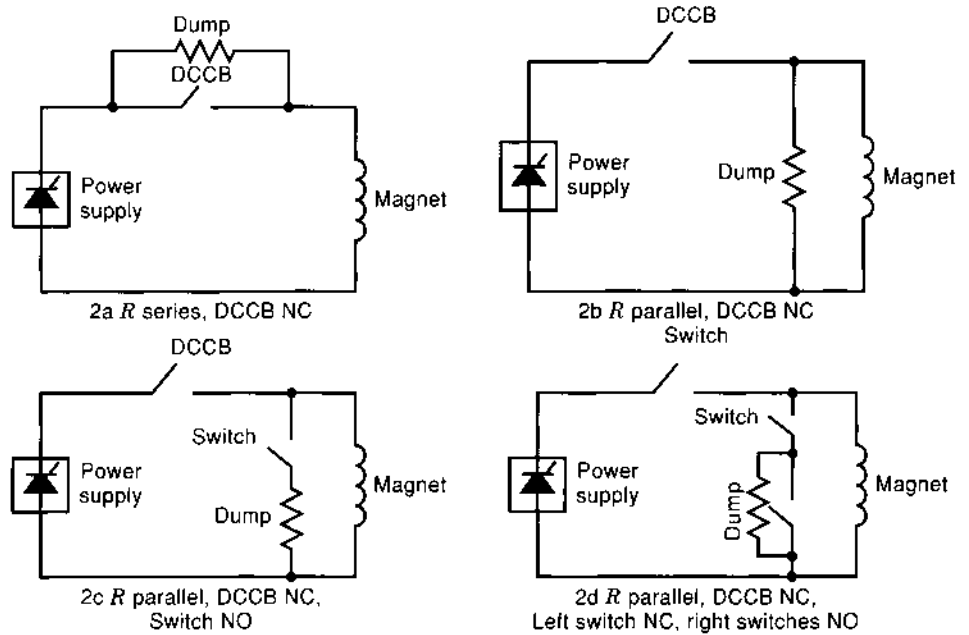


Figure 2. Equivalent superconducting magnet quench dump circuits.

Several equivalent simple dump circuits may be used, as shown in Fig. 2. Under normal operation, they all have the same effect on the magnet. The tradeoffs are in cost versus reliability and in the specifics of the power supply and magnet grounding system. A magnet ground/interrupter switch configuration should be selected that allows the switch and magnet to float on a single short to the magnet case ground without drawing large fault currents.

Neumeyer represents a typical counterpulse interrupter with the circuit schematic shown in Fig. 3.

The operating coil current flows through the power supply (PS) and the normally closed switch. When a quench is detected, the counterpulse capacitor  $C$  is discharged by closing the switch DS, producing a current zero in OS. The inductor SR is a saturable reactor, which desaturates near current zero, decreasing the  $dI/dt$ . This helps to extinguish arcs or to restore voltage-holding capability in a solid-state switch. The key parameter, set by selecting the counterpulse capacitor, is the time during

which current must be reversed and held near zero. This is on the order of  $10 \mu s$  for vacuum bottle interrupters,  $5 \mu s$  to  $50 \mu s$  for thyristors, and  $50 \mu s$  to  $200 \mu s$  for air-blast interrupters. The faster the interrupter clears, the less expensive the counterpulse circuit. For solid-state switches with antiparallel diodes, the energy stored in the commutating capacitor must be

$$\frac{1}{2}CV_c^2 = \frac{1}{2}t_{off} \frac{V(0)I(0)}{\sin \alpha} (\pi - 2\alpha) \quad (1)$$

where  $t_{off}$  is the specified time during which reverse voltage is maintained across the solid-state switch ( $s$ ),  $\alpha$  is the phase angle in radians when the switch current is zero, and  $V(0) = I(0)R$ .

The alternative to counterpulsed circuits are dc switches that can sustain a high enough voltage to force current zero without a resonant pulse. Here, the opening switch develops a high enough arc or resistive voltage to drive the device current close enough to zero to extinguish itself and shunt the current to a dump resistor. The most commonly used forced current zero devices are air-blast breakers and explosive switches. The air-blast breakers use a blast of compressed air and rapidly parting contacts to create a very long, contorted arc with high-voltage and a high tendency to quench. The blast also cools and further constricts the arc. Explosive switches use a redundant number of explosive charges and small arcs with a moderately high voltage in each arc. Both types of switch can also be counterpulsed to increase the probability of current interruption. However, even in a forced-current zero circuit with no capacitor, an inductor is still needed as a snubber, in order to limit the  $dV/dt$  rise across the superconducting magnet, as discussed next.

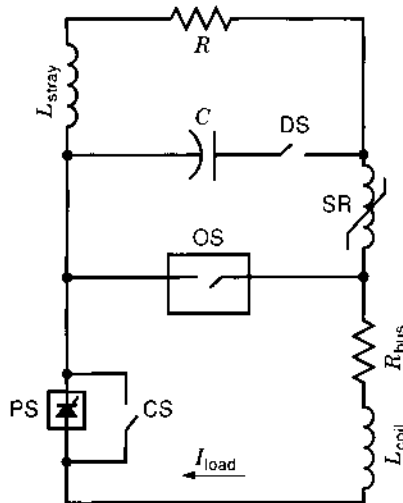


Figure 3. Typical Artificial Zero Counterpulse Circuit [Neumeyer et al. (1)].

**Interrupters.** In order to protect a superconducting magnet with an external dump resistor, an absolutely reliable current interruption switch is needed. Several current interruption switches have been used in magnet design for quench dump or other high voltage pulses. Whatever technology is selected, the interrupters will usually have two opening switches in series

in order to provide adequate reliability. The dump strategy will then be either to open both series switches at once or to detect a failure to open in a nondestructive opening switch (e.g., a solid-state switch), then to open a more reliable but destructive switch (e.g., a series explosive switch). An interrupter may also include a switch to shunt conductor current during normal operations in order to reduce the steady-state ampacity requirement of the main interrupter. In this case, the high-current, inexpensive, normally closed mechanical switch transfers current into the quench dump interrupter after quench detection. The interrupter then carries the magnet current only as long as is needed to open and transfer its current to the external dump resistor.

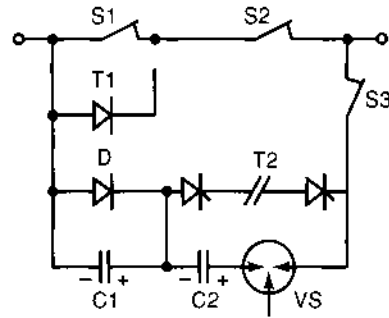
Interrupting switch technologies include:

1. Thyristor breakers with counterpulse circuits
2. Mechanical breaker (air, air blast, vacuum, vacuum and magnetic field)
3. Explosively actuated breaker (fuse and fuseless)
4. Water cooled fuses (activated by water flow interruption)
5. Gate turn off (GTO) thyristor breaker
6. Insulated gate bipolar transistor (IGBT) switches
7. Superconducting switches

Turn-on switches that are used for the counterpulse circuit include:

1. Ignitrons
2. Thyristors
3. Vacuum switches

In the past, mechanical interrupters were favored for large magnets, because of the high power handling capabilities of a single device (e.g., up to 73 kA  $\times$  24 kV in the JET air-blast interrupter (2)). However, with mechanical interrupters, the inevitable electrode erosion by current interruption arcs tends to limit the number of reliable operations to  $\sim 10^4$  operations with periodic maintenance every  $10^3$  interruptions. The probability of failure (to interrupt current) in a mechanical interrupter has also been typically  $10^{-3}$  to  $10^{-4}$  at best, although this is clearly dependent on the specific design. For example, Yokota reported vacuum bottle tests in which there were five interruption failures in 10,000 interruptions with single bottles and no failures in 10,000 with two seriesed bottles (3). Clearly, any degree of reliability can be achieved with mechanical switches with adequate redundancy and maintenance. However, high-reliability requirements are usually met by the use of solid-state devices. Because they have no moving parts and are erosion-free, their lifetimes can easily exceed  $10^6$  operations, limited only by thermal fatigue. While individual solid-state devices used to be limited to the range of 1 kV  $\times$  1 kA, it is now possible to purchase thyristors with ratings of 6 kA  $\times$  6 kV. In the case of thyristor solid-state switches, the reliability of interruption will probably be set by the counterpulse circuit with the failure mechanism being either capacitor burnout or failure to close of the counterpulse circuit switch. The counterpulse circuit and its reliability limitations can be eliminated by the use of GTOs. They have always had less power-switching capability than normal thyristors, but are currently available with ratings of 3.3 kV  $\times$  (4 kA, turn-off, 1.2 kA, ss). A new technology, IGBTs, is beginning to



**Figure 4.** An arc-free current interrupter with pulse-rated solid-state components.

be used in high-power applications requiring fast switching, with device ratings up to 3 kV  $\times$  (1.2 kA, turn off  $\times$  400 A, ss). IGBTs can be switched an order of magnitude faster than conventional thyristors, making them useful in switching converters that reduce the amount of filtering or voltage ripple on the magnets.

Explosive fuses are now capable of operating with very high reliability. They are inexpensive and incorporate redundancy in a single unit by including several in series explosive charges and arcs. They are frequently counterpulsed for further redundancy. Explosive fuses have the particular problem that they won't interrupt currents below a certain level. They are therefore most appropriately used in quasi-steady-state operations, in which low current quench is highly unlikely, or in those magnets that can guarantee passive internal absorption of the quench energy. They also favor applications where quench is highly unlikely, because they can only be used once.

A clever hybrid was proposed by Kuchinski to minimize the total cost of mechanical and solid-state interrupters (4), as shown in Fig. 4. This circuit allows all of the solid-state components to be pulse rated, while eliminating arcing in all of the mechanical components. S1 and S2, low-voltage mechanical switches, carry the magnet operating current. To interrupt, the thyristor T1 fires, suppressing any arcs, while S1 opens. T2 fires to initiate the counterpulse through T1 that turns it off. With D and T2 carrying magnet current with a low voltage drop, S2 opens. Then the vacuum switch is ignited to provide the reverse counterpulse through T2 to switch it off. All switches are interrupted and magnet current flows through the dump resistor.

**Dump Resistors.** The most common dump resistor is a meander of steel bars. Alternative dump resistor concepts include:

| Dump Resistor                      | Advantages   |
|------------------------------------|--|
| Steel bars in air                  | Simplicity, cost, maintainability  |
| Steel bars or ribbon in water (5)  | Simplicity, energy density   |
| Liquid rheostats (6)               | Energy density, elimination of solid structure;  |
| Voltage clamps (Zeners, MOVs, ZnO) | Disadvantage: negative temperature coefficient<br>Faster dump for fixed peak voltage;<br>Disadvantage: High cost/Joule |

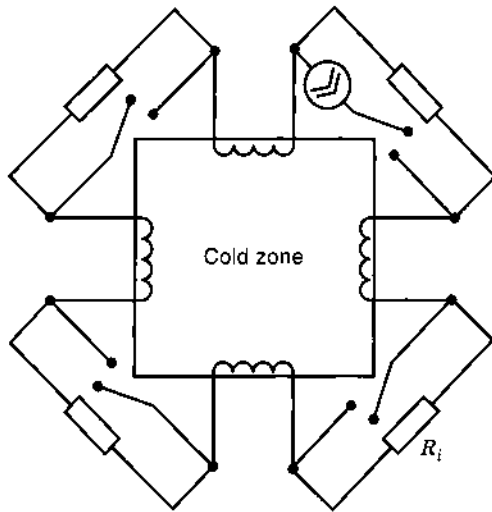


Figure 5. Series interleaf protection circuits [Dudarev et al. (8)].

The voltage across a linear resistor declines with the current. Since electrical integrity is limited by voltage, it would be more efficient to use a resistor that discharged at constant voltage. With a perfect voltage source, either the peak dump voltage could be lowered by one-third or the dump time could be improved one and a half times. This can be approximated by highly nonlinear resistors, such as Zener diodes, Metal Oxide Varistors (MOVs) or Zinc Oxide (ZnO) arresters (7). At very high energy levels, these are prohibitively expensive. An inexpensive alternative with a useful degree of nonlinearity is stainless steel, which has a resistance temperature coefficient of  $\sim 0.001/\text{K}$ . If the temperature of a stainless steel resistor is allowed to rise 500 K by the end of a dump, its resistance will have increased by 50%. With nickel-iron alloys, the resistance can be quadrupled by the end of a pulse (1).

**External Quench of Multiple Magnets.** Both magnets and switches are limited in voltage and current. When a magnet system becomes too large, the dump circuit must subdivide into several parallel, series, or independent circuits. A particularly elegant circuit topology is that of the series interleaf, used in the tokamak systems T-15 and the Tokamak Fusion Test Reactor (TFTR), a normal magnet system. The series interleaf connection is shown in Fig. 5 (8).

This circuit has two advantages over independent or parallel protection circuits. With the interleaf, the voltage drops alternatively up and down, because of the alternation of superconducting inductors and external resistors, as shown in Fig. 6. This prevents high voltage from building up through the coil system, as it would do if there was only a single dump resistor. If the large coil system had simply been broken into the same number of independent or parallel dump circuits, there would be a possibility of unbalanced currents and forces, during a quench dump. With a floating power supply, the interleaf circuit also prevents unbalanced forces during a single ground fault.

Alternative options for dumping energy in a mutually coupled multicoil system are summarized in the table.

| Dump Strategy   | Advantages  | Disadvantages   |
|---|---|---|
| Dump all coils in series  | Simplicity, no unbalanced currents or forces even with single ground fault  | High terminal voltage   |
| Dump all coil in series interleaf   | No unbalanced currents or forces even with single ground fault; low voltage | As many VCLs and dump circuits as interleafs  |
| Dump all coils in parallel; dump all coils independently                      | Low voltage; independent control  | Possibility of unbalanced currents and forces; large number of VCLs; least reliable for given component reliability |
| Dump faulted coil externally; other coils remain in persistent mode (ORNLEPR) | Smallest refrigeration requirement for recool                               | Overcurrents in persistent magnets when coupled to dump coil  |

### Internal Dump

Magnets are dumped internally when it is desirable to eliminate helium loss through the vapor-cooled leads and when terminal voltages have to be severely limited, as with commercial products, such as MRI magnets, where users have to be in the same room as the magnet. The least aggressive action that is typically taken with a pool-boiling magnet is simply to disconnect the superconducting coil from the power-supply and allow its current to die down slowly as it freewheels through a superconducting switch or normal joint. This would be an appropriate response to a low liquid-level reading, in which an active dump might do more harm than good.

Cold dump resistors may still be either internal or external to the magnet. That is, a large fraction of the energy can be deposited into resistors or diodes that are outside of the magnet, but inside the cryostat. However, since there is now no benefit in refrigeration or cooldown requirements in removing energy from the magnet, it is almost always desirable to return as much of the heat as possible to the magnet, thereby creating longer quench zones and more uniform heat deposition with the magnet. This can be done by making the resistive element a heater, closely coupled to the outer layer of the magnet or cowound with the magnet superconductor.

**Cold Dump Circuits.** In a cold dump circuit, either the interrupter or the resistor or both will be inside the cryostat at the magnet temperature. Cold switches include superconducting switches and fuses, cold diodes, and cold transistors. Cold resistors include cowound, insulated normal metal, surface

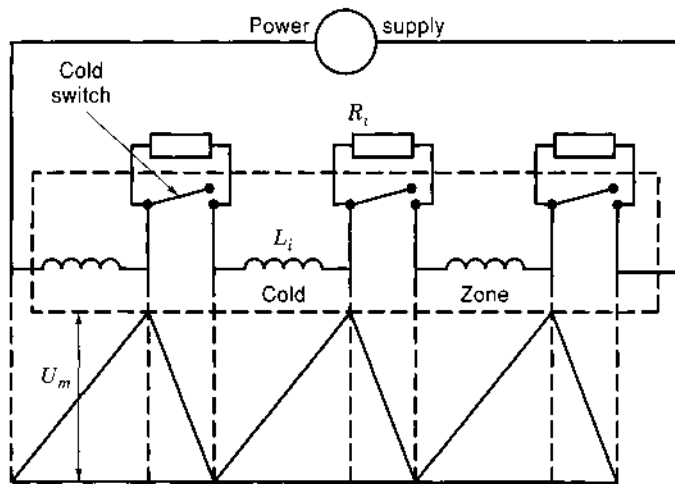


Figure 6. Series interleaf circuit dump waveform [Dudarev et al. (8)].

heaters, and power dissipation in the switch itself. One benefit of having all elements of the dump circuit cold, when the switch is superconducting, is that the vapor-cooled leads can be detached, allowing current to circulate losslessly in the magnet. Cold heaters have two additional benefits, especially when they are cowound through the entire magnet. They can prevent hot spots by reducing the peak local/average heating, and they can also cancel the  $dI/dt$  voltage of the magnet with the resistive voltage, greatly reducing internal voltages in the coil. Both the thermal peaking factor and the internal voltage can be reduced by orders of magnitude from the external dump option in a well-designed internal dump circuit.

The disadvantage of a cold resistor is that all of the magnet energy is absorbed at cryogenic temperature, greatly increasing the time and cost for recool. Therefore, strings of accelerator magnets, which involve a very large number of magnets and training (see SUPERCONDUCTORS, STABILITY IN FORCED FLOW) quenches, use cold switches, but dump externally in order to achieve a large number of rapid cooldowns. CICC magnets and absolutely stable pool-boiling magnets are not supposed to quench. If they do, the engineering postmortem would generally take more time than cooldown, and recool is not a major consideration. The disadvantage of a cold switch is that the power-handling capability of an individual switch is much smaller than that of a warm switch. Above approximately  $200 \text{ V} \times 15 \text{ kA}$ , cold-switching becomes impractical.

Cold switches can be triggered actively or passively. Cold diodes begin to conduct when a forward voltage higher than a threshold is applied. Cold transistors begin to conduct when a forward voltage higher than the turn-on voltage is applied to their base. The transistor gain allows them to be over an order of magnitude more sensitive than the diodes. Superconducting switches can be driven normal by external heaters or induction coils, coupled closely with the switch. These triggers may either be driven from a small, external power supply, or be driven passively by the magnetic induction of the quench itself.

A fully actively driven resistor doesn't actually require any switch. Usually, the entire magnet can be driven normal, by applying heat that is less than 1-2% of the magnet stored energy.

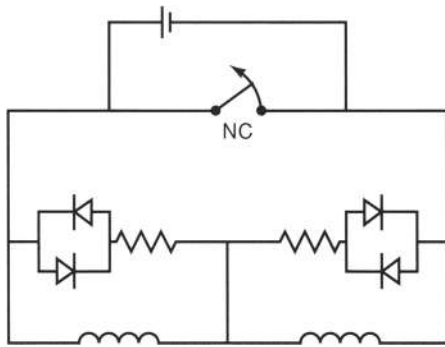
A small external circuit can drive the current in a cowound resistor or a heating pad on the outer layer in order to ensure a large quench zone. A totally nonresistive method for creating a long quench in a large coil has also been proposed, in which helium would be rapidly drained from the cryostat (9). Advantages and disadvantages of each method are listed in the table.

| Cold Dump Circuit                         | Advantages  | Disadvantages  |
|---|---|--|
| Metal strips                              | Low \$/kg; uniform heating easy   | Extremely massive, bulky   |
| Superconducting switches                  | Ideal for low energy coils with persistent currents   | Much too expensive for large coils; both massive and high \$/kg  |
| Superconducting fuses                     | Can develop higher voltage, faster dump than superconducting switches   | Much too expensive for large coils   |
| Internal resistive heaters                | Guarantee uniform heating of coil, without imposing high quench propagation velocities. High degree of redundancy can be designed in. No bulky dump structures outside coil | Extraction of leads from CICC must be leaktight, resistive drop must be low enough to prevent arcing within sensor |
| Fast helium drain to induce global quench | No bulky dump structures outside coil   | Slower, less uniform heating of coil than resistive heater. Not applicable to CICC                                 |

It is our opinion that cowound, resistive heaters have the fewest theoretical limits in almost all cases. They will always achieve the least peaking in energy deposition and internal voltage. They can also be used in all applications as advanced quench detection sensors, as discussed below; and they can sometimes be used as structural backing elements.

**Cold Switches.** A typical cold switch is shown in Fig. 7. Here, by splitting a coil into two halves it is possible to apply a resistive dump voltage to both sides without opening the normally closed switch across the coil as a whole.

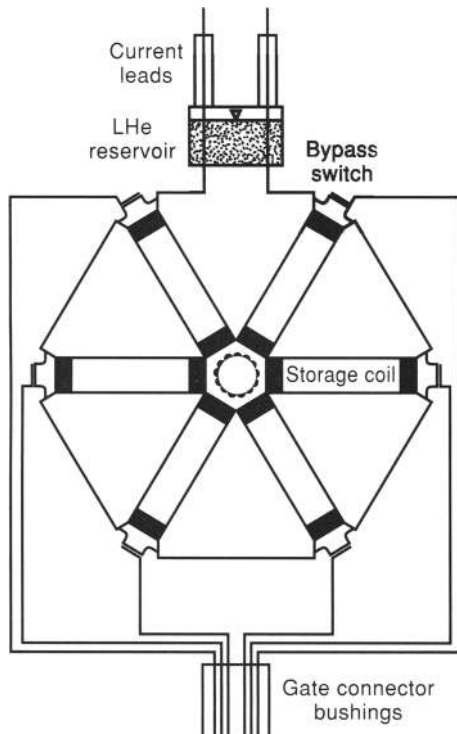
In this circuit, the magnets are charged through the power supply, which is then shunted through the normally closed switch, allowing current to circulate through the magnets. If a quench begins in either half of the center-tapped coil, the



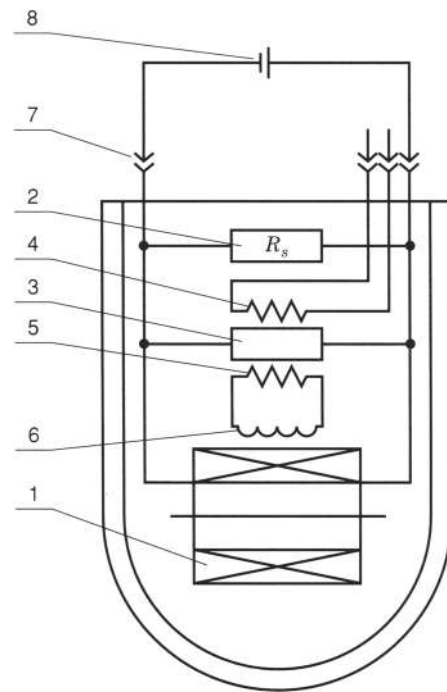
**Figure 7.** A representative dump circuit using cold diodes.

voltage across the switch is zero, but the inductive voltage on the unquenched half is equal and opposite to the inductive and resistive voltage across the quenched half. Above the cold diode threshold of 10 V to 15 V, two of the four diodes begin conducting. The resistor is sized so that its voltage drop is much larger than the 1 V to 1.5 V forward drop across the diode, but small enough to satisfy Underwriter Laboratory safety limits, typically 100 V. The resistor itself might be a heater pad on the outer layer of the winding, inducing a longer quench zone and more uniform heat distribution in the coil.

In the circuit shown in Fig. 8, cold transistors are used, instead of cold diodes, in order to block high voltages during both charging and discharging of the magnet system. Kaerner (10) found that the only active devices blocking bipolar voltages at helium temperatures are NPT (non-punchthrough) IGBTs (Insulated Gate Bipolar Transistors). In 1995, a single device would carry 300 A at 5.9 K and block 1200 V. In this system,



**Figure 8.** Arrangement of cold bypass switches in toroidal SMES system [Kaerner et al. (10)].



**Figure 9.** Highly sensitive passive protection circuit: (1) Superconducting magnet, (2) Shunt resistor, (3) Superconducting switch, (4) Main switch heater, (5) Auxiliary switch heater, (6) Secondary winding, (7) Detachable current leads, (8) Current supply [Anashkin et al. (11)].

after quench is detected, the “weak” coil is shunted by the IGBT, the coil is rapidly dumped by internal heaters, then the rest of the coils are ramped down. With this scheme, only  $1/n$  coils of the stored energy has to be dumped at helium temperatures, but there is only one set of vapor-cooled leads.

Probably the most commonly used cold switch is a normally closed superconducting switch. During coil charging, it is held normal by an actively driven heater. The heater is then turned off and frequently the charging leads are then disconnected to reduce losses. These can be used in conjunction with a cold transformer in order to lower the voltage threshold for passive dumping. Anashkin et al. designed a passive circuit that is capable of responding to very low rates of current decay (11). The circuit in Fig. 9 demonstrated a superconducting switch normal transition for a magnet field decay of  $2 \times 10^{-4}$  T/s.

The current decay due to a normal zone induces current in the secondary winding and auxiliary heater which drives the superconducting switch normal. Most of the magnet current is now forced through the shunt resistor. The shunt resistor can be placed either inside or outside the cryostat. An external resistor would be favored for applications with a large number of magnets or expected training quenches, such as accelerator rings, in which cooldown time and refrigeration requirement dominate. Internal placement is favored for very high performance magnets, where the resistor can also be used as a magnet heater to force uniform quench.

## QUENCH DETECTION

Superconducting magnets have traditionally used relatively simple methods for detecting a quench. Voltage taps on the

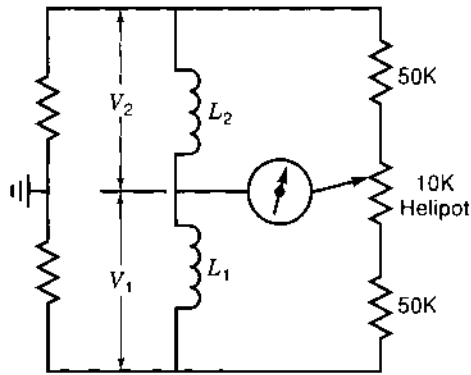


Figure 10. Quench detection bridge circuit [Purcell et al. (12)].

them, such as the voltage across two layers or two pancakes. This method, being the crudest and most likely to be overwhelmed by inductive noise, is used with very slow-charging dc magnets. The signal-noise ratio of voltage taps is most frequently enhanced by the use of a bridge circuit to cancel out the inductive signal. Two nearly equal inductive signals, such as the inductance of two adjacent double pancakes or an external inductance with the same or proportional  $dI/dt$  as the coil are put in two arms of the bridge. An external resistance is then balanced against the coil section whose resistance is being measured, and the ramped superconductor signal is zeroed out. Any voltage should then be equal to the resistive drop across the magnet during a quench. A typical bridge circuit design by Purcell (12) is shown in Fig. 10.

surface of a winding are the most common, and changes in temperature, pressure, and flow have also been used, sometimes as supplements to voltage taps. These methods become inadequate when a magnet has to operate in a strongly pulsed field. The problem of induced noise voltage is exacerbated by large size and large transients in flow or temperature. In the future of large commercial systems, requiring high reliability, quench detection is likely to be the weak link in the magnet protection system. While series redundancy can provide arbitrarily high reliability in external protection circuits, it will require advanced concepts to guarantee high signal-noise ratio, rapid detection, along with leak-free and discharge-free quench detection under all operating conditions.

**Voltage Sensors**

**Conventional Voltage-Taps and Bridges.** The voltage across a coil or a section of a coil is measured by tapping into the external surface of the conductor through the insulation. This is commonly done at sections of the winding pack that are physically close to each other, in order to avoid unnecessary inductive pickup, but which may still have significant voltages between

**Pick-Up Coils.** If a coil is connected to a low transient impedance external circuit, it is possible to detect a quench by inductive pickups to an overall magnet change in current. An external dump circuit can then be triggered to accelerate the coil dump. A typical use of a pickup coil quench detector by Sutter (13) is shown in Fig. 11.

A variation on this technique is to place higher-order pickup coils around a superconductor as a multipole antenna (14). During a quench, circulating supercurrents will be suppressed more rapidly than overall current, causing a rapid collapse of diamagnetism in the conductor that can be detected by the antenna. Both techniques have the advantage of not interrupting the coil insulation.

**Detection Circuits.** The basic detection circuit by St. Lorant (15), shown in Fig. 12, would trigger a quench dump, based on a resistive voltage level that exceeds a preset threshold, typically 10 mV to 200 mV. In a bridge circuit, resistive voltages of either polarity must be expected. Low-pass filters prevent false positive signals due to ambient noise. The comparator is the quench detector itself. The rectifier allows either positive or negative voltages to be used as unipolar digital triggers. A

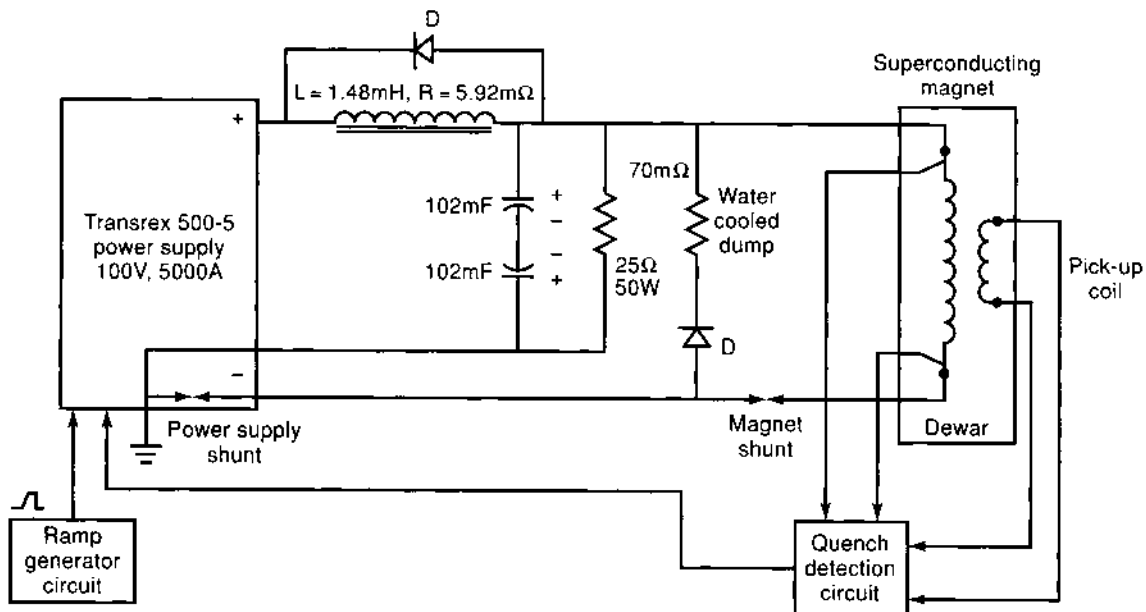
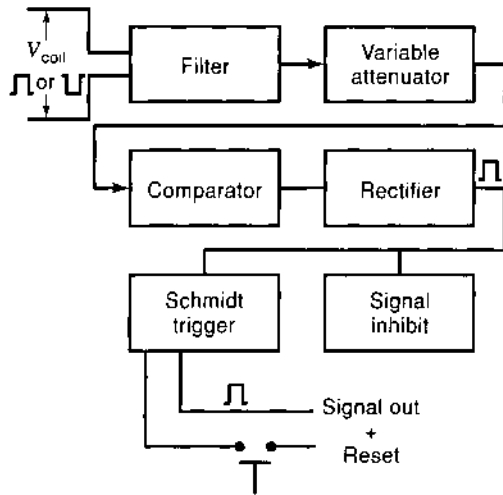


Figure 11. Use of pickup coil quench detector in the Fermi National Accelerator Laboratory's energy doubler magnet [D. F. Sutter et al. (13)].



**Figure 12.** Basic quench detection circuit block diagram [S. J. St. (Lorant) et al. (15)].

signal inhibit may be used to prevent quench triggers during coil ramping, and the Schmidt trigger creates a trigger pulse of fixed amplitude and duration.

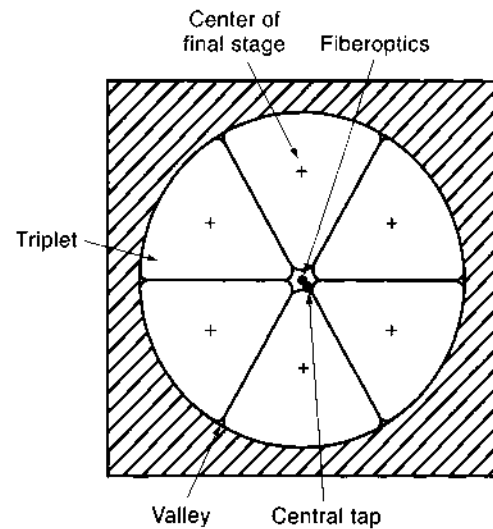
#### Advanced Detection Techniques

In a large magnet, the terminal voltage during pulsing may be as high as 5 kV to 25 kV. The POLO coil set a world's record for CICC at 23 kV (16), while recent designs of large CICC coil systems, such as ITER (10 kV), TPX (7.5 kV), and NAVY SMES (10 kV) have specified voltages in this range, while the important pool-boiling Anchorage SMES system is being designed to withstand 4.2 kV. By contrast, a large number of quench simulations have shown that in order to hold hot-spot temperatures to 150 K, a quench must be detected at a threshold voltage of 0.2 V to 1.0 V (17,18). If, as in the TPX design (19), a desired value of signal/noise of 10:1 is specified, the quench detection sensors must be capable of reducing noise levels to ~20 mV to 100 mV. This implies that the voltage rejection capability of the quench detection system should be on the order of 100,000–500,000:1. Although this may seem optimistic, recent experiments at MIT, the Lawrence Livermore Laboratory, and the Ecole Polytechnique Federale de Lausanne have demonstrated the feasibility of such high levels of voltage noise rejection (17,20).

New techniques that promise the greatest cancellation of inductive noise and the highest signal/noise ratios include the use of internal sensors, digital differencing and signal processing, and fiberoptic temperature sensors.

**Advanced Voltage Sensors.** Several noise rejection techniques can be used simultaneously in order to obtain ultrahigh system noise rejection. Individual concepts that can be used include:

1. Cowinding an insulated voltage sensor with the CICC cable
2. Extracting voltage sensors at joints, but terminating them within the winding; thus subdividing the terminal voltages and localizing quench information, without exacerbating electrical integrity or leak tightness
3. Placing the cowound sensor in the part of the cable best calculated to reject transverse, longitudinal, and self-field



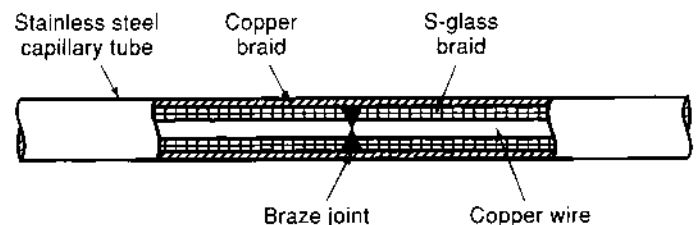
**Figure 13.** Placement of sensors within cable.

voltages. Placement of voltage taps in the center of the final stage subcable was favored by TPX, as shown in Fig. 13 (17)

4. Terminating the voltage sensor internally, further localizing and subdividing the winding into sections, and permitting another level of voltage differencing, such as simple differencing and central difference averaging. The two halves of the sensor form one solid wire that can be cabled and wound with the rest of the conductor. A method for forming an internal termination in a continuous sensor is shown in Fig. 14
5. Further signal processing, such as using integrated volt-seconds, rather than simple voltage thresholds as detection criteria

These techniques are now described in more detail.

A cowound sensor has been used previously on the US-Dual-Purpose Coil (US-DPC) coil, where an insulated wire was wound along an edge of the conduit on the outside (21). A cowound, insulated aluminum strip in the B&W SMES coil is also being used as structural reinforcement. If the sensor is cabled on the inside of the conduit, as shown in Fig. 13, the degree of noise cancellation will improve by at least another order of magnitude. The ITER QUELL (Quench on Long Lengths) experiment demonstrated the cabling and extraction of voltage sensors on the surface of a cable, but inside the conduit, for a 100 m length of conductor. In the QUELL experiment, the rejection of transverse voltage was more than 400 times better than



**Figure 14.** Two-sided voltage sensor, showing internal termination between two halves.



that of conventional voltage taps (21). Two TPX experiments, one with a copper cable, the other with a NbTi cable, demonstrated cabling of sensors in different positions of a full-scale cable (17). Depending on the sensor position, noise rejection ranged from 600–60,000:1.

**Extracting Voltage Sensors at Joints.** Joint extraction from the ends of a continuous winding, while difficult, is much less intrusive than extraction every two layers or pancakes through the winding pack. The joints must already be accessible for servicing and capable of accommodating helium stubs and lines and instrumentation feedthroughs. The biggest problem with extracting voltage sensors at the joints is ensuring that they will not leak helium and that they won't be subject to electrical breakdown. The approach to preventing leaks is two-fold: (1) injecting a moderately sized seal area with a mineral-filled sealant, such as Stycast, with a good thermal match to the seal metal, and (2) designing the seal to be long, in compression, and mechanically redundant in terms of interrupting individual microcracks, and (3) made up of a separate, weldable piece that can be thermal shock tested or cycled before installation. A simple version of this joint was demonstrated at MIT that, after two cooldowns to nitrogen, was capable of holding 70 bars on one side and vacuum on the other (17).

Design against electrical breakdown is ensured by designing the CICC itself, so that there can never be 160 V across the voltage sensor. In this case, the sensor insulation will be below the Paschen minimum for helium, irrespective of uncertainties in helium pressure due to leaks or in stray transverse

or longitudinal magnetic fields. Almost any large magnet will satisfy this design approach automatically. For example, a normal zone as long as 1 km with a stabilizer current density of 200 A/m<sup>2</sup> at an average field as high as 13 T, would have a resistive voltage of only 133 V between joints. Depending on the time needed for internal current dump, the current density can be lowered slightly, if needed, in order to compensate for resistivity rising with temperature. This should seldom be necessary.

The surface of the cable is not the best position for emulating the trajectory and therefore the flux linkage of a typical strand in a cable. The best job would be done by a sensor that was cabled into a first triplet, as though it were a strand. The second best position, identified so far, is to place the sensor in the center of the final stage of the cable, which is easier to cable and not vulnerable to conduit broaching or welding. Both positions are far superior to the center of the cable or a natural cabling valley on its surface. Equations for approximating the induced noise voltage for different types of field and sensor placement were derived by Martovetsky (22) and are summarized in Table 1.

**Terminating the Voltage Sensor Internally.** Internal termination of sensors localizes the signal by subdividing the winding into sections and permits another level of voltage differencing, such as simple differencing and central difference averaging. In TPX, the voltage sensors were terminated internally at 1/6, 1/3, 1/2, 2/3, 5/6, and completion of the distance through the winding pack, as shown in Fig. 15. Even before using differencing techniques, this should further reduce the noise voltage by

**Table 1. Noise Voltages for Different Sensor Positions**

|                               | Transverse Field  | Self Field   | Parallel Field  |
|-------------------------------|---|--|---|
| Wire in center                | Grad-B<br>$V = \frac{1}{2} \int_0^L \frac{\partial \hat{B}_z}{\partial r} r_{\text{last stage}}^2 dl$ Geometry<br>$V = \hat{B} \frac{N_{\text{turns}} d_{\text{cable}}^2}{2}$                             | $V = \frac{\mu I r^2}{4\pi R^2} L$                                 | $V = \int_0^L \hat{B}_l \sum_{n=1}^4 \frac{\pi r_{\text{subcable},n}^2}{l_{p,n}} dl$  |
| Wire in valley                | Grad-B<br>$V = \frac{1}{2} \int_0^L \frac{\partial \hat{B}_z}{\partial r} r_{\text{last stage}}^2 dl$ Geometry<br>$V = f(\text{Inhom}) \times \frac{1}{2\pi} \int_0^L \hat{B}_z r_{\text{cable}} dl$      | $V = \frac{\mu I L}{4\pi} - \frac{\mu I L}{12\pi}$                 | $V = \int_0^L \hat{B}_l \left( \sum_{n=1}^4 \frac{\pi r_{\text{subcable},n}^2}{l_{p,n}} - \frac{\pi r_{\text{cable}}^2}{l_{p,\text{cable}}} \right) dl$ |
| Wire in triplet               | Grad-B<br>$V = \frac{1}{2} \int_0^L \frac{\partial \hat{B}_z}{\partial r} r_{\text{last stage}}^2 dl$ Geometry<br>$V = f(\text{Inhom}) \times \frac{1}{2\pi} \int_0^L \hat{B}_z r_{\text{last stage}} dl$ | $V = f(\text{Inhom}) \times \frac{\mu I L}{12\pi}$                 | $V = f(\text{Inhom}) \int_0^L \hat{B}_l \sum_{n=1}^4 \frac{\pi r_{\text{subcable},n}^2}{l_{p,n}} dl$  |
| Wire in center of final stage | Grad-B<br>$V = \frac{1}{2} \int_0^L \frac{\partial \hat{B}_z}{\partial r} r_{\text{last stage}}^2 dl$ Geometry<br>$V = f(\text{Inhom}) \times \frac{1}{2\pi} \int_0^L \hat{B}_z r_{\text{last stage}} dl$ | $V = \left( \frac{\mu I L}{12\pi} - \frac{\mu I L}{16\pi} \right)$ | $V = \int_0^L \hat{B}_l \sum_{n=1}^3 \frac{\pi r_{\text{subcable},n}^2}{l_{p,n}} dl$  |

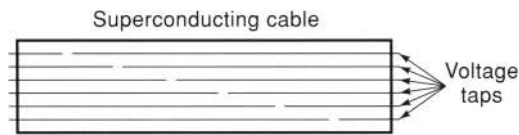


Figure 15. Inexpensive voltage sensor concept.

another factor of six, in a way similar to putting voltage taps on each one of six double pancakes. Taking a simple difference between, say, the signal at 1/6 and 1/3 is frequently ineffective, since there may be a systematic gradient through the winding pack due to eddy currents being turned around by a break on one side of the winding pack, but not the other (23). However, central difference averaging, in which one-half the first and third sensor signal are subtracted from the middle sensor, as first proposed by Yeh and Shen (24), can be effective in canceling out gradients. For example, if voltages are measured across double pancakes, the threshold voltage across the first six pancakes would then be  $0.5V_{1-2} - V_{3-4} + 0.5V_{5-6}$ . In simulations of TPX, the degree of further cancellation using CDA was 5–20 (17). Internal termination has the additional benefit of localizing quench initiation and of following the sequence of quench propagation through a winding pack.

Using a sensor insulation that is compatible with inexpensive seam-welding of the steel capillary tube permits the redundant use of multiple voltage sensors in a cable. Fabrication of sensors with S-glass braid, as shown in Fig. 16, was demonstrated in the QUELL experiment and the ITER CS Model Coil. Some combination of Formvar, teflon, and/or kapton would be used for NbTi. Teflon would probably be most desirable, because of its ability to eliminate seams at modest temperature and pressure, thus eliminating tracking. XMPI Kapton, at higher temperature and pressure, should have even higher performance.

**Further Signal Processing, Such as Using Integrated Volt-Seconds, Rather Than Simple Voltage Thresholds as Detection Criteria.** This is particularly effective in screening out short disturbances, such as stick-slip, flux jumps, and plasma disruptions. In TPX simulations, an additional factor of 10–20 in signal/noise was achieved by using a volt-second window, instead of a voltage threshold (17). Further improvements in signal/noise ratio through signal processing have been proposed through the use of carrier signals and synchronous detectors.

In summary, although there is clearly a broad range of design specific signals and signal/noise ratios, the following rules of thumb might be used for preconceptual design of a voltage noise reduction system with the goals of achieving a signal/noise level improvement of  $>10^5$  (i.e., 10 kV down to 100 mV): Split the winding pack into sections  $(5-10) \times$  Place internal sensor in cable final stage  $(100-1000) \times$  central difference

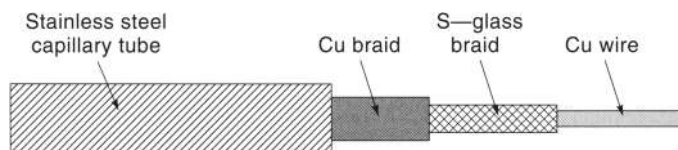


Figure 16. Six internal voltage sensors, terminated at equal distances through cable.

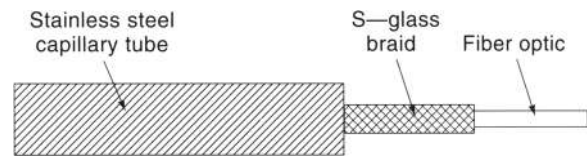


Figure 17. Fiber-optic sensor.

average  $(5-10) \times$  filter, integrate, signal process  $(5-10) = 10^4-10^6$ . Two TPX noise injection experiments, one with copper, the other with NbTi cable, demonstrated transverse field noise rejections of up to 60,000 (17). The ITER QUELL coil had both conventional and cowound voltage taps on the outside of the cable in its natural valleys. In this experiment, the voltage rejection of the cowound sensor was 80 times better than that of the noninductive winding with a voltage rejection of 6:1 or 500:1 total. Another noise-rejection method used in ITER of differencing multiple in-hand windings was calculated to achieve noise rejection ratios of 300–400. A better way to normalize the results to design for all noise sources is to show that the  $V/(km \cdot T/s)$  are  $\ll 1$  for transverse field and  $\ll 10$  for parallel and self-field, corresponding to  $\ll 1$  V for 1 T/s over a kilometer. Placement of the sensor in the center of the final stage or in a triplet have, so far, met this goal in all experiments.

**Advanced Fiberoptic Temperature Sensors.** Conventional temperature sensors have a number of disadvantages: they cannot be inserted inside a cable-in-conduit, don't measure cable temperature directly, and have a significant time lag. Like conventional voltage taps, they can't give coverage to a winding without frequent penetration of the insulation system. Furthermore, thermocouples are insensitive at helium temperatures, while carbon glass resistors and resistance temperature detector (RTD) sensors are insensitive at quench temperatures. The use of fiber optic temperature sensors, protected by a steel capillary, as shown in Fig. 17, has several advantages over conventional sensors: (1) They can be inserted directly into the helium flow channel with a very short thermal time constant, (2) they are insensitive to pulsed magnetic fields (25), helium flow, and pressure, (3) they are very small ( $<50 \mu m$ ), so as many fibers as desired can be placed in a single steel can, (4) the length of a fiber between joints is practically unlimited, 50 km being a routine commercial length, and (5) they have great scientific potential in the use of signal processing to provide a complete profile of temperature and field versus length and time.

Fiberoptic temperature sensing works on the principle of measuring optical path length, taking advantage of the temperature dependence of the glass index of refraction. The change in phase from either effect is

$$\Delta\phi = \frac{2\pi}{\lambda} [n\Delta L + \Delta nL] \quad (2)$$

where  $\Delta\phi$  is the phase shift and  $n$  is the index of refraction.

The major problem with this technique is the rejection of path length changes due to mechanical strain in the glass. One strain-rejection technique is to decouple the glass from the conduit containing it, so that they don't share strain. At M.I.T., a fiber was inserted in a 1.0 mm stainless steel capillary

**Table 2. Sensitivity of Dual-Mode and Dual-Polarity Fiber to Temperature and Strain**

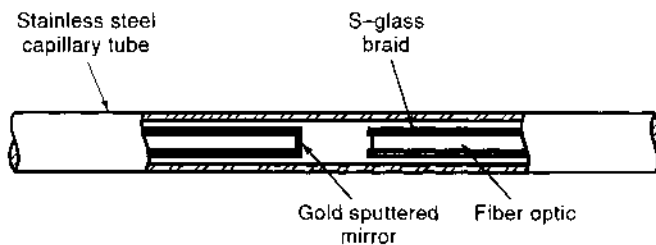
| Method       | Temperature Sensitivity (radians/m-K) | Strain Sensitivity (radians/m-ε) |
|--------------|---------------------------------------|----------------------------------|
| 2 Polarities | 1.2                                   | $5 \times 10^3$                  |
| 2 Modes      | 2.18                                  | $55 \times 10^3$                 |
| 2 Colors     | 0.1                                   | $0.5 \times 10^3$                |

tube and several turns were wound on an 80 mm steel tube. The assembly was heated to 700°C, then cooled to 4 K without damage, thus demonstrating strain decoupling and the absence of a capstan multiplying effect on fiber tension (26). The QUELL experiment demonstrated that the relatively loose fit of a copper-clad fiber can reduce the strain-sharing by a factor of 10–100. However, mechanical strain was still a dominant effect, since the changes in index of refraction are <1%. It has been demonstrated by Smith that the mechanical strain and temperature-dependent signals can be almost totally decoupled by using two independent signals with different strain and temperature dependences (27). This can be accomplished by the use of polarization maintaining (PM) optical fiber, two color operation, or two mode operation. Each polarity, color, or mode has a different, calibrated ratio of strain and temperature dependence, so that the gain on one phase shift can be adjusted to ‘tune out’ the strain signal. This method has the additional benefit that it can also be used to measure the integrated conductor strain. Overall system cost-performance analysis indicates that two mode operation will probably be the most cost-effective long-term solution. Typical sensitivities are shown in Table 2.

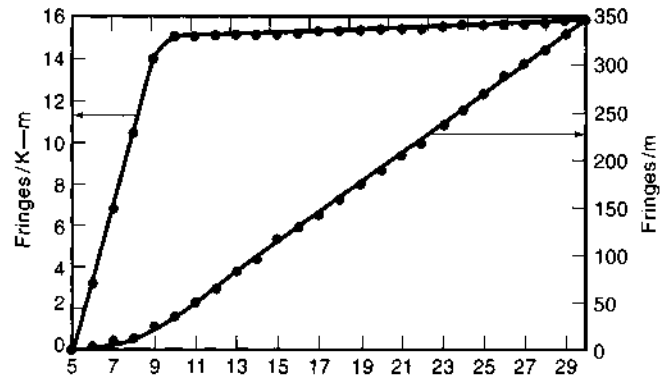
Glass-fiber must be clad in order to prevent the entrance of water during handling. Commercial acrylic fibers are adequate for NbTi, but will not survive Nb<sub>3</sub>Sn, NbAl, or BSSCO heat treatments at >600°C. Although the surrounding steel can contain any spattering of the cladding, this solution is, at best, messy. Coatings of copper, gold, or graphite can protect the glass fiber and survive heat treatments. Unjacketed acrylic coatings should be used outside of the joint region, so that high-voltage cryostat feedthroughs won’t be necessary.

The same design principles of internal termination applied to voltage sensors within a winding pack apply to fiber optic temperature sensors. If they are internally terminated by sputtering mirrors on polished cuts, as shown in Fig. 18, the integrated noise temperatures are reduced by the number of subdivisions, then reduced further by differencing techniques.

Subdivision also allows greater localization of quench events. In the special case of irradiated magnets, it also helps keep the signal attenuation manageable (<60 dB). The method for subdividing a fiber is to cut the fiber at the desired length, polish and silver the end. The fiber end is then a mirror and



**Figure 18.** Internal termination of fiber optic temperature sensor.



**Figure 19.** Temperature sensitivity of glass fibers.

the laser light signal returns to the splitter and detector to form the information-carrying half of a Michelson interferometer.

After converting phase shifts to voltages, the same signal processing concepts used for voltage sensors can also be used for fiber optics. Integrating the number of fringe shifts over a time window, such as one second, can achieve an order of magnitude improvement in signal-noise ratio. The signal-noise ratio of the fiber optic quench detector is naturally enhanced by the increase in sensitivity with rising temperature. Simulations of quench and disturbances in TPX showed minimum signal/noise ratios of 600:1 for quench detection within a second (26). This was both within design criteria and superior to internal voltage sensors. The temperature sensitivity of the fibers has been measured with and without cladding and can be characterized by an initial quadratically increasing fringe count according to the equation:

$$\int_5^T a dT = 1.75T^2 - 17.5T + 43.25; 5K < T < 9K \quad (3)$$

The increasing sensitivity saturates at about 15 fringes/m-K, and then only increases to about 25 fringes/m-K at room temperature, as shown in Fig. 19.

If the fiber optic sensor is considered to be a length-temperature rise integral measurement, the sensitivity is 200 times higher at 30 K than at 5 K. Therefore, a global disturbance that raises the helium temperature throughout a 1 km winding from 5 K to 5.2 K will give a signal that is smaller than a quench that raises 1 m of conductor to 30 K.

For an unclad fiber, the sensitivity is almost exactly one-tenth that of a clad fiber, the thermal strain of the plastic cladding acting as an amplifier of the temperature signal. The curves in Fig. 19 can be used for design of temperature sensors with NbTi; but fringes/m-K should be multiplied by 0.1 for design with Nb<sub>3</sub>Sn.

Fiber optic sensors would use the same sort of prefabricated and pretested seal as the voltage sensors. However, a few inches of clearance would be needed between a joint and the initial position of the seal piece, in order to use a handheld field splicer. The optical fiber can then be coiled into the pocket of the seal piece.

**Quench Detection Conclusions**

1. Fiber optic temperature sensors and internal voltage sensors have been shown by simulation and experiment to

improve signal/noise ratios in quench detection systems by several orders of magnitude.

2. A leak-free method for extracting sensors has been demonstrated, and a redundant and replaceable sealing system has been designed.
3. A method for coil/sensor electrical design has been defined that is robust against arbitrary helium pressures and magnetic fields. Enormous safety margins are feasible with NbTi and fused teflon or kapton insulation.
4. Advanced quench detection sensors can also be used as scientific instruments, measuring the internal properties of CICC conductors.

## MAGNET PROTECTION CRITERIA

### Adiabatic Protection Criterion

A popular and conservative protection criterion is to assume that there is no heat transfer from the local hot spot where quench is initiated and that all Joule heating is absorbed by the stabilizer. In this case, the relation between the peak allowable hot spot temperature and the  $J^2t$  integral of the conductor stabilizer during a coil dump is a unique property of the stabilizer material, usually copper. The maximum allowable current density is then determined by the peak allowable hot spot temperature and the peak allowable terminal voltage for a coil dump. Typical values of peak temperature allowables are 80 K to 200 K. If a coil is completely supported in compression by external structure, as an accelerator coil in a large iron yoke, it is possible to design up to 450 K or the melting point of solder (28). A typical allowable terminal voltage for a pool-boiling magnet is 1 kV to 3 kV, and 3 kV to 20 kV for a CICC magnet. However, as discussed in the section on electrical protection, the fundamental limits on voltage for both topologies are strong functions of specific design.

Irrespective of the coil temperature allowables, the selection of the peak terminal voltage fixes the minimum  $L/R$  dump time constant (s) at

$$\tau_D = \geq \frac{2W_m}{V_{\max} I_{\text{cond}}} \quad (4)$$

where  $W_m$  is the stored energy in the magnet (J),  $V_{\max}$  is the peak allowable terminal voltage on dump (V), and  $I_{\text{cond}}$  is the conductor current (A). For a system dumping its energy into a linear external resistor with negligible resistive voltage within the coil system, the magnet current after switching is simply

$$I(t) = I_0 \exp\left(-\frac{R_D t}{L_m}\right) \quad (5)$$

where  $R_d/L_m$  is the dump time constant  $\tau$  (s).

Since a quench will not be detected immediately, there is also a delay time before any action is taken. This has sometimes been specified as 1–2 s as a design goal, but it is a function of the signal/noise ratio of the quench detection system, as explained in that section. The allowable current density in the copper then is fixed by the combination of delay plus dump time as

$$j_{0cu}^2 \left( t_{\text{delay}} + \frac{\tau_D}{2} \right) = Z(T_f) \quad (6)$$

where  $t_{\text{delay}}$  is the maximum possible time delay (s),  $T_f$  is the final temperature at the end of dump (K), and  $G(T_f)$  or, alternatively in some texts  $Z(T_f)$ , is a property, unique to each material or combination of materials, defined as

$$Z(T_f) = \int_{T_0}^{T_f} \frac{C(T)}{\rho(T)} dT \quad (7)$$

This integral can also be thought of as the  $J^2t$  integral of the current in the conductor during a dump, and is thus usually represented with the units ( $A^2/m^4\cdot s$ ). An analytical approximation for  $G(T_f)$  for copper from a bath temperature of 4.2 K is

$$Z(T_f) = \left( \frac{1363}{T_f + \frac{4173}{T_f + 47.89}} \right) \times 10^{16} \quad (8)$$

Another way of stating the design constraint, derived by Iwasa and Sinclair (29) defines  $Z(T_f)$  as the integral from the detection temperature to the final temperature and models the presence of materials other than substrate by a correction factor  $a$ , which is the enthalpy change ratio of substrate/all other materials, which can be determined by table lookup or approximated as a volumetric ratio. The maximum allowable current density in the substrate is then

$$j_0 = \sqrt{\left( \frac{1+a}{a} \right) \frac{V_D I_0 Z(T_D, T_f)}{E_M}} \quad (9)$$

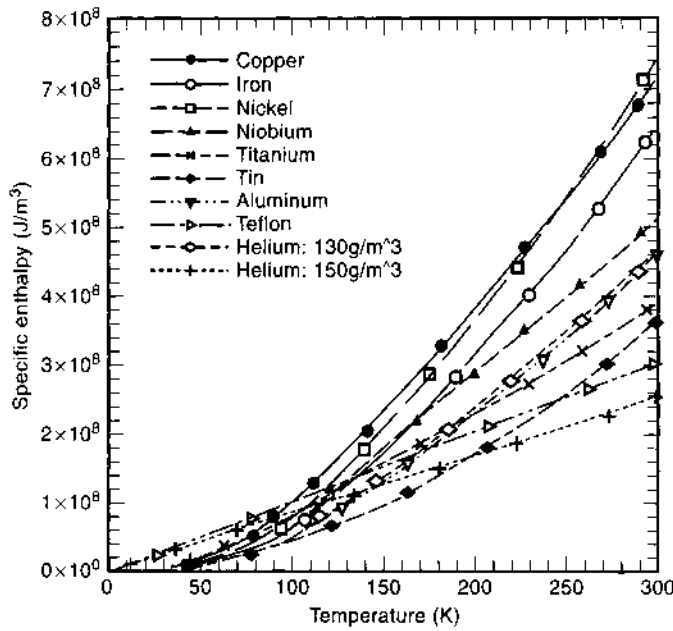
where  $V_D$  is the dump voltage (V),  $I_0$  is the operating current (A), and  $E_m$  is the stored energy (J).

Iwasa's  $a$ -factor can be quickly estimated with a curve of volumetric specific enthalpies and the volumetric ratios of the different constituent materials in a design. The specific enthalpies of copper, aluminum, iron, nickel, niobium, titanium, tin, teflon, helium at 130 kg/m<sup>3</sup> and helium at 150 kg/m<sup>3</sup> are plotted from 4.5 K to 300 K in Fig. 20.

Copper and nickel are the two best materials above a maximum temperature of 100 K to 120 K. However, there is no more than a factor of two difference between the best and the worst metals. Copper is about 60% better than aluminum; but since aluminum is three times lighter than copper, it is twice as good as copper, if energy/mass is a more important consideration than energy/volume. Copper is clearly also the best material from an adiabatic  $J^2t$  criterion. Therefore, if the design is limited by the hot spot temperature in a given volume, copper is the material of choice.

The curves of  $Z(T_f)$  versus  $T$  (K) in Fig. 21 appear in Iwasa's casebook (30).

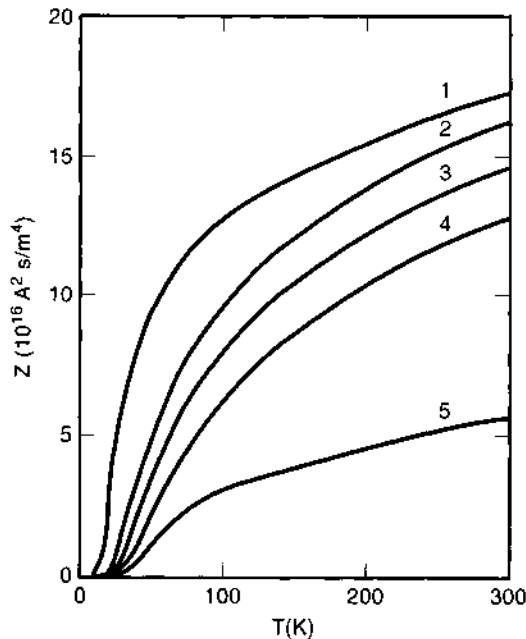
While high-purity silver has the best  $Z(T_f)$ , far-less expensive oxygen-free coppers are nearly as good. Copper is two-three times as good as high purity aluminum. The  $Z(T_f)$  versus  $T_f$  (K) curves in Fig. 22 were calculated specifically for a tradeoff between copper and stronger, less high purity aluminums. The curves show that copper is 20 to 25 times better than aluminum alloys. The implication is that the cross-sectional area of copper needed for protection in a design where the available enthalpy is dominated by the stabilizer would be four to five times less than that of aluminum. In a design in which other components,



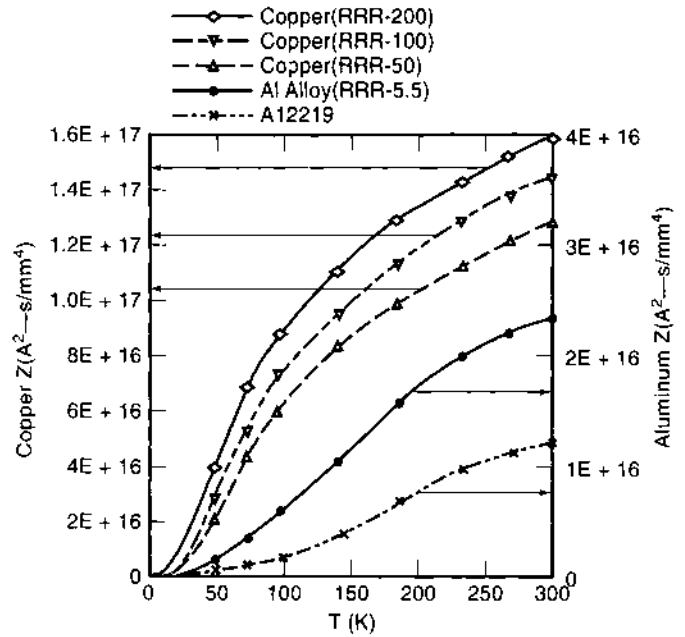
**Figure 20.** Specific enthalpies of magnet materials ( $J/m^3$ ) versus temperature (K).

such as superconductor, helium, or other structures were important, the superiority of copper would be reduced. Magnetoresistivity would also reduce the quantitative superiority of copper. In magnetic energy storage designs, the main factor that improves the relative position of aluminum is that an aluminum such as 2219 can also be used as a structural material. In a design that requires a structural cross-section that is several times larger than the cross-section needed for protection, the volumetric advantage of copper would disappear.

The values of  $Z(T_f)$  from Fig. 22 are listed in Table 3.



**Figure 21.**  $Z(T_f)$  functions: (1) Silver (99.99%); (2) Copper (RRR 200); (3) Copper (RRR 100); (4) Copper (RRR 50); (5) Aluminum (99.99%) (30).



**Figure 22.**  $Z(T_f)$  functions: (1) Copper (RRR = 200), (2) Copper (RRR = 100), (3) Copper (RRR = 50), (4) Aluminum alloy (RRR = 5.5), (5) Aluminum alloy 2219.

**QUENCH PROPAGATION**

**Adiabatic (Potted) Magnets**

Adiabatically cooled magnets are selected for applications that are relatively small, and quasi-steady state, since they are incapable of absorbing large amounts of local energy. They are particularly suited to applications where there can't be any cryogenic fluid within the magnet and where small, compact winding packs are required. Iwasa has argued convincingly (31) that most high-temperature superconductor (HTS) magnets are also likely to be adiabatically cooled, because the local energy absorption is improved by orders of magnitude at higher temperatures.

Adiabatically cooled magnets may be protected either by internal or external energy dumps. If the simplifying assumption is made that the thermal conductivity and heat capacity of all materials is temperature independent, the longitudinal quench propagation velocity is expressed by the balance of constant local heating density and thermal diffusion through the winding pack as (30)

$$v_{\text{propagation}} = J \sqrt{\frac{\rho_n k_n}{C_n C_s \left( \frac{T_{cs} + T_c}{2} T_{op} \right)}} \tag{10}$$

where  $v_{\text{propagation}}$  is the longitudinal quench propagation velocity in the winding direction (m/s),  $J$  is the current density in the composite wire ( $A/m^2$ ),  $\rho_n$  is the electrical resistivity of the composite wire (W-m),  $k_n$  is the thermal conductivity of the normal wire (W/m-K),  $C_n$  and  $C_s$  are the heat capacities of the wire in its normal and superconducting states respectively ( $J/kg\cdot K$ ), and  $T_{op}$ ,  $T_c$ , and  $T_{cs}$  are the operating and superconductor transition temperatures, respectively (K) (32). Again, because the heat capacity of all materials rises

**Table 3.  $Z(T)$  Functions for Copper and Aluminum**

| Temperature | Copper,<br>RRR = 200   | Copper,<br>RRR = 100   | Copper,<br>RRR = 50    | Al Alloy<br>(RRR = 5.5) | Al2219-T85             |
|-------------|------------------------|------------------------|------------------------|-------------------------|------------------------|
| 4           | 0                      | 0                      | 0                      | 0                       | 0                      |
| 10          | $3.173 \times 10^{14}$ | $1.597 \times 10^{14}$ | $7.953 \times 10^{13}$ | $2.280 \times 10^{12}$  | $5.688 \times 10^{11}$ |
| 20          | $4.776 \times 10^{15}$ | $2.481 \times 10^{15}$ | $1.256 \times 10^{15}$ | $2.584 \times 10^{13}$  | $6.422 \times 10^{12}$ |
| 50          | $3.847 \times 10^{16}$ | $2.798 \times 10^{16}$ | $2.005 \times 10^{16}$ | $9.466 \times 10^{14}$  | $2.517 \times 10^{14}$ |
| 75          | $6.757 \times 10^{16}$ | $5.457 \times 10^{16}$ | $4.283 \times 10^{16}$ | $2.898 \times 10^{15}$  | $8.521 \times 10^{14}$ |
| 100         | $8.633 \times 10^{16}$ | $7.274 \times 10^{16}$ | $5.990 \times 10^{16}$ | $5.538 \times 10^{15}$  | $1.827 \times 10^{15}$ |
| 145         | $1.097 \times 10^{17}$ | $9.567 \times 10^{16}$ | $8.206 \times 10^{16}$ | $1.050 \times 10^{16}$  | $4.055 \times 10^{15}$ |
| 190         | $1.285 \times 10^{17}$ | $1.142 \times 10^{17}$ | $9.793 \times 10^{16}$ | $1.504 \times 10^{16}$  | $6.451 \times 10^{15}$ |
| 240         | $1.410 \times 10^{17}$ | $1.268 \times 10^{17}$ | $1.116 \times 10^{17}$ | $1.941 \times 10^{16}$  | $9.016 \times 10^{15}$ |
| 273         | $1.505 \times 10^{17}$ | $1.360 \times 10^{17}$ | $1.192 \times 10^{17}$ | $2.195 \times 10^{16}$  | $1.062 \times 10^{16}$ |
| 300         | $1.561 \times 10^{17}$ | $1.416 \times 10^{17}$ | $1.247 \times 10^{17}$ | $2.386 \times 10^{16}$  | $1.186 \times 10^{16}$ |

much more quickly with temperature than resistivity does, the quench propagation velocity of HTS quenches should be much lower than that of LTS quenches and the coils will be harder to protect.

### Pool-Boiling Magnets

**Pressure Rise.** The quench pressure rise during the quench of a pool-boiling magnet involves a design tradeoff. The maximum pressure can most easily be controlled by the setting of external pressure relief valves or rupture disks (33). The additional pressure in the magnet due to pressure drops in the vent lines and the disks or valves can be solved by the time-dependent model of Krause and Christensen (34), assuming frictional, adiabatic (Fanno) flow in the vent lines. An alternative, recommended by Powell (35) is repressurization with warm helium gas. The motive for maintaining a high pressure, discussed in the section on electrical integrity, is that it will improve the dielectric strength of warm helium. The disadvantage is that it will increase the required thickness of the liquid helium cryostat, the pulsed eddy currents in the thicker cryostat, and the conduction losses through the cold mass supports.

**Quench Propagation.** Under normal conditions of cooling by a pool of liquid helium, the propagation velocity should obey the proportionality (36):

$$v_p = a(j_0 - j_r) \quad (11)$$

where  $j_r$  is the current density at which the conductor would recover ( $A/m^2$ ), and  $j_0$  is the operating current density ( $A/m^2$ ). If the quench condition causes local dry out, then the propagation velocity is simply (37)

$$v_p = aj_0 \quad (12)$$

### CICC Magnets

Quench propagation in CICC conductors is usually treated as a one-dimensional problem, quench propagating from an interior normal zone toward the inlet and outlet of a hydraulic channel. Bottura has written a general three-dimensional numerical solution for quench propagation (38), which, to the best of our knowledge, is also the only commercially available general quench propagation solution for any coil topology. However, since three-dimensional effects have only a second-order effect on the key design parameters of temperature, pres-

sure, and expulsion velocity, one-dimensional solutions are still used.

**Whole Coil Normal.** In the extreme case of the whole coil going normal at once, quench propagation isn't an issue. This may help to place conservative upper bounds on peak pressure and expulsion velocity for design purposes. Note that simultaneous quench of a whole coil is not a worst case for hot spot temperature. If the coil is designed for internal energy absorption, it is a best case, because the peak/average energy absorption in the coil would be 1.0. If it is designed for external energy absorption, it may be a worst case for the refrigerator by absorbing energy at cryogenic temperature, instead of externally at room temperature; but in terms of the hot spot defined by the local  $J^2t$ , it is not a worst case, because the nonlinear dump time can only be accelerated by adding a larger internal coil resistance to the external resistance. Dresner (39) derived a simple expression for the pressure rise, if an entire hydraulic channel goes normal at once. This case would correspond to practical designs in which a resistive heater was used to ensure uniform internal energy absorption or to an external, uniform energy source being applied to a layer with nearly uniform temperature and field:

$$P_{\max} = 0.65 \left( \frac{Q^2(L_{\text{coil}}/2)^3 f}{D_h} \right)^{0.36} \quad (13)$$

where  $Q$  is the volumetric heating of the helium ( $W/m^3$ ),  $l$  is the half-length of the channel (m), and  $f$  is friction factor, and  $D_h$  is the hydraulic diameter (m). A cable-in-conduit has approximately three times the friction factor of a smooth tube with the same Reynold's number. For a quench pressure wave, Dresner adopts an approximate value of  $f = 0.013$ . When the pressure rise is not much greater than the initial pressure  $P_0$  (Pa), the more exact formulation is

$$P_{\max} = 0.65 \left( \frac{Q^2(L_{\text{coil}}/2)^3 f}{D_h} \right)^{0.36} \left( 1 - \frac{P_0}{P_{\max}} \right)^{-0.36} \quad (14)$$

The same assumptions also predict (40) a helium expulsion velocity of

$$v_{\text{expulsion}} \simeq 0.952 \left( \frac{Q\beta c_0}{\rho C_p} \right)^{2/3} \left( \frac{D_h l}{f} \right)^{1/3} \quad (15)$$

where  $Q$  is the volumetric heating of the helium ( $\text{W/m}^3$ ),  $\beta$  is the constant pressure thermal expansion coefficient,  $c_0$  is the isentropic sound speed ( $\text{m/s}$ ),  $\rho$  is the helium density ( $\text{kg/m}^3$ ),  $C_p$  is the helium specific heat at constant pressure ( $\text{J/kg-K}$ ),  $D_h$  is the hydraulic diameter ( $\text{m}$ ),  $f$  is friction factor, and  $t$  is the time since quench initiation ( $\text{s}$ ). This solution is valid only for the beginning of a quench, since it assumes constant helium properties and neglects inertia and frictional heating.

**Time-Dependent Normal Zone.** Quench propagation scaling for long coils and uniform helium properties were first derived by Dresner (40). Dresner's scaling laws remain valid for describing the early stages of quench in a long hydraulic channel. However, the engineering limits of CICC coils are usually defined by the hot spot temperature, peak pressure, and peak helium expulsion flow toward the end of a quench. By this time, the material properties have changed significantly and heat absorption is dominated by the cable-and-conduit metal, rather than the helium. In this regime, the scaling of quench behavior is more complicated than that described by Dresner and the Shajii/Freidberg theory described next should be used.

Dresner's time-dependent equation for the pressure rise is

$$p - p_0 = C\rho c \left[ \frac{4fZ}{D} \right]^{3/2} \left( \frac{D}{4fct} \right) \quad (16)$$

where  $C$  is a function of quench zone acceleration  $O(1)$ ,  $p_0$  is the initial pressure ( $\text{Pa}$ ),  $\rho$  is the helium density ( $\text{kg/m}^3$ ),  $c$  is the velocity of sound in helium ( $\text{m/s}$ ),  $f$  is the friction factor,  $Z$  is the length of the hot helium piston ( $\text{m}$ ),  $D$  is the hydraulic diameter ( $\text{m}$ ), and  $t$  is the quench time ( $\text{s}$ ). In order to calculate the equivalent piston size, Wachi (41) makes the substitution:

$$CZ^{3/2} = 0.95t^{2.01} \quad (17)$$

Dresner's equation implies that the quench zone  $Z$  is

$$Z = \left[ \frac{3(\gamma - 1)}{2C(3 + 4\gamma)} \right]^{2/3} \left( \frac{Q^2 Dt^4}{\rho^2 c^2 f} \right)^{1/3} \quad (18)$$

Similarly, Dresner solves for central pressure as function of time as

$$p = \frac{3(\gamma - 1)}{3 + 4\gamma} Qt, \quad p \gg p_0 \quad (19)$$

For helium,  $\gamma = 5/3$ , and if  $C = 0.83$  for helium, as proposed by Dresner, these equations reduce to

$$Z = 0.25 \sqrt[3]{\frac{Q^2 Dt^4}{\rho^2 c^2 f}} \quad (20)$$

and:

$$p = 0.207 Qt, \quad p \gg p_0 \quad (21)$$

Shajii and Freidberg (42) rewrite the time dependences of Dresner's quench propagation equations for temperature, pres-

sure, and quench propagation velocity as

$$T_{\text{Dresner}}(t) = \frac{0.10}{R\rho_0 L_{IQZ}} \left( \frac{4d_h}{f\rho_0^2 c_0^2} \right)^{1/3} (\eta_0 J^2)^{5/3} t^{7/3} \quad (22)$$

where  $\eta_0$  is the resistivity of copper ( $\Omega\text{-m}$ ).

$$p_{\text{Dresner}} = 0.21\eta_0 J^2 t \quad (23)$$

and

$$V_{q\text{Dresner}} \equiv \frac{5}{4} X_q = 0.42 \left( \frac{4d_h}{f\rho_0^2 c_0^2} \right)^{1/3} (\eta_0 J^2)^{2/3} t^{1/3} \quad (24)$$

**Shajii Quench Theory.** The Dresner equations were shown by Shajii and Freidberg to apply only in the operating space of short times, low conductor temperatures ( $<25$  K), and long initial quench zones (43). Since this operating regime almost never includes the regimes of greatest interest for design (hotspot temperature, peak pressure, peak expulsion velocity), it was necessary to develop solutions for other, more relevant regimes. Shajii and Freidberg derived analytic expressions for temperature, pressure, and quench zone propagation velocity for five other regimes, identified as the (1) short coil, low  $\Delta p$ , (2) short coil, high  $\Delta p$ , (3) long coil, low  $\Delta p$ , (4) long coil, high  $\Delta p$ , and (5) thermal hydraulic quenchback (THQB) regimes. In a short coil, the coil length is much shorter than the diffusion length of the quench zone, so that quench propagation is affected by end conditions. In a long coil, the mass of the coil channel is constant and propagation is unaffected by the ends. In the low  $\Delta p$  regime, the pressure rise due to the quench is  $\ll p_0$ , while it is  $\gg p_0$  in the high  $\Delta p$  regime. An actual quench may have a trajectory in quench regime space that traverses two or more of these regimes. We use the nomenclature long coil and short coil here, because the usage has become accepted. However, it should be clarified that we are always discussing the length of a hydraulic channel, which is typically an order of magnitude shorter than a coil length. In all four of the constant mass quench zone solutions, Shajii derived scalings for the temperature and pressure, valid in the high-temperature ( $>20$  K) regimes that are of interest as coil allowables. Assuming that the heat transfer coefficient  $h$  ( $\text{W/m}^2\text{-K}$ ) is large:

$$\Delta T(t) = \frac{A_w \rho_w C_w}{2hP_w} \alpha(\bar{T}) J^2 (1 - e^{-t/\tau_w}) \quad (25)$$

where  $\bar{T}$  is the average of the cable and conduit temperatures ( $\text{K}$ ), and  $\tau_w$  is a characteristic time constant for heat exchange between the cable and conduit ( $\text{s}$ ):

$$\frac{1}{\tau_w} = hP_{\text{wall}} \left( \frac{1}{A_{\text{cable}} \rho_{\text{cable}} C_{\text{cable}}} + \frac{1}{A_{\text{wall}} \rho_{\text{wall}} C_{\text{wall}}} \right) \quad (26)$$

where  $h$  is the wall heat transfer coefficient ( $\text{W/m}^2\text{-K}$ ),  $P_{\text{wall}}$  is the conduit wetted perimeter ( $\text{m}$ ),  $A_{\text{cable}}$  and  $A_{\text{wall}}$  are the cable and wall cross-section areas ( $\text{m}^2$ ),  $\rho_{\text{cable}}$  and  $\rho_{\text{wall}}$  are the mass densities of the cable and wall ( $\text{kg/m}^3$ ), and  $C_{\text{cable}}$  and  $C_{\text{wall}}$  are the specific heats of the cable and wall ( $\text{J/kg-K}$ ), respectively.

**Short Coil Solutions.** The criterion for a short coil solution is that the length of the coil is longer than the length of the quench zone, but less than the thermal diffusion length of the quench:

$$L < \sqrt{\frac{24d_h c_0^2 t_m}{f V_q}} \quad (27)$$

where  $c_0$  is the initial sound speed (m/s),  $d_h$  is the hydraulic diameter (m),  $V_q$  is the quench front velocity (m/s), and  $t_m$  is the time needed to reach the maximum allowable temperature  $t_m \approx T_{\max}/\alpha_0 J^2$ .

**Short Coil, High Pressure.** A short coil quench will be in the high-pressure rise regime ( $\Delta p > p_0$ ) when

$$\frac{\Delta p}{p_0} \approx \frac{R \rho_0 \alpha_0 J^2 L_{IQZ}}{2 p_0 V_q} > 1 \quad (28)$$

where  $R$  is the universal gas constant ( $=8314.3$  J/kg-mole-K,  $\text{He} = 4.003$  kg/mole),  $\rho_0$  is the background helium mass density (kg/m<sup>3</sup>),  $J$  is the current density in the stabilizer (A/m<sup>2</sup>),  $L_{IQZ}$  is the initial length of the quench zone (m),  $p_0$  is the background helium pressure (Pa),  $d_h$  is the hydraulic diameter (m),  $f$  is the friction factor, assumed to be a constant in the range of 0.06–0.08, and  $\alpha_0$  is a diffusion constant of the conduit material (typically  $5\text{--}7 \times 10^{-6}$  m<sup>4</sup>-K/A<sup>2</sup>s), defined as

$$\alpha_0 \approx \min \left[ \frac{A_{cu} \eta_c(T)}{A_c \rho_c C_c(T) + A_w \rho_w C_w(T)} \right] \quad (29)$$

where  $\eta_c$  is the thermal resistivity (W/m-K).

For a short coil in the high-pressure regime, the asymptotic quench velocity, once quench has been well initiated, is

$$V_q = \sqrt[3]{\frac{2dR}{fL_{\text{coil}}} L_{IQZ} \alpha_0 J^2} \quad (30)$$

where  $d$  is the hydraulic diameter (m) and  $f$  is the friction factor. The position of the forward quench front  $X_q$  (m), using the convention that all quenches are symmetrically centered about  $x = 0$ , is

$$X_q = \left[ \left( \frac{L_{IQZ}}{2} \right)^{3/2} + (V_q t)^{3/2} \right]^{2/3} \quad (31)$$

**Short Coil, Low Pressure.** A short coil quench will be in the low-pressure rise regime ( $\Delta p < p_0$ ) when

$$\frac{\Delta p}{p_0} \approx \left( \frac{f \rho_0 L_{\text{coil}}}{4 d p_0} \right) V_q^2 < 1 \quad (32)$$

For a short coil in the low-pressure regime, the quench velocity is

$$V_q = \frac{R \rho_0 \alpha_0 J^2 L_{IQZ}}{2 p_0} \quad (33)$$

The short coil limit also provides a simple analytical expression for the density outside the normal zone as a function of length

and time:

$$\rho(x, t) = \rho_0 + \left( \frac{\rho_0 f V_q^2}{2 d_h c_0^2} \right) \left( \frac{L_{\text{coil}}}{2} - x \right) \quad (34)$$

where  $\rho_0$  is the initial density (kg/m<sup>3</sup>). The density decreases linearly with  $x$ , while the velocity outside the normal zone is a constant versus both space and time.

**Long Coil Solutions.**

**Long Coil, High Pressure.** The long coil solution is defined by the two criteria that the coil length is much greater than the length between diffusion edges  $(L/2)^2 \gg X_D^2(t_m)$ , and that the length of the quench zone is much less than the length between diffusion edges  $[X_q^2 \ll X_D^2(t_m)]$ . The criterion for the long coil solution can then be stated as

$$L_{\text{coil}}^2 \gg \frac{24 d_h c_0^2 t_m}{f V_q} \gg 4 V_q^2 t_m^2 \quad (35)$$

The long coil quench is in the high-pressure rise regime ( $\Delta p > p_0$ ) when

$$\frac{\Delta p}{p_0} \approx \frac{R \rho_0 \alpha_0 J^2 L_{IQZ}}{2 p_0 V_q(t_m)} > 1 \quad (36)$$

The position of the forward quench front in the high-pressure regime is

$$X_q = \left[ (L_{IQZ}/2)^{5/3} + (V_q t)^{5/3} \right]^{3/5} \quad (37)$$

where  $V_q$  is the asymptotic quench propagation velocity (m/s):

$$V_q = 0.613 \left( \frac{2d_h}{f} \right)^{1/5} \left( \frac{R L_{IQZ} \alpha_0 J^2}{c_0} \right)^{2/5} \frac{1}{t^{1/5}} \quad (38)$$

(The multiplier of 0.613 differs from the two published values of 0.766 by a factor of 5/4, removing an ambiguity in interpreting the Shajii equations self-consistently.) The helium velocity (m/s) in the region outside the quench zone is

$$v(x, t) = \frac{0.8 V_q t}{t + \lambda_1^2 (x - L_{IQZ}/2)^2} \quad (39)$$

where  $\lambda_1^2 = \rho_0 0.8 V_q / 3 v_0^2$ . The density profile for the density  $\rho$  in the outer region (kg/m<sup>3</sup>) is given by

$$\begin{aligned} \rho(x, t) &= \rho_0 + \frac{9 v_0^2}{2 K^{3/2}} \left( \cot^{-1} \frac{\xi}{\sqrt{K}} \frac{\sqrt{K} \xi}{K + \xi^2} \right) \\ &\approx 3 v_0^2 \frac{\lambda_2^3 t^2}{[t^{3/4} + \lambda_2^{3/2} (x - L_{IQZ})^{3/2}]^2} \end{aligned} \quad (40)$$

where  $\lambda_2(t) = (3\pi/4)^{1/3} \lambda_1$  and  $v_0$  is

$$v_0 = \sqrt{\frac{2 d_h \rho_0 c_0^2}{f}} \quad (41)$$



Equation (11) implies that the expulsion velocity is given by

$$v(x = L_{\text{coil}}/2, t) \simeq \frac{24d_h c_0^2}{fL_{\text{coil}}^2} t \quad (42)$$

This is the same expulsion velocity as that predicted by Dresner. However, with the exception of the expulsion velocity, Dresner's solutions for temperature, pressure, and quench velocity in Eqs. (16–18) have functional dependencies that are significantly different from Shajii's, even in the long-coil case.

**Long Coil, Low  $\Delta p$  Regime.** In the small  $\Delta p$  regime, the pressure rise remains small in comparison with the initial pressure ( $\Delta p \ll p_0$ ), corresponding to a weak quench in which the helium coolant removes most of the heat generated by a quench. In the small  $\Delta p$  regime, the length of the quench zone is

$$X_q \approx \frac{R\rho_0 L_{IQZ}(\bar{T} + \Delta T - \bar{T}_{t=0+})}{2p_0} + \frac{L_{IQZ}}{2} \quad (43)$$

where  $\bar{T} = (T_{\text{cable}} + T_{\text{wall}})/2$  and  $\Delta T = (T_{\text{cable}} - T_{\text{wall}})/2$ .  $t = 0+$  means the time immediately after the initial quench zone has been established, assuming sudden energy deposition.

The quench velocity is then

$$V_q \approx \frac{R\rho_0 \alpha_0 J^2 L_{IQZ}}{2p_0} \quad (44)$$

The low  $\Delta p$  itself is given by

$$\Delta p(t) = 9v_0^2 c_0^2 t^2 \int_{x_q}^{L_{\text{coil}}/2} \frac{dx}{[(x - L_{IQZ}/2)^2 + X_D^2]^2} \quad (45)$$

where the leading edge of the diffusion front  $X_D$  (m) is

$$X_D^2 = \frac{3v_0^2 t}{\rho_0 V_q} \quad (46)$$

For the combined short coil plus small  $\Delta p$  case, the solution for  $\Delta p$  reduces to Eq. (23). For the combined long coil plus small  $\Delta p$  case, the solution for  $\Delta p$  reduces to

$$\Delta p(t) = 1.36 \left( \frac{f}{2d_h} \right)^{1/2} \rho_0 c_0 t^{1/2} V_q^{3/2} \quad (47)$$

**Universal Scaling Regimes.** Shajii recast the preceding criteria into a universal form that predicts the boundaries between the four quench regimes. Two dimensionless variables  $x$  and  $y$  are defined and all quench regimes are shown as filling four corners of  $x - y$  space. First a dimensionless variable  $\lambda$  and a dimensional variable  $L_0 J^{4/3}_0$  ( $\text{A}^{4/3}/\text{m}^{5/3}$ ) are defined:

$$\lambda = 1.7 \left( \frac{\rho_0 R T_{\text{max}}}{p_0} \right) \left( \frac{c_0^2 \rho_0}{p_0} \right) \quad (48)$$

where  $\rho_0$ ,  $c_0$ , and  $p_0$  are the density ( $\text{kg}/\text{m}^3$ ), sound speed (m/s), and pressure (Pa) of the background helium.  $R$  is the universal gas constant ( $8314.3 \text{ J}/\text{kg}\cdot\text{mole}\cdot\text{K}$ ), and  $T_{\text{max}}(x, t)$  is the maxi-

imum temperature of the quench zone (K).

$$L_0 J_0^{4/3} = \frac{2.6}{R} \sqrt[3]{\frac{p_0^5 d}{f \alpha_0^2 c_0^2 \rho_0^5 T_{\text{max}}}} \quad (49)$$

$\lambda$  and  $L_0 J_0^{4/3}$  tend to be relatively constant over a wide range of time and space for a given design. In order to distinguish better between the widely different quench regimes, these variables are reordered by being normalized to the strongly varying  $J$  ( $\text{A}/\text{m}^2$ ) and  $L_q$  (m):

$$x = \frac{\lambda L_q}{L} \quad (50)$$

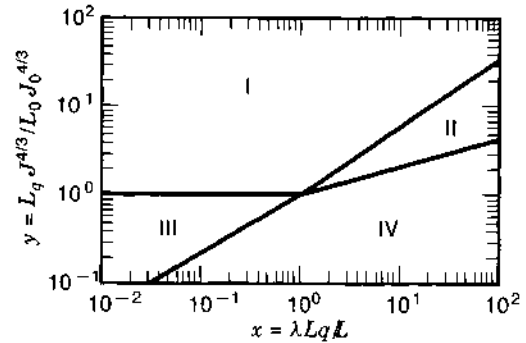
$$y = \frac{L_q J^{4/3}}{L_0 J_0^{4/3}} \quad (51)$$

This now allows the criteria for entry into each of the four quench regimes to be written in the following simple universal form:

| Quench Regime             | Pressure Condition | Length Condition |
|---------------------------|--------------------|------------------|
| Long coil, high pressure  | $y > 1$            | $y > x^{5/6}$    |
| Short coil, high pressure | $y > 1.2x^{1/3}$   | $y < 1.1x^{5/6}$ |
| Long coil, low pressure   | $y < 0.8$          | $y > x^{2/3}$    |
| Short coil, low pressure  | $y < 1.2x^{1/3}$   | $y < x^{2/3}$    |

The four universal scaling regimes are shown in Fig. 23.

**Thermal-Hydraulic Quenchback.** The misnomer thermal-hydraulic quenchback (THQB), which is certainly thermohydraulic, but has nothing to do with quenching backwards, has become sufficiently popular as a term that we won't try to rename it Joule-Thomson quench propagation. It refers to a condition in which compression heating of helium in front of a quench zone leads to rapid propagation of quench in which the thermal/quench wave travels much faster than the mass flow



**Figure 23.** Boundaries in  $x$ - $y$  space defining the four quench regimes: (I) long coil-high pressure rise, (II) short coil-high pressure rise, (III) long coil-low pressure rise, and (IV) short coil-low pressure rise [Shajii, 1995 (44)].

of helium expulsion in front of the quench zone. A key dimensionless parameter in predicting whether there can be a rapid onset of THQB, again introduced by Shajii (44), is the safety margin between current sharing and background temperature  $M$ :

$$M = \left[ \frac{C_h(T_0)}{C_\beta(T_0)} \right] \left( \frac{\rho_0 c_0^2}{\rho_0} \right) \left( \frac{T_{cs} - T_0}{T_0} \right) \quad (52)$$

Typically  $M \sim 2$  to 5 for practical coils. Another parameter  $\beta$  is needed to account for the finite ratio of frictional to compressive heating in the THQB regime:

$$\beta = \frac{C_\beta(T_0) \rho_0 T_0}{\rho_0} \quad (53)$$

$\beta$  is typically of order unity.

THQB cannot exist long in the low pressure rise regime. In the long-coil, high pressure rise regime, the temperature just ahead of the quench front can be written as:

$$T_f(t) \approx T_0 + \frac{1}{2} \left[ \frac{C_\beta(T_0)}{C_h(T_0)} \right] \frac{RT_0 \alpha_0 J^2 L_q}{c_0^2 V_q(t)} \quad (54)$$

This expression includes the effect of compression heating, but not of frictional heating.

The condition for THQB to occur before  $T = T_{\max}$  is

$$y > M^{5/3} \quad (55)$$

For the short coil-low pressure rise regime, the condition for THQB to occur before  $T = T_{\max}$  is

$$y > \beta^{1/3} M^{1/3} \quad (56)$$

The temperature ahead of the quench front is then

$$T_f(t) = T_0 + \left[ \frac{f V_q^3}{2 d C_h(T_0)} \right] t \quad (57)$$

In the short coil-high pressure rise regime, frictional heating must again be included, and the condition for THQB to occur before  $T = T_{\max}$  is

$$y > \frac{x^{1/3}}{2} \left\{ 1 + [1 + 4(x - M)]^{1/2} \right\} \quad (58)$$

where  $\beta$  and other numerical coefficients have been set to unity.

The intersection of the THQB regime with the four universal regimes for a typical case of  $M = 4$  is shown in Fig. 24.

Unique THQB shaded regions can be calculated for every value of  $M$ .

## ELECTRICAL INTEGRITY

A superconductor has no voltage in the quiescent state and may have a very low voltage, during startup and shutdown, depending on the application. In the majority of applications, the superconducting magnet will experience its highest design voltages during a quench dump. It is also during a quench dump that the helium temperature will be highest, helium density in

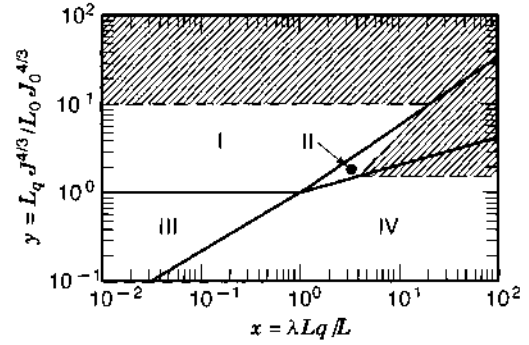


Figure 24. Scaling diagram showing dimensionless quench regime boundaries for  $M = 4$ .

electrical isolators may be decreased by pressure relief valves, and liquid will be boiled into vapor. All of these effects can simultaneously increase electrical fields, while decreasing the ability of helium or insulation voids to withstand the electrical fields. Arcs have developed during operation of real superconducting magnets and are a known cause of failure and life limitation.

## Breakdown in Helium

In a pool-boiling magnet, helium is the primary insulation. Design against helium breakdown must include the winding-pack, joints, supports, vapor-cooled leads, feedthroughs, and room-temperature isolators. In CICC magnets, the problem is restricted to specially designed helium isolators, providing the electrical isolation between the magnets and its grounded helium headers. However, CICC magnets have the special problem of protecting feedthroughs and leads against low-density helium breakdown, in the event of a helium leak. The primary motivation for dry superconducting magnets is to avoid this problem altogether, while accepting a low energy margin against disturbances.

The dielectric strength of liquid helium is comparable to that of air at standard temperature and pressure. Unfortunately, since breakdown accompanies heating due to a normal event and rapidly heats local helium, the actual breakdown strength of liquid helium in a magnet is hard to interpret. It is conservative and probably correct to always consider gaseous helium to be the insulator in a pool boiling magnet.

The breakdown strength of gaseous helium at ambient temperature is only a small fraction of that of air because the electrons can gather kinetic energy from electrical field drift up to the ionization level in the noble gas helium. Paschen's law should hold for gaseous helium at any temperature. At all density-gap products that are well above the minimum of the Paschen curve, the dielectric strength of helium is at least a factor of 10 worse than that of nitrogen or air.

In the high-pressure regime of the Paschen curve, Olivier showed that a direct exponential correlation can predict the voltage breakdown for a broad range of gases, including helium, in a uniform field over a broad range of gap lengths and pressures (45). At any temperature, the breakdown voltage between spherical electrodes is

$$V_{\text{breakdown}} = K \rho^a d^b \quad (59)$$

**Table 4. Paschen Minimum Gaps and Gaps at 20 kV, According to Olivier Equation**

| Pressure (atm) | Temperature (K) | Density (kg/m <sup>3</sup> ) | Gap at Paschen Minimum (mm) | Gap at 20 kV (mm) |
|----------------|-----------------|------------------------------|-----------------------------|-------------------|
| 1.0            | 4.229           | 127, liquid                  | $7.9 \times 10^{-5}$        | 0.133             |
|                |                 | 15.2, vapor                  | $6.6 \times 10^{-4}$        | 1.17              |
| 3.0            | 5.0             | 117                          | $8.5 \times 10^{-5}$        | 0.144             |
| 3.0            | 150             | 0.973                        | 0.011                       | 20.0              |
| $10^{-4}$      | 5.0             | $10^{-3}$                    | 10                          | 20,000            |
| 1.0            | 273             | 0.178                        | 0.056                       | 112               |

where  $K$  is a constant,  $\rho$  is the mass density of helium (moles/liter),  $d$  is the gap (mm), and  $\alpha$  and  $\beta$  are the exponents. According to Olivier, for helium  $\alpha = 0.878$ ,  $\beta = 0.901$ , and  $K = 4310$ . Thus, for helium at STP, where one mole = 22.4 L and the density of helium is  $4 \text{ g}/22.4 \text{ L} = 0.178 \text{ kg/m}^3$ , the general equation for voltage breakdown with the gap  $d$  in m, the breakdown voltage  $V$  in volts, and the density  $\rho$  in  $\text{kg/m}^3$  would be

$$V_{\text{breakdown}} = 645,375 \rho^{0.878} d^{0.901} \quad (60)$$

The density-gap product that produces the minimum of 160 V in helium is  $10^{-5} \text{ kg/m}^3\text{-m}$ . The gaps at the Paschen minimum and the gaps predicted for breakdown at 20 kV by the Olivier equation are shown in Table 4.

A Paschen curve through several sets of experimental data is shown in Fig. 25.

At cryogenic temperatures the breakdown strength of helium in uniform field allows for a straightforward design in pool boiling magnets, where it can be guaranteed that breakdown will occur in the high density-distance regime. Figures 26 and 27 (46) show that with minimum breakdown fields on the order of 10 MV/m, each millimeter of gap should be adequate to protect against 10 kV.

However, breakdown in helium is highly sensitive to electric field nonuniformities and to field polarity. Figure 28 (47)

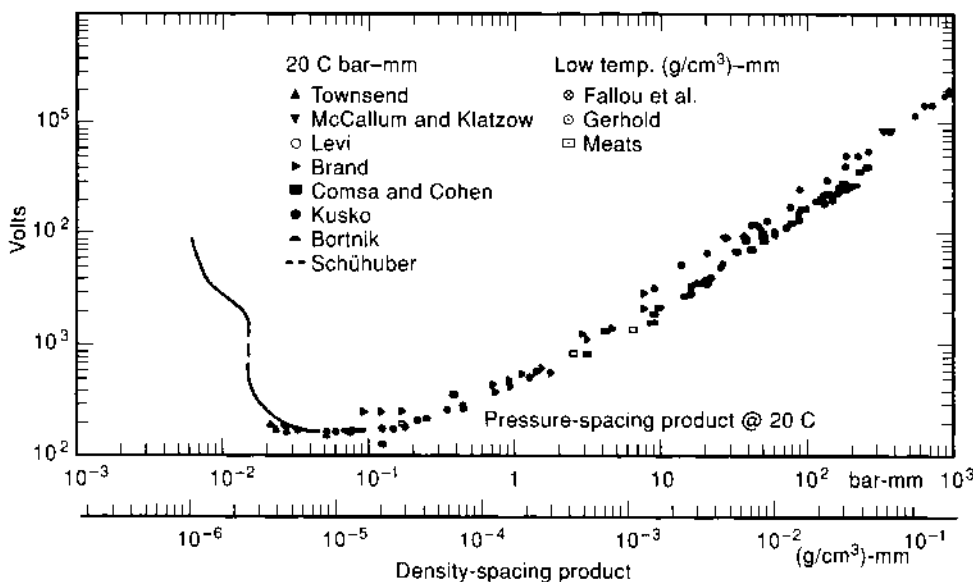
depicts low temperature dc breakdown characteristics for helium gas under nonuniform (point to plane) field conditions at atmospheric pressure (48).

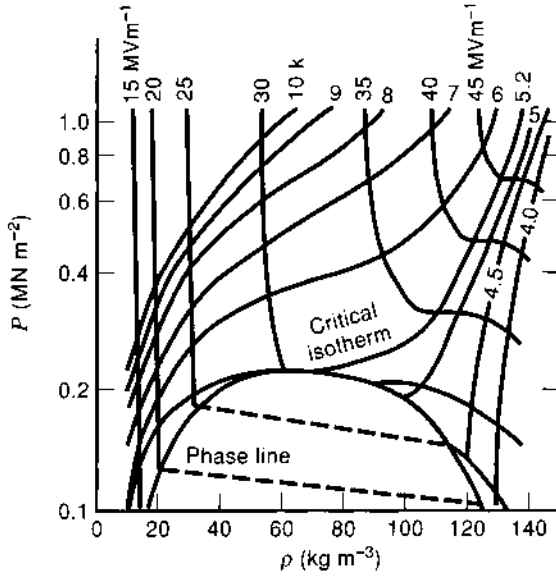
The strong polarity dependence is evident in this figure, suggesting that the design of the electrode configuration to maintain a field as close to uniform as possible is particularly important, when feasible. The breakdown level is also a function of electrode surface material and conditions (49), conditioning, and electrode polarity (50).

A design approach proposed to avoid striking an arc across a helium gap is to use the dielectric strength of helium at the design temperature and atmospheric pressure for a sharp point at a 3 mm gap between the point and the plain (51). At 4 K, the breakdown strength of helium is 60 V/mil (2.36 kV/mm), at 10 K, the breakdown strength is 24 V/mil (944 V/mm), and at 100 K, the breakdown strength is 2.4 V/mil (94 V/mm).

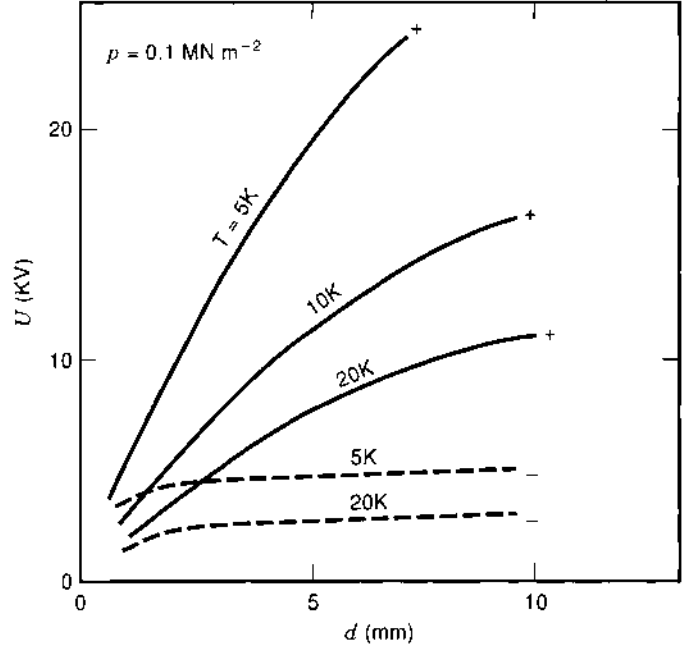
### Electric Field Concentrations

Solid insulations are seldom designed with electrical fields higher than 10–20% of their intrinsic dielectric strength. This is not merely an expression of extreme conservatism toward catastrophic punchthrough and arcing, but also reflects the reality of electrical field concentrations in practical designs. While many solid insulations have dielectric strengths >100 kV/mm, it has been shown that the presence of small voids

**Figure 25.** Paschen curve for helium.



**Figure 26.** Depicts the breakdown strength of helium as a function of pressure and density, under uniform field conditions [Gerhold, 1979 (48)].



**Figure 28.** Helium nonuniform field breakdown voltages in low temperature helium gas at atmospheric pressure [Gerhold, 1979 (48)].

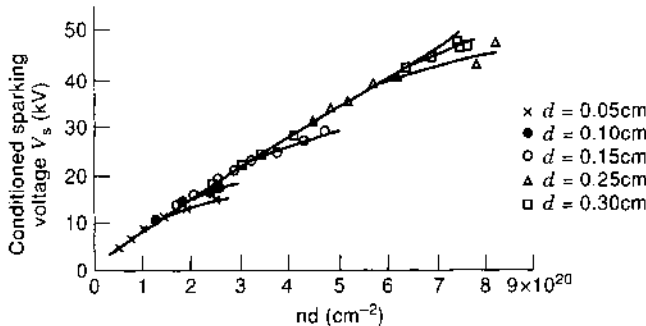
in a solid will cause the inception of particle discharges in low temperature helium at electric fields of 10 kV/mm to 20 kV/mm (52). There is some evidence that the minimum breakdown voltage in a helium void may saturate at low density and that there is no left-hand side of the Paschen curve in voids. Measurements by Hiley and Dhariwal (53) appear to saturate at 4.5 kV/mm in polyethylene at zero density for a cavity depth of 0.2 mm (900 V). In epoxy resin, they appear to saturate at 1.2 kV/mm for a cavity depth of 0.2 mm (240 V).

Partial discharges in solid insulation voids are caused by high electric fields in the voids. These will not destroy the insulation, if they are infrequent, but can cause erosion of organic insulation in a pulsed application.

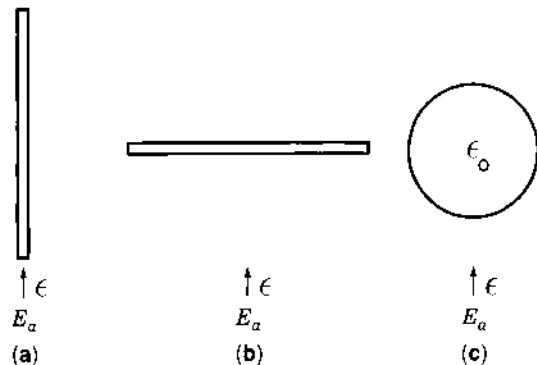
The electric field in a void tends to be higher than that of the solid insulation itself because the dielectric constant of solid insulations is always higher than that of free space. Further electric field multipliers are caused by shape factors in the void, where breakdown can be further enhanced by nonuniform electric field within the void itself. Analytic solutions are available for the electric field concentrations in planar, spherical, and

cylindrical voids. For other shapes, numerical techniques are available to calculate the electric field concentrations for both 2-D and 3-D shapes. Engineering rules of thumb also exist for the most likely void shapes. Zahn (54) developed the following derivation of the electric field concentrations in spheres and cylinders.

**Void in Solid Insulation.** The simplest model assumes that a void region exists within a uniform dielectric that has an essentially uniform electric field in the vicinity of the void. We consider three simple cases shown in Fig. 29 of uniform electric field incident on (1) a long, thin void, either planar or cylindrical, with the long axis in the direction of the electric field; (2) a long thin void, either planar or cylindrical, with the long axis perpendicular to the electric field; and (3) a spherical void. Cases (1) and (2) form easily at an interface between dissimilar materials, while case (3) can arise due to gassing. If the void region is air at ambient temperature and pressure, the electrical breakdown strength is  $E_b \sim 3$  kV/mm. Fields above  $E_b$



**Figure 27.** Depicts low temperature (<10 K) dc breakdown characteristics for helium gas under uniform field conditions [Meek, 1979 (47)].



**Figure 29.** Three void orientations versus electric field in dielectric medium.

will result in spark discharges. The maximum external electric field outside the void can then be calculated that would keep the voidage electric field below  $E_b$ .

**Long Thin Void Along Electric Field.** Because the tangential component of an electric field is continuous across an interface, the electric field

$$\bar{E}_v = \bar{E}_a \quad (61)$$

This result is valid for planar and cylindrical voids.

#### Long Thin Void Perpendicular to Electric Field.

**Planar Void.** For highly insulating dielectrics with dielectric relaxation times ( $\tau = \epsilon/\sigma$ ) much greater than the time scales of a quench, the normal component of the electric displacement field,  $\bar{D} = \epsilon\bar{E}$  must be continuous across an interface. Thus

$$\epsilon\bar{E}_v = \epsilon\bar{E}_a \rightarrow \epsilon\bar{E}_v = \frac{\epsilon}{\epsilon_0}\bar{E}_a \quad (62)$$

The higher the dielectric permittivity of the insulation surrounding the void, the lower the applied electric field must be to keep  $\bar{E}_v < E_b$ .

**Cylindrical Void.** In the absence of any volume charge in a uniform permittivity dielectric, Laplace's equation can be solved in cylindrical coordinates for the electrical field inside and outside the cylindrical void:

$$\begin{aligned} \bar{E} &= 2 \frac{\epsilon}{\epsilon + \epsilon_0} \bar{E}_a \bar{i}_x & 0 < r < R \\ \bar{E} &= E_a \left[ 1 + \frac{R^2 (\epsilon_0 - \epsilon)}{r^2 (\epsilon_0 + \epsilon)} \cos \varphi \bar{i}_r \right. \\ &\quad \left. - \left( 1 - \frac{R^2 (\epsilon_0 - \epsilon)}{r^2 (\epsilon_0 + \epsilon)} \sin \varphi \bar{i}_\varphi \right) \right] & r > R \end{aligned} \quad (63)$$

The internal electric field is purely  $x$  directed, while the external electric field has the applied uniform electric field plus a line-dipole field. To avoid breakdown in the cylindrical void

$$2 \frac{\epsilon}{\epsilon + \epsilon_0} E_a < E_b \quad (64)$$

**Spherical Void.** The general form of solution to Laplace's equation in spherical coordinates for a uniform  $z$  directed electric field is

$$\varphi = \left( Ar + \frac{B}{r^2} \right) \cos \theta \quad (65)$$

The total electric field inside and outside the cylindrical void is

$$\begin{aligned} \bar{E} &= \frac{3\epsilon}{2\epsilon + \epsilon_0} E_a \bar{i}_z & 0 < r < R \\ \bar{E} &= E_a \left[ \left( 1 + 2 \frac{R^3 (\epsilon_0 - \epsilon)}{r^3 (2\epsilon + \epsilon_0)} \right) \cos \theta \bar{i}_r \right. \\ &\quad \left. - \left( 1 - \frac{R^3 (\epsilon_0 - \epsilon)}{r^3 (2\epsilon + \epsilon_0)} \right) \sin \theta \bar{i}_\theta \right] & r > R \end{aligned} \quad (66)$$

**Table 5. Dielectric Constant  $\epsilon$**

| Material                      | Dielectric Constant ( $\epsilon_r/\epsilon_0$ ) |
|-------------------------------|---|
| Kapton polyimide film 120CI-1 | 3.5   |
| Kapton polyimide film 135RCI  | 3.8   |
| G10-CR                        | 4.9–5.0   |
| G11-CR1                       | 5.1–5.2   |

The internal electric field is purely  $z$  directed, while the external electric field has the applied uniform electric field plus a point dipole field. To avoid breakdown in the spherical void:

$$\frac{3\epsilon}{2\epsilon + \epsilon_0} E_a < E_b \quad (67)$$

Dielectric constants of some widely-used solid insulations are listed in Table 5.

The electric field multipliers for other commonly found void shapes are listed in Table 6.

The multipliers for laps and voids in corners are typical values, based on numerical analysis.

A conservative way to look at the design rules for solid insulation would be to design so that partial discharges in the solid insulation were impossible, irrespective of the partial pressure in the insulation voids. As an example, in a superconducting magnet, the worst gas and the most likely gas to have a partial pressure is helium, whose Paschen minimum is 160 V. Adopting 2 kV/mm as a typical allowable for glass-epoxy and a geometry/dielectric constant mismatch multiplier of 2, this design rule would then put the burden on the mechanical design to assure that the largest possible void in the solid insulation would be

$$d_{\max, \text{allowable}} \leq \frac{V_{\text{Paschen min}}}{E_{\max, \text{allowable}} R_{\text{mult}}} = \frac{160 \text{ V}}{2 \frac{\text{kV}}{\text{mm}} \times 2} = 0.040 \text{ mm} \quad (68)$$

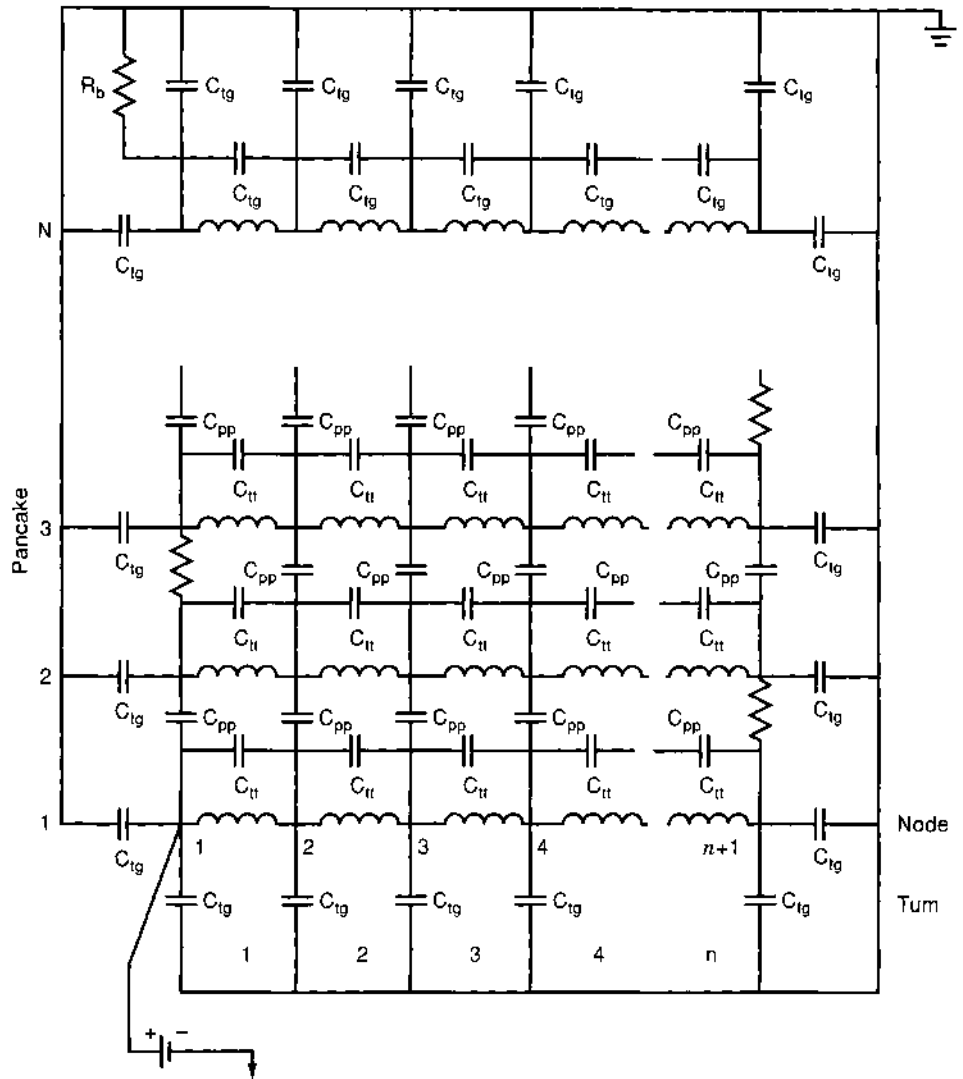
For example, if a conduit with a slip-plane wrap bowed in 10  $\mu\text{m}$ , this would be acceptable, since breakdown of helium would still be impossible.

#### Stray Coil Capacitance

The turn-turn and layer-layer voltages across a winding pack are not exactly equal, even in the absence of any resistance. The turns in a winding pack don't all link exactly the same amount of flux, particularly in multicoil systems. The peak-average ratios in coil systems we have investigated were in the range of 1.1–1.3:1. A more serious, but avoidable, problem comes from the possibility of overvoltages due to distributed capacity charging of the winding pack. There is

**Table 6. Electric Field Multipliers for  $\epsilon = 3$  Insulation System**

| Shape   | Multiplier  |
|---|-------------|
| Infinite plane or cylinder, parallel to field | 1           |
| Infinite plane, transverse to field           | 3           |
| Infinite cylinder, transverse to field        | 1.5         |
| Sphere  | 9/7 = 1.286 |
| Cusp/crescent (debonding at a rounded corner) | 1.6–1.7     |
| Triangle (e.g., epoxy/kapton lapping)         | 1.6–1.8     |



**Figure 30.** Distributed capacitance/inductance network model of a winding pack and ground insulation.

a distributed capacitance between every turn of the winding pack and to ground, as illustrated by the circuit model, shown in Fig. 30. Mutual capacitance to nonadjacent turns is neglected.

The peak local electrical field is higher than what would be predicted considering only the coil resistance because of stray coil capacitance in the insulation. The voltage enhancement factor is a function of  $\alpha$ :

$$\alpha = \sqrt{\frac{C_g}{C_s}} \tag{69}$$

where  $C_g$  is the shunt capacitance to ground and  $C_s$  is the series capacitance. For fast rise times and large values of  $\alpha$ , half of the terminal voltage can appear between the first two turns of the coil. Measurements on the POLO coil (55) and analyses at Karlsruhe (KfK) and M.I.T. showed that large voltage enhancements could be avoided, if the voltage rise time were long enough. This is a tradeoff, since switch losses or counterpulse circuit size is proportional to the voltage rise time. Optimization of the cost/performance trade can be done for a specific design, using commercial circuit codes. However, since the ringing time is primarily a function of conductor size only, analysis indicates

that it is safe over a broad range to limit the voltage rise time to  $>100 \mu s$ , if the conductor is  $<5 \text{ cm}^2$  square.

**Tracking Along an Insulating Surface**

Tracking is the leakage current due to the formation of a conducting path across the surface of an insulation. In most cases, the conduction results from degradation of an organic insulation. The conducting film is usually moisture from the atmosphere absorbed by some form of contamination, such as dust. Tracking does not depend upon Paschen breakdown and can occur at well below 100 V in air, while a gaseous discharge in air cannot exist below 380 V. Degradation of the insulation is caused by heat from tracking, which either carbonizes or volatilizes the insulation. This implies that insulations in helium or vacuum at liquid helium temperatures should be more immune to tracking than magnets in air, because of the absence of chemical reactions. Carbonization results in a permanent extension of the electrodes and usually takes the form of a dendritic growth; but erosion of the insulation also occurs.

Configurations for which tracking is particularly important are those where a high voltage conductor emerges from an insulated lead. Generally a conducting electrostatic shield or

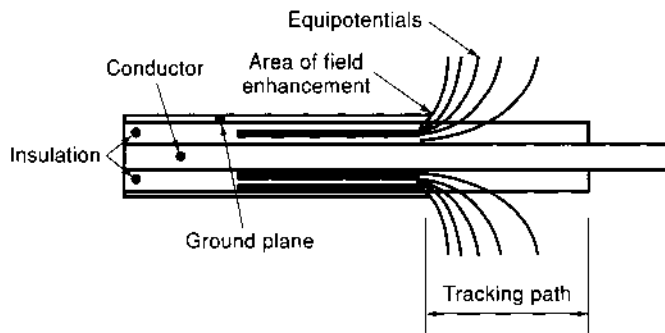


Figure 31. Tracking at high voltage lead termination.

ground plane will surround the lead insulation, but will be terminated at some point, leaving a tracking path along the surface of the lead insulation from the end of the ground plane to the conductor itself. The electric field is enhanced at the end of the electrostatic shield in this configuration, and can lead to local ionization which effectively extends the ground electrode, due to space charge, and reduces the useful length of the tracking path, such that the discharge advances along the surface toward the lead, with eventual flashover. Figure 31 illustrates tracking at a high voltage lead termination.

Common practice is to utilize a stress cone (a specially shaped electrode) or a stress grading material or coating on the surface of the lead insulation just after the termination of the conducting electrostatic shield. The effect of the stress grading is to reduce the maximum field gradient based on the non-linear voltage versus current characteristic of the stress grading material. Silicon carbide is a commonly used material for stress grading, and is available in tape and paint binders. Heat shrinkable stress grading tubing is also available, and is normally used in the termination of high voltage cables. This material could offer a solution to the problem of temporary stress grading during test, when the termination of high voltage leads is not in its final configuration.

Yet another technique is to cover the entire termination in an insulating jacket. The effect is to contain the areas of highest stress in a solid dielectric, and to increase the tracking path length that must be taken by a discharge which would directly connect the high voltage electrode with the ground electrode.

To the extent that the high voltage conductors can be fully insulated and enclosed by an electrostatic shield, the tracking problem in the near vacuum environment can be avoided.

However, the high voltage conductors such as the main high current leads and any instrumentation (e.g., voltage tap) conductors must eventually emerge from the insulation.

The helium supply lines which must bridge the gap between the high voltage conductors and ground present another potential tracking path. Considering a configuration consisting of a metallic pipe emerging from the main conductor which delivers the coolant to an insulating tube, which in turn connects to a grounded metallic pipe associated with the manifold. Tracking on the outer surface of the insulating tube could be a problem, particularly in the cryostat vacuum environment which can be compromised by helium leaks. This problem can be solved by encapsulating the region of transition from high voltage, through the insulator, to ground, in a solid dielectric contained within an electrostatic shield. In this case the path which would otherwise be available for tracking, and the entire region subject to electric stress, can be fully contained in solid insulation, as illustrated in Fig. 32.

**Experimental Basis for Tracking Allowables.** Migliori, Schermer, and Henke (56) measured dielectric tracking fields in liquid helium for four gaps in a G-10 circuit board. The minimum tracking strength that they measured was 13.7 kV/mm at 0.48 mm. They reported that the cleanliness of the sample, that is, whether it was abraded with copper or wet-wound with metal inclusions, made little difference in the tracking strength. They also reported that the breakdown field declined as  $d^{-0.25}$  and that the tracking strength was that of helium vapor, rather than that of liquid helium. This is a somewhat misleading conclusion, because the tracking strengths that they measured were similar to those measured earlier by Haarmon and Williamson (57) in liquid helium, and orders of magnitude superior to Haarmon and Williamson's measurements in gaseous helium. The earlier experiment reported nearly identical tracking breakdown fields for paper phenolic, polyethylene, nylon, and teflon. Their measurements of the tracking strength for series of five 60 kV, 500 ms, half sine-wave pulses with descending gap are shown in Table 7.

For design purposes, the relevant numbers are those for gaseous helium at 4 K and 293 K. Fortunately, the temperature sensitivity is rather small, varying only a factor of two at most over a temperature range of 75:1. A factor of four safety margin would imply that tracking should be limited to 100 V/mm (2.5 V/mil). Designs have been identified with tracking fields in the range of 60 V/mm to 200 V/mm.

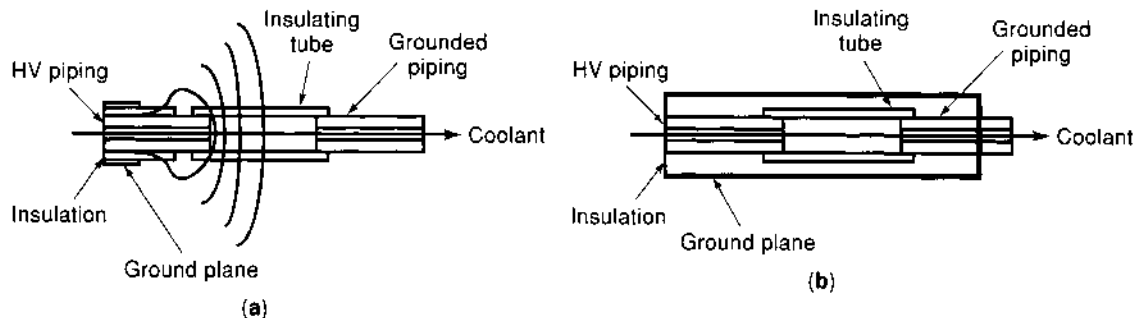


Figure 32. Avoidance of tracking with full-insulated coolant connections. (a) Tracking at coolant connection. (b) Fully insulated coolant connection.

**Table 7. Tracking Strength for 60 kV Pulse Waveforms (kV/cm)**

| Material       | Pressure (atm) | LHe 4 K | GHe 4 K | GHe 293 K | LN2 77 K | GN2 77 K | GN2 293 K |
|----------------|----------------|---------|---------|-----------|----------|----------|-----------|
| Mylar          | 0.9            | 79      | <3.6    | 2.7       | 158      | 39       | 4.7       |
|                | 1.6            | 79      |         | 2.8       | 189      | 79       | 6.8       |
| Paper phenolic | 0.9            | 95      | 4.3     | 2.3       | 158      | 32       | 4.3       |
|                | 1.6            | 118     | 4.3     | 3.3       | 189      | 39       | 7.9       |
| Polyethylene   | 0.9            | 118     | 3.8     | 2.6       | 158      | 30       | 5.9       |
|                | 1.6            | 158     | 3.9     | 2.7       | 236      | 59       | 7.9       |
| Nylon          | 0.9            | 79      | 3.9     | 2.0       | 158      | 47       | 4.7       |
|                | 1.6            | 118     | 3.9     | 2.0       | 236      | 59       | 7.9       |
| Teflon         | 0.9            | 95      |         | 2.8       | 79       | 43       | 4.7       |
|                | 1.6            | 158     |         | 3.1       | 158      | 53       | 7.9       |
| Permalin       | 0.9            | 68      |         | 2.2       | 118      | 30       | 4.7       |
|                | 1.6            | 79      |         | 2.8       | 158      | 68       | 7.9       |

### Electrical Design Allowables

For design purposes, typical electrical design limits in a cryogenic environment are summarized in Table 8.

### Breakdown Due to Helium Leaks

High voltages can be maintained over short distances in a good vacuum. Typically, in the cryostat surrounding CICC and potted magnets, or the vacuum tank surrounding pool-boiling magnets, a vacuum of  $10^{-6}$  to  $10^{-5}$  torr can be maintained. In vacuum systems, where residual gas density is low, the breakdown voltage is to the left of the Paschen minimum. Therefore, electrical integrity within the cryostat can be compromised by helium leaks into the vacuum space. Helium leaks can be caused by diffusion through welds in a conduit, helium feed stubs, and feedthrough seals. If there is a gas leak, the gas density increases and the breakdown voltage is lowered toward the minimum, which could lead to a spark discharge. This is the usual situation in cryostats and junction boxes, where leaks from helium coolant lines or extraction of sensor wires are difficult to suppress completely. Irmisch has confirmed low voltage breakdown in helium gas at cryogenic temperatures and low pressures (58), simulating the effect of a helium leak into a vacuum space with a 10 mm gap. At 6 K, the Paschen minimum was 240 V, instead of the classic minimum of 160 V.

Leaks are suppressed by the avoidance of nonwelded seals, design layouts that permit field repair of leaks, and vigilant leak detection at room temperature and at cryogenic temperatures. Some rules of thumb for leak rate testing are that the leak rate at 1 atmosphere and room temperature is approximately 1000 times better than the leak rate at 10 atmospheres and 4 K. A state-of-the-art measurement, using mass spectrometry,

might test helium stubs down to leak rates of  $10^{-10}$  torr-L/s at room temperature, corresponding to  $\sim 10^{-7}$  torr-L/s at 4 K.

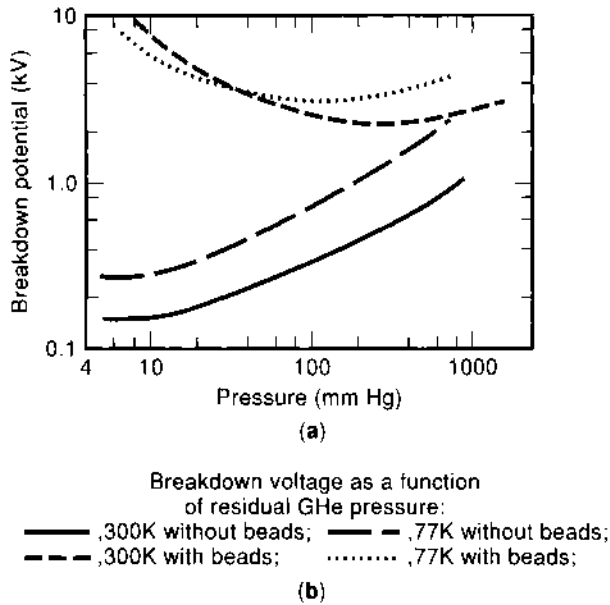
It is generally acknowledged that quality assurance alone won't guarantee an absence of electrical discharges in vacuum, because of the difficulty of guaranteeing zero leaks and because of the deleterious effects on breakdown of modest transverse magnetic field. Three design ideas have been proposed that can decrease the probability of discharges by several orders of magnitude: (1) grading of the insulating ground planes, (2) the use of a guard vacuum, and (3) the use of insulating beads.

1. The use of graded ground planes around all insulating surfaces has the greatest generality as a design concept. If all metallic surfaces that are facing each other with high potential differences are insulated, then there can't be an arc between those surfaces acting as electrodes; although there could still be partial discharges that would gradually degrade the insulation. However, if each insulating surface or wrap has a ground plane, the resistance of the ground plane can be graded, so that the insulating or ground plane surfaces facing each other never have a potential difference greater than 160 V. In this case, no combination of pulsed fields, capacitive charges, or helium pressures can cause a breakdown.
2. The use of a guard vacuum around a winding pack can be very effective, where space constraints aren't significant, as in the Mirror Fusion Test Facility-B (MFTF-B) coils (59). The secondary vacuum space greatly decreases the probability of a leak into the cryostat by requiring two series independent leaks in the coil case and the guard vacuum shell. It can further lessen the possibility of operational problems by the use of independent on-line leak detection and/or differential vacuum pumping. In MFTF-B, it also had a fringe benefit of being used as an auxiliary helium channel to reduce thermal stresses in the winding during cooldown.
3. Fast and Hart (60) have shown that filling an evacuated box for power feedthroughs box with glass beads increased the Paschen minimum voltage and increased the pressure at the Paschen minimum, as shown in Fig. 33. Testing in helium as a function of pressure from 1 Pa to 0.13 MPa ( $10^{-5}$  to 1 atm) at room temperature, 77 K, and 4.2 K in helium, they found that the minimum breakdown voltage across a 4.8 mm gap was increased by more than a factor of 10 at all three temperatures. Insulating beads

**Table 8. Typical Electrical Insulation Sizing Rules**

| Type  | Units | Design Criterion |
|---|-------|------------------|
| Flashover along insulating spacer in helium | V/mil | 5-7              |
|   | V/mm  | 100-300          |
| Tracking along insulating wrap              | V/mil | 1.5-5            |
|   | V/mm  | 60-200           |
| Breakdown through epoxy                     | V/mil | 50-150           |
|   | kV/mm | 1.5-6            |
| Breakdown through kapton wrap               | V/mil | 125-500          |
|   | kV/mm | 5-20             |





**Figure 33.** Paschen curve for helium. (a) Measured in uniform gap and (b) measured in gap with and without glass bead particles [Fast and Hart (60)].

could be used to raise the Paschen minimum by an order of magnitude, either in evacuated spaces, such as junction boxes, or in high pressure chambers, such as helium isolators with filters.

## CASE HISTORIES

The majority of actual failures to protect superconducting magnets have had little relation to the design principles described in this chapter. Errors in operation have accounted for more failures than errors in fabrication, and errors in design have been less important than either. Several studies and workshops have addressed the problems of real-life magnet protection (61).

Operational errors causing magnet component failures have included:

- Bypassing protective circuits or not replacing them when they fail
- Wiring magnet or instrumentation leads backwards
- Interpreting anomalous or changed sensor readings as indicators of interesting experimental phenomena, instead of burnout
- Poorly written, long and boring operating instruction manuals, with incorrect or ambiguous instructions; the ignoring of instructions
- Absence of routine inspection and replacement of components with finite lifetimes
- Energization of coils in an incorrect sequence (the classic error is energizing an outer solenoid and crushing an inner solenoid with no strength in radial compression through induced diamagnetic currents)
- Increasing current when critical properties proved to be better than expected, leading to burnout of vapor-cooled lead

- Unattended dewar operation leading to ice blocks, causing overpressure in the dewar
- Removal of a rubber stopper in a dewar, leading to inadequate vapor flow to the leads; failure to monitor lead gas flow or voltage drop

Manufacturing QA problems that led to coil or prototype failure included:

- Helium leaks at fittings
- Voids at the conductor/monolith interface leading to mechanical rupture of the conductor
- Small metallic slivers being sprung from a conduit by the tube mill
- Metallic inclusions in an open, pool-boiling magnet causing turn-turn shorts and arcing
- Insulation tearing during winding, especially of kapton
- Joint installation without solder
- Weld contamination due to contamination by injected polyurethane spacers
- Turn-turn short due to mechanical abrasion of insulation during transportation and installation
- Vacuum leaks through mechanical (i.e., nonwelded) seals
- Exposure of Incoloy conduit to oxidants during heat treatment, leading to stress aggravated grain boundary oxidation (SAGBO) cracking

Reliability problems caused by inadequate redundancy in the design of the protection circuits have included:

- Dump circuit interrupters failing to clear
- Power supplies failing to shut down on receiving a correct trigger signal from the protection circuit
- False positive dump signals leading to arcing
- Programming error in a programmable logic controller leading to deliberately applying overcurrent to the coil. Same problem with overvoltage
- Arcing in a dump resistor causing a dump that was too slow

Structural design errors have included:

- An overly complex load path causing structural failure of a brittle, aluminum primary structure
- Inadequate mechanical protection of sensor insulation
- Inadequate mechanical support of leads, causing lead shorting or arcing
- Peeling of solder lap joints from ends, neglecting strain incompatibility

## SUMMARY

Superconducting magnet protection during a quench is an important but special case of the larger problem of magnet protection against structural, electrical, and thermal failures. The usual method of guaranteeing protection is to build conservative magnets with more structure, stabilizer, and superconductor than they need. The techniques of providing adequate

conservatism in the design of magnet protection circuits, quench detection, protection criteria, quench propagation modeling, and electrical integrity have already been described. However, experience with magnet failures suggests that new design approaches may be more successful in the future.

In most cases, redundancy should be favored, where possible, as a method of guaranteeing reliability, over just adding more material, because of its ability to multiply low probabilities of failure. Redundant techniques include the use of in series, independent interrupters; paralleled independent quench detectors and paralleled independent cowound heaters. It also includes the use of leak-tight welds and guard vacuums, insulation systems with no voids below the Paschen minimum and long discharge or tracking life, and simultaneous use of signal-noise improvement techniques, such as selecting the best spot in the cable for a cowound sensor, differencing, and filtering the optimized signals.

Given the certainty that data processing will continue to improve more rapidly than magnet technology, future magnets are likely to concentrate on using redundancy and intelligent signal processing in order to improve performance and reliability simultaneously. Smart interlocks should check for wiring errors, while simulators check and calibrate sensors. Reviewed projects must adopt more stringent review standards, insisting on a complete design of the protection system, not only of the magnets, but also the leads, bus, and instrumentation feedthroughs. The goals of these improvements would be to achieve orders of magnitude decreases in the probability of failure to detect a quench rapidly, to successfully interrupt current in a magnet, and to simulate and adequately predict the ability of quench propagation to absorb magnet energy without damage.

## BIBLIOGRAPHY

1. C. Neumeyer et al., Quench protection circuits for superconducting magnets, *Proc. IEEE Symp. Fusion Energy*, U. Illinois, p. 1275, Sept 30–Oct 5, 1995.
2. P. Dokopoulos and K. Kriechbaum, DC circuit breaker for 73 kA, 24 kV. *Electrotech. Zeitschrift A*, **97** (8): translated by H. Vogel, Los Alamos Report LASL-TR-77-27, 1976.
3. S. Yokota, Poloidal field power supply using vacuum circuit breaker, *IAEA Second Large Tokamak Meeting*, Princeton, New Jersey, Nov. 1976.
4. V. Kuchinski et al., High-power fast switches, *IPEC-Yokohama*, 899, 1995.
5. Argonne National Laboratory Superconducting Magnet Group, Final design of a superconducting MHD magnet for the CFFF, Argonne National Laboratory, ANL-MHD-79-6, pp. 250–251, March, 1979.
6. C. Neumeyer, Liquid rheostat dump resistors, Princeton Plasma Physics Laboratory 40-940217-PPPL/CNeumeyer-01, Feb 17, 1994.
7. T. Ishigohka and Y. Kushiuro, Quench protection of superconducting magnet using ZnO arrester, *Cryogenics*, **31**: 562, 1991.
8. A. Dudarev, V. E. Keilin, and Y. Kuroedov, Quench protection of very large superconducting magnets, *IEEE Trans. Appl. Supercond.*, **5**: 226–229, 1995.
9. S. M. Schoenung et al., Liquid helium dump concept for a large scale superconducting magnetic energy storage plant, *Adv. Cryo Eng.*, **31**: 1986.
10. J. F. Kaerner et al., A protection system for small high power SMES with power semiconductors working at cryogenic temperatures, *IEEE Trans. on Appl. Supercond.*, **5**: June 1995.
11. O. P. Anashkin et al., Superconducting magnet protection system, *Adv in Cryogenic Eng.*, **37**: Part A, 339, Plenum Press, 1992.
12. J. Purcell et al., The superconducting magnet system for the 12-foot Bubble Chamber. Argonne National Laboratory Report ANL/HEP6813, 1968.
13. D. F. Sutter et al., Electrical protection of superconducting magnet systems, Fermi National Laboratory Technical Memorandum TM-559, March 1975.
14. T. Ogitsu et al., Quench observation using quench antennas on RHIC IR quadrupole magnets. *IEEE Trans. Mag.*, **32**: 3098, July 1996.
15. S. J. St. Lorant communication reported in J. Allinger, G. Danby, S. Y. Hsieh, J. Kean, J. Powell and A. Prodel, Fusion magnet safety studies program superconducting magnet protection system and failure: Interim Report. Brookhaven National Laboratory, BNL 20787, Nov 1975.
16. M. Darweschsad, The POLO coil, a prototype tokamak poloidal field coil, design features and test results. MT-14, Tampere, Finland; June 1995.
17. J. H. Schultz, Feasibility of the TPX voltage sensor quench detection system. M.I.T. Plasma Science and Fusion Center Report PSFC/RR-97-3, March 5, 1997.
18. N. Mitchell et al., ITER No: N11 R1 03 04-09-07 W1, Magnet design description: Appendix A: conductor design, April 14, 1995.
19. I. Zatz et al., TPX No: 94-950117-PPPL-IZatz-01, TPX structural and cryogenic design criteria, Rev. 1, Jan 17, 1995.
20. J. H. Schultz, ITER No ITER/US/96/EV-MAG/J.H.SCHULTZ/5.31/-1 J. H. Schultz, P. W. Wang, and S. Smith, Interpretation of voltage sensor data in the QUELL experiment, May 31, 1996.
21. M. M. Steeves et al., The US Demonstration Poloidal Coil. *IEEE Trans. Mag.*, **27**: 2369–2372, 1991.
22. Nicolai N. Martovetsky and Michael R. Chaplin, Normal-zone detection in tokamak superconducting magnets with co-wound voltage sensors. *IEEE Magnet Technol. Conf.*, Tampere, Finland, July, 1995.
23. P. W. Wang, TPX Memo: 13-930719-MIT-PWANG-01, TF voltage tap noise due to plasma disruption. July 19, 1993.
24. H. T. Yeh, J. S. Goddard, and S. Shen, Inductive voltage compensation in superconducting magnet systems, *8th Symp. Eng. Probs. Fus. Res.* 1802, 1979.
25. S. Smith and S. Ezekiel, TPX No:1314-941130-MIT/SSmith-01, Improved fiberoptic quench detection at the Francis Bitter National Magnet Laboratory. Rev. 1, Nov 30, 1994.
26. J. H. Schultz and S. Smith, TPX Memo 1314-950815-MIT-JSchultz-01, Feasibility of the fiber optic temperature-sensor quench detection system. Aug 15, 1995.
27. S. P. Smith, TPX No:1314-950424-MIT-SSmith-01, Report on fiberoptic strain sensitivity reduction experiments at the Francis Bitter National Magnet Laboratory on 3/16/1995, April 24, 1995.
28. K. Koepke, P. Martin, and M. Kuchnir, Doubler system quench detection threshold. *IEEE Trans. Mag.*, **MAG-19**: 696, May 1983.
29. Y. Iwasa and M. Sinclair, Protection of large superconducting magnets: maximum permissible undetected quench voltage. *Cryogenics*, Dec. 1980.
30. Y. Iwasa, *Case Studies in Superconducting Magnets: Design and Operational Issues*. New York: Plenum Press, 1994.
31. Y. Iwasa, Design and operational issues for 77 K superconducting magnets, *IEEE Trans. Mag.*, **MG-24**: 1211, 1988; Y. Iwasa

- et al., Stability and quenching in high-temperature superconductors, *IEEE Trans. Appl. Supercond.*, **5**: 389, 1995.
32. C. H. Joshi and Y. Iwasa, Prediction of current decay and terminal voltages in adiabatic superconducting magnets, *Cryogenics*, **29**: 157, 1989.
  33. L. R. Turner, Safety of superconducting fusion magnets: twelve problem areas, Argonne National Laboratory Report ANL/FPP/TM-121, May 1979.
  34. R. P. Krause and E. H. Christensen, Calculation of quench pressures in pool boiling magnets, *IEEE Trans. Mag.*, 1979.
  35. J. Powell, ed., Aspects of safety and reliability for fusion magnet systems. Brookhaven National Laboratory, BNL 50542, 72-75, Jan 1976.
  36. J. R. Miller, J. W. Lue, and L. Dresner, *IEEE Trans. Mag.*, **MAG-13** (24): 1977.
  37. Z. J. J. Stekly, *Adv. Cryo. Eng.*, **8** (ed.) Timmerhaus K.D. New York, Plenum Press, 1963, 585.
  38. L. Bottura and O. C. Zienkiewicz, Quench analysis of large superconducting magnets, *Cryogenics*, **32** (7): 659, 1992; L. Bottura, A numerical model for the analysis of the ITER CICC's, *J. Comp. Phys.*, **125** (26): 1996; L. Bottura, Numerical aspects in the simulation of thermohydraulic transients in CICC's. *J. Fusion Eng.*, **14** (1): 13, 1995.
  39. L. Dresner, Protection considerations for force-cooled superconductors, *Proc. 11th Symp. Fusion Eng.*, **II**: 1218-1222, Austin TX; Nov. 18-22, 1985.
  40. L. Dresner, Thermal expulsion of helium from a quenching cable-in-conduit conductor, *Proc. 9th Symp. Eng. Prob. Fusion Res.*, 618, Chicago, 1981; reformulated in L. Bottura, Quench propagation and protection in cable-in-conduit superconductors. Encyclopedia of Applied Superconductivity, Comett 4090 Ch, 1996.
  41. Y. Wachi et al., Investigation of the pressure rise during the quench of a force cooled superconducting coil, *IEEE Trans. Mag.*, **25**: 1500-1503, 1989.
  42. A. Shajii and J. P. Freidberg, Quench in superconducting magnets. II. Analytic solution, *J. Appl. Phys.*, **76** (5): 1 Sept 1994; L. Dresner, *IEEE Trans. Mag.*, **25**: 1710, 1989; L. Dresner, *Proc 11th Symp Fus Eng, IEEE*, 1218, **2**: 1986.
  43. A. Shajii and J. P. Freidberg, *J. Appl. Phys.*, **76** (5): 1 Sept 1994.
  44. A. Shajii, Universal scaling laws for quench and thermal hydraulic quenchback in CICC coils, *IEEE Trans. Appl. Supercond.*, **5**: 477-482, 1995.
  45. G. Olivier, Y. Gervais, and D. Mukhedar, A new approach to compute uniform field breakdown of gases, *IEEE Trans. Power Appar. Syst.*, **PAS-97** (3): May/June 1978.
  46. Dielectric breakdown of cryogenic gases and liquids. J. Gerhold, *Cryogenics*, Fig. 22, 1979.
  47. Meek and Craggs, *Electrical Breakdown of Gases*. 662, Fig. 6.14, New York: Wiley, 1978.
  48. J. Gerhold, Dielectric breakdown of cryogenic gases and liquids, Fig. 10, *Cryogenics*, 1979.
  49. D. May and H. Krauth, Influence of the electrode surface condition on the breakdown of liquid helium. *IEEE Trans. Mag.*, **MAG-17** (5): Sept 1981.
  50. Ch. Olivier, Performance of electrodes at the first voltage breakdown in liquid helium, *IEEE Trans. Mag.*, **MAG-17** (5): Sept. 1981.
  51. Fallou, Dielectric breakdown of gaseous He at very low temperatures. *Cryogenics*, April 1970.
  52. R. J. Meats, Butt gap discharges in laminated insulation impregnated with very cold helium gas. *Third Int. Conf. Gas Discharges, IEEE*, London, 1974.
  53. J. Hiley and R. S. Dhariwal, Dielectric breakdown in high density helium and in helium impregnated solid dielectrics. *Cryogenics*, **25**: 334, June 1985.
  54. M. Zahn, Preliminary engineering design criteria to prevent spark discharges in TPX, Nov 27, 1994, TPX Memorandum No. 1314-941127-MIT-MZahn-01.
  55. A. M. Miri, Transient voltage oscillations in a large superconductive coil. *9th Int. Symp. High Voltage Eng.*, Graz, Austria, Aug 28-Sept 1, 1994.
  56. A. Migliori, R. I. Schermer, and M. D. Henke, Dielectric tracking in liquid helium. *Cryogenics*, 442, 1978.
  57. R. A. Haarman and K. D. Williamson, Jr., Electrical breakdown and tracking characteristics of pulsed high voltages in cryogenic helium and nitrogen. *Adv Cryo Eng.* **21** (102): Plenum Press, 1976.
  58. M. Irmisch et al., Breakdown characteristic of helium gas at cryogenic temperatures and low pressures with respect to a local helium leak. KFK Report 31-03-05 P02B, June 1992.
  59. C. Henning, Reliability of large superconducting magnets through design. *IEEE Trans. Mag.*, **MAG-17** (1): 618, Jan. 1981.
  60. R. W. Fast and H. L. Hart, Use of glass beads to increase the breakdown voltage in subatmosphere, cold helium gas. *Advances in Cryogenic Engineering*, **35**: Plenum Press, New York, 1990, 809-812.
  61. P. G. Marston et al., Magnet Failure Workshop, Journal de Physique, Colloque C1, Tome 45, Jan 1984, C1-637; F. Arendt and P. Komarek, Potential failures and hazards in superconducting magnet systems for fusion reactors. *Nuc Tech/Fusion*, **1**: Oct, 1981; R. J. Thome, J. Bart Czirr, and J. H. Schultz, Survey of selected magnet failures and accidents.

JOEL H. SCHULTZ  
Massachusetts Institute of  
Technology

## SUPERCONDUCTIVITY: CRITICAL CURRENT

In the application of superconductors, the superconducting critical current density is often the most important parameter in the design and engineering of practical devices. The reason is that the majority of applications for superconducting wires involve building electromagnets, which develop their magnetic field by virtue of the number of ampere-turns in the magnet winding. Examples of electromagnets presently in the marketplace include magnetic resonance imaging (MRI) magnets; high field research magnets; beam-bending, focusing, and detector magnets for high-energy physics research; superconducting energy storage systems; and superconducting motors and generators (1). As the critical current density of the superconducting wire increases, the achievable magnetic field increases, and the amount of superconductor needed for constructing the magnet decreases. Thus, there is a direct driving force to increase the critical current density of superconductors for magnet applications.

As with the critical temperature ( $T_C$ ) and critical magnetic field ( $H_{C2}$ ), the critical current density ( $J_C$ ) marks the limit of the superconducting state. For direct current (dc) densities less than the  $J_C$ , the current is carried without resistive losses, and thus no power input. For current densities greater than  $J_C$ , a voltage develops along the superconductor, and the zero resistance condition breaks down.

The  $T_C$  and  $H_{C2}$ , which are both determined by the chemistry and physics of the superconducting system, are relatively unaffected by the processing of the superconductor. The same

is not true for the  $J_C$ , which can be radically changed within a given superconducting material by varying the fabrication process and therefore the material's microstructure. For example, within the Nb-Ti alloy system, once the composition of the alloy has been chosen, the  $T_C$  and  $H_{C2}$  are essentially determined. However, by varying the metallurgical treatments of the alloy as it is processed into wire, it is possible to vary the  $J_C$  by factors of 1000 or more (2,3). The potential for controlling the critical current density through processing provides materials science researchers with hope for improving the properties of technical superconductors.

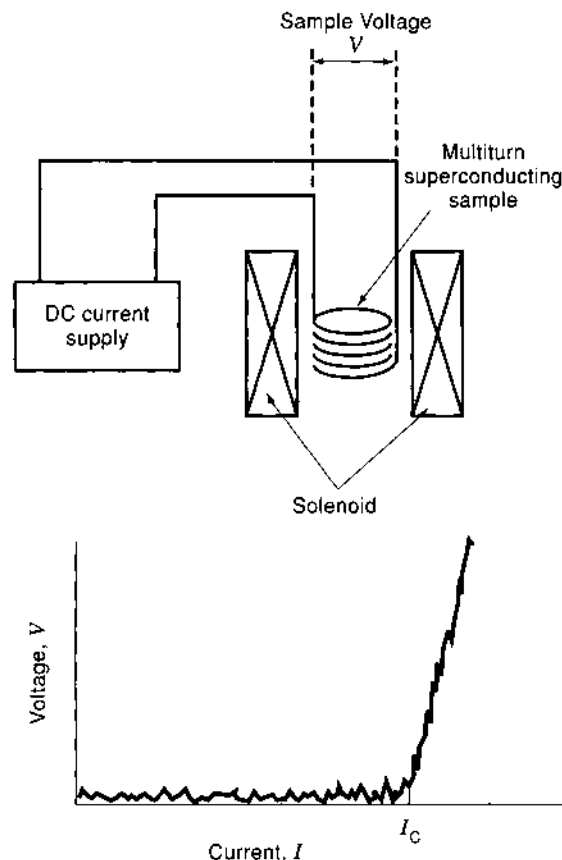
In describing the theory of critical current density in superconductors, it will be useful to consider two length scales. The first is the London penetration depth  $\lambda$ , which is the distance over which an externally applied magnetic field penetrates into a superconductor. This is essentially the distance over which we expect to see large changes in the magnitude of the magnetic fields inside the superconductor. The second length scale is the coherence length  $\xi$ , which is the distance over which the superconducting order parameter (or alternatively, the density of superelectrons) varies.

In the Ginzburg-Landau theory of superconductivity, the ratio of the penetration depth to the coherence length is called the Ginzburg-Landau parameter  $\kappa$ , where  $\kappa = \lambda/\xi$ . The Ginzburg-Landau parameter distinguishes between the two broad classes of superconductors; type I, for which  $\kappa < 1/\sqrt{2}$ , and type II, for which  $\kappa > 1/\sqrt{2}$ . The high critical current density superconductors are all type II materials, and  $\kappa$  is quite large for many of these materials, on the order of 50 or so.

The remainder of this article describes the measurement of  $J_C$ , the basic theory of critical currents in type I and type II superconductors, the flux-line lattice and flux pinning, and thermal effects in determining the  $J_C$ . As an overview of a large and complex subject that has occupied many researchers, this article cannot be complete. For a more detailed discussion of general superconductivity, see texts by Rose-Innes (4), Tinkham (5), and especially Orlando (6) for an excellent presentation from an electrical engineering perspective. Detailed surveys of flux pinning and critical current density may be found in the monographs by Campbell and Evetts (7), and Ullmaier (8), which form the basis of much of the discussion following. Although the majority of the references and discussion uses examples from low-temperature superconductivity (LTS), the principles described are also equally applicable to high-temperature superconducting (HTS) materials.

## THE RESISTIVE MEASUREMENT OF CRITICAL CURRENT DENSITY

The critical current is usually found by a simple four-point measurement, using the change in the resistance of the sample to determine the transition between superconductivity and normal conductivity (see Fig. 1). The sample, a long coil of wire, is placed in the bore of a high field electromagnet, and a dc current is passed through it. The voltage along the length of the superconducting wire is measured. A zero voltage indicates superconductivity, whereas a nonzero measurement indicates resistive dissipation and loss of superconductivity. This "resistive" measurement technique may be used to measure critical temperatures by passing a small constant current, varying the sample temperature, and measuring the sample voltage. It can



**Figure 1.** Schematic of the resistive measurement of the critical current density. The superconducting sample in the center of a high magnetic field solenoid produces a voltage when the current exceeds the critical current. The plot shows the  $V(I)$  characteristic of the superconducting wire. The critical current  $I_C$  is defined at the appearance of a measurable voltage.

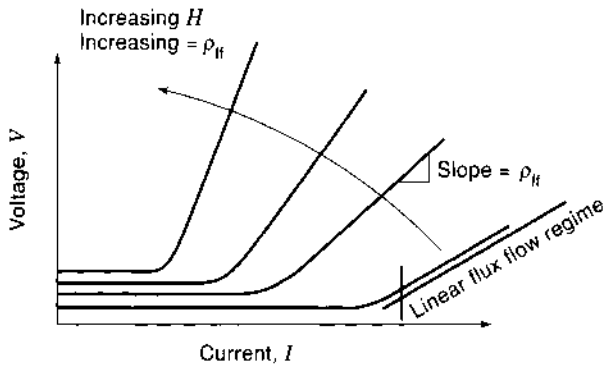
also be used to measure the critical magnetic field  $H_{C2}$  of type II superconductors by varying the magnetic field on the sample. Of more concern to the present discussion is measuring the critical current  $I_C$ . This is done by holding the sample in a constant magnetic field and increasing the current through the superconductor until a voltage increase is measured (see Fig. 1). In this way the critical current is determined.

As the current is increased through the superconducting wire, the voltage along the wire slowly increases from zero until a rapid increase occurs near the critical current. If the  $V(I)$  characteristic is measured over a large enough voltage range, the curve looks like that shown in Fig. 2. Early experiments (9) showed that the  $V(I)$  characteristic at high currents becomes linear, and the resistivity depends on the applied magnetic field, roughly following

$$\rho_{ff} = \left( \frac{H}{H_{C2}} \right) \rho_n \quad (1)$$

where  $\rho_{ff}$  and  $\rho_n$  are the high current resistivity (flux flow resistivity) and normal state resistivity of the superconductor, respectively. This linear region at high currents is called the "flux-flow" regime for reasons that are described following.

Eventually, the increasing power dissipation due to the flux flow resistivity causes the temperature of the superconductor



**Figure 2.** Schematic of the voltage versus current measurement of the critical current of a superconducting wire as a function of the applied magnetic field. The zero voltage points have been offset for clarity. As the field increases, the critical current decreases, and the high current slope  $\rho_H$  increases.

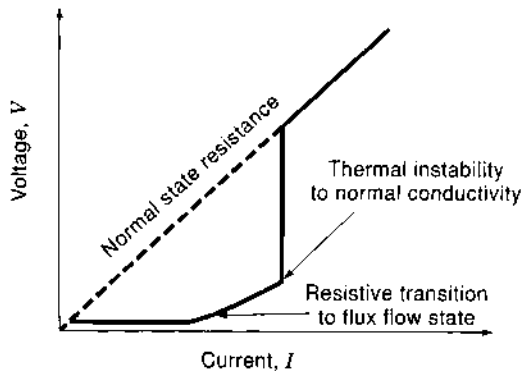
to rise above  $T_C$  and there is a phase transition to the normal state (Fig. 3). An important point to note is that within the flux flow regime the sample is still superconducting, though it is no longer in the zero dissipation (or zero resistance) condition.

Empirically it has been found that the shape of the  $V(I)$  curve at low voltages (well below the linear flux-flow regime) can be described by a power law

$$(I) = K \left( \frac{I}{I_C} \right)^n \tag{2}$$

so that a plot of  $\log V$  versus  $\log I$  from the  $I_C$  measurement yields a straight line whose slope is sample-dependent and also depends on the applied magnetic field (10). Recent measurements on high temperature superconducting (HTS) materials also show this power law behavior. Typically HTS materials are measured over a much larger range of current and voltage than is usual in the LTS materials (11,12).

Over the years of development of LTS wires, it has been empirically determined that larger slopes of the  $\log V$  versus



**Figure 3.** Schematic of the full  $V(I)$  characteristic of a superconducting wire. At low currents the superconductor exhibits zero resistance and zero voltage. At the critical current there is a transition into the flux-flow regime, and the voltage increases. At larger currents, ohmic heating causes the temperature to rise above the critical temperature, and the sample thermally runs away to the normal state resistance behavior.

$\log I$  plot were correlated with higher current density or higher “quality” superconducting wires. This slope has come to be known as the “ $n$ -value” of a wire and is commonly reported by LTS wire manufacturers. Values of  $n$  between 40 and 100 are generally considered to indicate good quality wires. From Eq. (2) one can see that, as the  $n$ -value for a wire increases, the transition from the zero resistance state to the flux flow state becomes steeper and narrower.

Because of the gradual transition from zero voltage to the linear flux-flow regime, the critical current is usually determined by using a standard measurement criterion. For many years magnet designers preferred a constant resistivity criterion, for example,  $\rho = 10^{-14} \Omega\text{-m}$ . A line is drawn on the  $V(I)$  data plot with a resistive slope corresponding to  $10^{-14} \Omega\text{-m}$ , based on the dimensions of the sample. The intersection of this line with the  $V(I)$  measurement is the critical current  $I_C$ . Because the  $V(I)$  characteristic is curved, the measured value of the  $I_C$  depends on the criterion used, so this information must be provided along with the measurement value. Historically, the  $10^{-14} \Omega\text{-m}$  criterion has been used by magnet designers because it was found that early superconducting magnets thermally ran away to the normal state (quench) at a magnet resistivity of about  $10^{-14} \Omega\text{-m}$ . A second commonly used criterion is the constant electric field criterion, in which the  $I_C$  is given by the current at which the  $V(I)$  measurement exceeds a constant electric field value, for instance,  $10 \mu\text{V/m}$ .

It is important here to draw a distinction between the critical current  $I_C$ , which has units of amperes, and the critical current density  $J_C$ , which has units of amperes per unit cross-sectional area. The fundamental property of the superconducting state is the  $J_C$ , which is the maximum current per unit cross-sectional area of superconductor that is carried without resistive losses. The  $J_C$  is determined by measuring the critical current  $I_C$  of the specimen and dividing by the cross-sectional area of the superconductor:  $J_C = I_C/A$ .

An additional definition of importance is the engineering  $J_C$  which is defined as the maximum transport current per unit cross section of superconducting wire. Most technological superconductors are fabricated as a composite of superconducting and normal metal for thermal and mechanical stability (1,13). For the magnet designer, the engineering  $J_C$  (sometimes abbreviated as  $J_E$ ) determines the available current in the magnet windings. The distinction between  $J_C$  and  $J_E$  is especially important when magnet designers are working with superconducting composites in which the superconducting area is a small fraction of the total wire cross section, as in tape composites and many early HTS wires.

It is also worthwhile at this point to describe the difference between transport currents and shielding currents. Superconductors placed in a magnetic field exclude some or all of the magnetic flux from the bulk of the superconductor (known as the Meissner-Ochsenfeld effect). For this exclusion to occur, shielding currents flow on the surface of the superconductor such that the magnetic field produced by the shielding currents opposes the applied field and cancels it out. In general, the shielding currents flow in loops that are closed entirely within the superconductor. In contrast, transport currents are those currents applied from outside the superconductor using external current sources. The transport currents are the currents used to produce magnetic fields in the superconducting magnets and to make the resistive measurements of the critical parameters of superconductivity.

With this basic understanding of how the  $J_C$  is typically measured we can begin to discuss the physical mechanisms limiting  $J_C$  in practical materials.

### ULTIMATE LIMITS TO $J_C$ : THE DEPAIRING CRITICAL CURRENT DENSITY

In the Bardeen–Cooper–Schrieffer theory of superconductivity, the charge carriers are pairs of electrons bound together by a positive electron–phonon interactive force. The bonding energy of the superelectron pair at zero kelvin is denoted as the energy gap  $\Delta(0)$ . The critical temperature can be determined from the energy gap as the temperature at which the thermal excitation energy  $kT$  is equal to the energy gap bonding the superelectron pair together. More rigorously this relationship is

$$2\Delta(0) = 3.5 kT_C \quad (3)$$

The superconductivity stops because the thermal energy is sufficient to “depair” or “decouple” the superelectrons.

Similarly, the thermodynamic critical magnetic field  $H_C$  at zero kelvin can be determined from the energy gap by using the magnetic free energy difference between the normal and superconducting states:

$$\frac{1}{2} \mu_0 H_C^2 = \frac{1}{2} N(\epsilon_F) [\Delta(0)]^2 \quad (4)$$

where  $N(\epsilon_F)$  is the density of states at the Fermi energy and  $\mu_0$  is the permittivity of free space. Thus, the critical magnetic field is that magnetic field for which the magnetic energy is large enough to break apart the electron pair.

In a similar approach to the arguments used to estimate  $T_C$  and  $H_C$ , one would expect that there is an ultimate critical current density limited by the kinetic energy of the superelectron pairs in a transport current. When the kinetic energy of the electrons exceeds the energy gap, the pairs break apart and become normal (resistive) charge carriers. The kinetic energy of the superelectrons can be written as

$$\text{KE} = \frac{(m^* v_F^2)}{2} = \frac{p_F^2}{2m^*} \approx 2\Delta(0) \quad (5)$$

where  $m^*$  is the effective mass of the charge carriers (in this case  $m = 2m_e$ , two times the mass of an electron), and  $v_F$  and  $p_F$  are the Fermi velocity and momentum, respectively.

The current density produced by these charge carriers is given by

$$J = qn_s v = 2en_s v_F \quad (6)$$

where  $q$  is the charge of the current carrier ( $= 2e$  for electron pairs) and  $n_s$  is the density of superelectron charge carriers.

The depairing critical current density can be found by combining Eqs. (5) and (6):

$$J_D \approx \frac{10en_s \Delta(0)}{p_F} \quad (7)$$

Orlando (6) derives an equivalent form of Eq. (7) from

**Table 1. Theoretical Depairing Critical Current Density<sup>a</sup>**

| Superconductor  | $\lambda$ , nm | $\xi$ , nm | $J_D$ , A/m <sup>2</sup> |
|---|----------------|------------|--------------------------|
| NbTi  | 300            | 4          | $2 \times 10^{11}$       |
| Nb <sub>3</sub> Sn                                    | 65             | 3          | $8 \times 10^{12}$       |
| YBa <sub>2</sub> Cu <sub>3</sub> O ( <i>ab</i> plane) | 30             | 3          | $4 \times 10^{13}$       |
| YBa <sub>2</sub> Cu <sub>3</sub> O ( <i>c</i> plane)  | 200            | 0.4        | $6 \times 10^{12}$       |

<sup>a</sup> Calculated using Eq. (8) for several important superconductors.

Ginzburg–Landau theory as

$$J_D = \frac{\phi_0}{3\sqrt{3}\pi\mu_0\lambda^2\xi} \quad (8)$$

where  $\lambda$  and  $\xi$  are the penetration depth and coherence length, respectively, and  $\phi_0$  is the magnetic flux quantum. Using either Eqs. (7) or (8) we can calculate the depairing critical current density shown in Table 1 at zero kelvin and zero magnetic field for several superconducting systems.

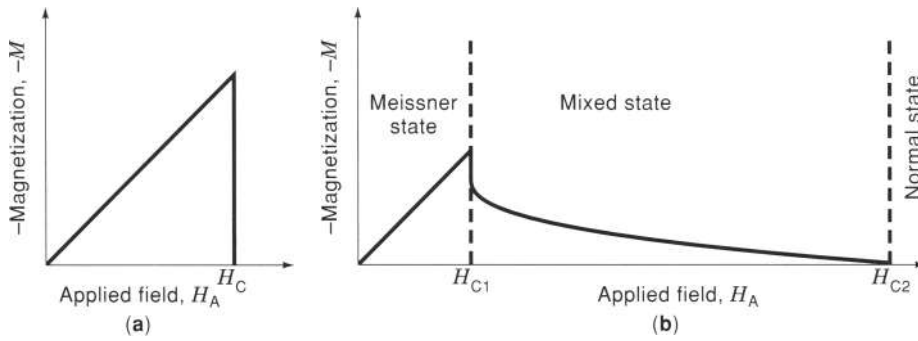
As can be seen from Table 1, the depairing critical current densities are quite large, especially when compared with the typical current density in standard copper household wiring of about  $10^7$  A/m<sup>2</sup>. The superconducting materials shown here have theoretical critical current densities at least  $10^4$  times larger than copper. It is because of these large values of critical current density with no resistive losses (and therefore no power dissipation) that superconductors are so important in large electromagnet applications (1).

It should be remembered that the values of  $J_D$  listed in Table 1 are calculated for zero kelvin and zero magnetic field, and in practice these conditions do not hold. In fact, these values of current density have never been reached because of other practical limitations. One of these limitations is the self-field produced by a wire carrying a transport current. As the transport current through the wire is increased, the self-field at the surface of the superconductor increases. At some point, the magnetic field due to the transport current becomes equal to the critical magnetic field of the wire, and the superconductivity breaks down. This model of the practical limit to  $J_C$  is known as Silsbee’s hypothesis (14), and is usually applied to find the critical current limit of type I superconductors.

Of greater technological importance than the Silsbee limit in type I materials is the limitation of the  $J_C$  in type II superconductors because type II superconductors display superconductivity up to larger magnetic field values than type I superconductors. To understand the factors limiting the  $J_C$  in type II materials, it is necessary to review the magnetic properties of these superconductors and introduce the concept of the flux-line lattice.

### THE MAGNETIC FLUX LINE LATTICE

The principal difference between type I and type II superconducting materials lies in their response to an applied magnetic field. Type I superconductors exclude an applied magnetic field from the body of the superconductor up to the thermodynamic critical field  $H_C$ . To exclude this magnetic flux, a shielding current is established on the surface of the superconductor that flows in a direction so as to produce a flux density equal and opposite to the applied field. This surface current flows in a



**Figure 4.** The magnetic behavior of type I and type II superconductors. Type I superconductors exclude the applied field from the bulk of the superconductor by producing a supercurrent on the surface to cancel the applied field, yielding the magnetization versus field plot shown in (a). Type II superconductors exclude the applied field up to a lower critical field  $H_{C1}$  and then allow the field to enter the bulk as flux quanta  $\phi_0$ , until the upper critical field  $H_{C2}$  is reached (b). At this point the superconductivity is destroyed by the applied field.

surface layer whose thickness is equal to the magnetic penetration depth  $\lambda$ . The magnitude of the surface current can be found by using Ampere's law which states that the spatial variation in the magnetic field is proportional to the current density flowing:

$$\nabla \times \mathbf{H} = \mathbf{J} \quad (9)$$

where both  $\mathbf{H}$  and  $\mathbf{J}$  are vector quantities. For the one-dimensional case of a semi-infinite slab of type I superconductor in the  $y - z$  plane, and an applied field  $H$ , parallel to the slab in the  $z$ -direction, Ampere's law becomes

$$dH_z/dx = J_y \quad (10)$$

Because the magnetic field decays over a length  $\lambda$  into the superconductor, we know that the applied field and the shielding current density are related approximately as

$$\frac{H_A}{\lambda} = J_y \quad (11)$$

As the magnitude of the applied field increases, the magnitude of the current density increases to shield the superconductor from the field. The maximum current density is obtained when the applied field at the surface of the type I superconductor is equal to  $H_C$ , in which case  $J_{MAX} = H_C/\lambda$ . This shielding current density is the same as the depairing current density described by Eqs. (7) and (8).

The shielding currents in the type I superconductor effectively provide a diamagnetic magnetization,  $M = -H_A$ , as shown in Fig. 4(a), called the Meissner-Ochsenfeld effect.

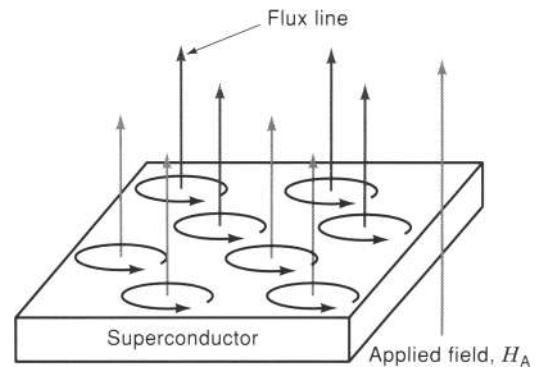
For type II superconductors, the magnetic response is somewhat different. Up to a lower critical field  $H_{C1}$ , the magnetic response of type II superconductors is the same as that of type I and shows a full flux expulsion, with  $M = -H_A$  (see Fig. 4b). In this region, the superconductor is said to be in the Meissner state. For magnetic fields higher than  $H_{C1}$ , the magnetic free energy balance of the superconductor makes it energetically favorable for the magnetic field to enter the bulk of the superconductor. As the magnetic flux enters the bulk superconductor, it breaks into quantized units of flux  $\phi_0$ , variously called the flux quantum, fluxon, fluxoid, flux vortex, or flux line. The flux quantum has a magnitude of  $\phi_0 = 2.0679 \times 10^{-15} \text{ T}\cdot\text{m}^2$ .

The individual flux quanta, or flux lines, orient themselves parallel to the applied field and effectively reduce the magnetization of the type II superconductor below that of the perfect diamagnetism of the Meissner state. This state of lower

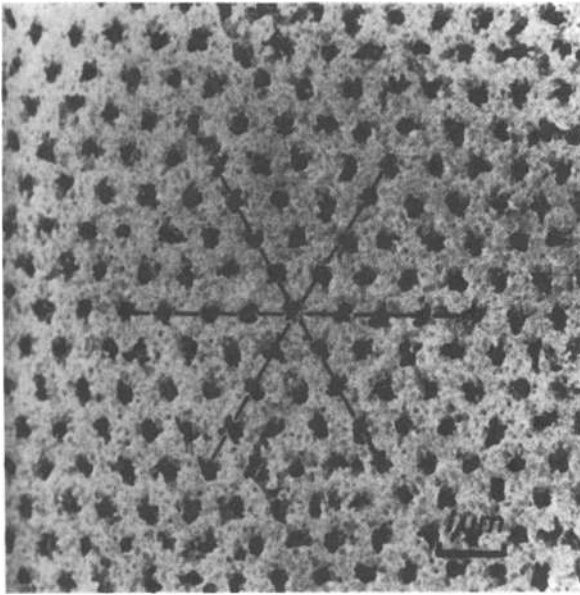
magnetization is called the mixed state. As the applied field increases, the number of flux lines per unit area increases in the superconductor and  $M$  approaches zero. Eventually the flux lines touch one another, and the field inside the superconductor becomes equal to the applied magnetic field, driving the magnetization to zero at the upper critical magnetic field  $H_{C2}$ . The superconducting material remains superconducting up to large values of the applied magnetic field ( $H_{C2}$ ), and this is one of the main reasons that the type II materials are used for electro-magnet applications.

The interaction of the individual flux lines with one another is similar to that of two parallel bar magnets. Because the orientation of the field in the flux lines is the same, they repel one another strongly. This causes the flux lines to distribute themselves in a periodic lattice to minimize the interflux line interactions (Fig. 5). This periodic structure, called the flux line lattice (FLL), was theoretically predicted by Abrikosov (15) using extensions of the Ginzburg-Landau theory of superconductivity. Abrikosov found that the lowest free energy configuration for the FLL is a triangular or hexagonal "crystal." The FLL has been experimentally verified in several ways, including magnetic particle decoration techniques (16) and diffraction from the flux line crystal by using the magnetic moment of neutrons (17). The magnetic decoration technique, in particular, provides a striking visualization of the periodicity of the flux line lattice, as shown in Fig. 6.

We can model an isolated flux line as a cylindrical core of normal-phase material in which the superconductivity has



**Figure 5.** Schematic of the magnetic flux line lattice in a type II superconductor for applied fields between  $H_{C1} < H_A < H_{C2}$ . The flux lines arrange themselves in a triangular or hexagonal lattice due to the inter-flux-line magnetic repulsive forces.



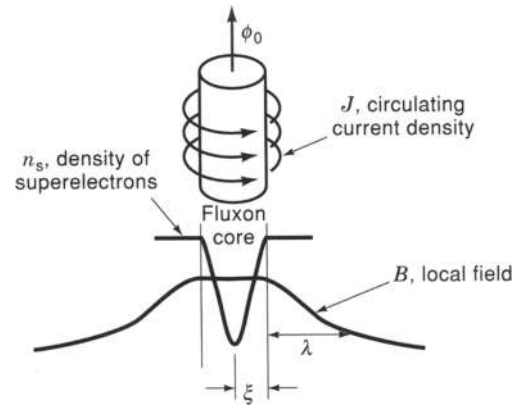
**Figure 6.** The magnetic flux line lattice imaged using a magnetic particle decoration technique. The small scale magnetic particles are attracted to the regions of large flux gradient at the flux line cores and then can be imaged by transmission electron microscopy. The periodic triangular flux line lattice can be clearly distinguished. Reprinted with permission from H. Trauble and U. Essmann, *J. Appl. Phys.*, **39**(9): 4052–4056, 1968. Copyright 1968, American Institute of Physics.

been destroyed by the magnetic field and which is surrounded by a circulating supercurrent. The magnetic flux resides within the core and decays into the bulk of the superconductor over a distance of the magnetic penetration depth  $\lambda$  (Fig. 7). Within this range, the local magnetic field strength  $H$  is changing, and therefore, by using Ampere’s law [Eq. (9)], there is a current flowing in the superconductor. This current is analogous to the shielding current that flows on the superconductor surface to exclude the magnetic field. In this case it is a circulating current that flows around the flux-line core and has an orientation and magnitude needed to produce the  $\phi_0$  of magnetic flux in the core (Fig. 7). This circulating current is the origin of the name “flux vortex.”

The cylindrical core has a diameter twice the coherence length ( $2\xi$ ). The coherence length is the distance over which the superconducting order parameter  $|\psi|^2$  (or the density of superconducting electron pairs  $n_s$ ) changes from its maximum value at the core radius to zero in the center of the core (Fig. 7). The field strength in the core can be estimated as the magnetic flux divided by the cross sectional area of the flux line:

$$H_{\text{CORE}} = \frac{\phi_0}{\mu_0 \pi \xi^2} = H_{C2} \quad (12)$$

As we introduce more flux lines into the interior of the superconductor, the circulating supercurrents of the neighboring flux lines begin to interact and repel one another, leading to the periodic structure of the flux-line lattice. The flux density at any point in the superconductor is found from the number density of flux lines as  $B = n_\phi \phi_0 / A$ , where  $n_\phi$  is the number of flux quanta in the cross-sectional area  $A$ . In a homogeneous



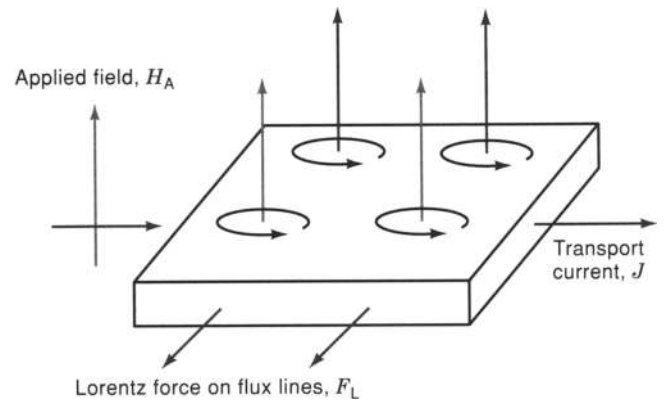
**Figure 7.** Model of an isolated flux line as a core of normal material containing the magnetic flux quantum  $\phi_0$ . The magnetic field falls off over a distance of  $\lambda$ , the penetration depth. The core has a radius of  $\xi$ , the superconducting coherence length, and the density of superelectrons falls to zero at the center of the flux line core.

type II superconductor in the mixed state (i.e., in an applied field between  $H_{C1}$  and  $H_{C2}$ ), the magnetic flux breaks up into flux lines, each containing one quantum ( $\phi_0$ ), of magnetic flux that are periodically arranged in this two-dimensional “crystal” lattice.

If a transport current is applied to such a superconducting wire, where the current flow is along the axis of the wire and the applied magnetic field is perpendicular to the axis (as shown in Fig. 8), there is an interactive force between the flux lines and the transport current. Lorentz’s law for the force on a charged particle moving through a magnetic field is given by

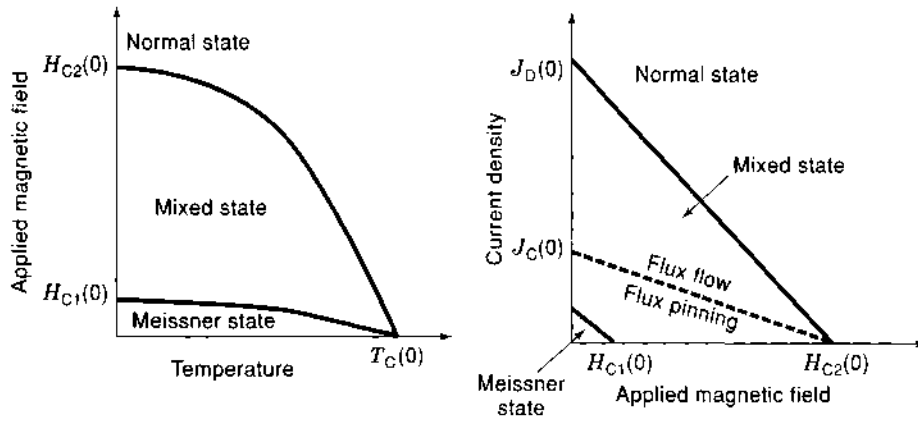
$$\mathbf{F}_L = \mathbf{J} \times \mathbf{B} \quad (13)$$

where  $\mathbf{F}_L$  is the Lorentz force density acting between the current of  $\mathbf{J}$  and the flux density  $\mathbf{B}$ . The  $\mathbf{F}_L$  has units of newtons per cubic meter and acts in a direction perpendicular to both the flux density  $\mathbf{B}$  and the transport current density  $\mathbf{J}$  (Fig. 8).



**Figure 8.** The orientation of the magnetic field, transport current, and Lorentz force acting on the flux line lattice. The Lorentz force between the flux lines and the transport current causes the flux lines to move across the superconductor.





**Figure 9.** The  $H$ - $T$  and  $J$ - $H$  phase diagrams for type II superconductors. At low applied fields the superconductor is in the Meissner state. At higher fields the superconductor enters the mixed state with the creation of the flux line lattice (FLL). As the transport current is increased from zero, the Lorentz force on the FLL eventually causes it to move, causing flux-flow dissipation and a resistive voltage, shown as the dotted line. The superconducting state does not end until the current density exceeds the depinning current density or the temperature rises above  $T_C$ . As flux pinning increases, the transition to flux flow occurs closer to the depairing critical current density limit.

The result of the Lorentz force acting on the FLL is to push the flux lines across the superconductor. The movement of the flux lines corresponds to a change in the flux density within the superconducting circuit with time (Fig. 1), and from Maxwell's equations for such a case (6),

$$\frac{d\mathbf{B}}{dt} = -\nabla \times \mathbf{E} \quad (14)$$

In other words, the moving magnetic flux lines produce an electric field gradient (or voltage) in the direction of the transport current flow. As the FLL moves, a voltage is generated that must be supplied by the external power supply. The consequence of this flux motion is that the superconductor no longer supports a transport current with zero dissipation, and therefore the zero resistance state no longer exists.

It is important here to draw a distinction between the loss of the superconducting state and the loss of the zero resistance condition. As long as the superconductor remains in fields less than the upper critical field  $H_{C2}$  at temperatures lower than  $T_C$  and carries transport currents less than  $J_D$ , it is in the superconducting state. The nonzero resistance occurs only because the FLL is moving under the Lorentz force produced by the transport current and changing the magnetic flux linked by the superconductor. This is the "flux-flow" regime described earlier. If we could prevent the FLL from moving because of the Lorentz force, the zero resistance condition would persist to higher transport currents, effectively increasing the critical current density.

The bulk pinning force density  $F_P$  (N/m<sup>3</sup>), is defined for samples carrying a transport current in a transverse field as the critical Lorentz force:

$$\mathbf{F}_P = |\mathbf{F}_{LC}| = \mathbf{J}_C \times \mathbf{B} \quad (15)$$

where  $\mathbf{J}_C$  is the current density at which voltage losses occur in the superconductor.

In essence, holding the FLL against the Lorentz force adds a transition line for the change from flux pinning to flux flow to the phase diagram of the type II superconductor in Fig. 9. The solid lines represent thermodynamic phase transitions between the superconducting (Meissner and mixed) states and the normal state, and the dashed line shows the transition from the flux-pinning (zero resistance) condition to the dissipative flux-flow condition.

Ideally, one would like to move the dashed line up in current density as close to the phase transition line (determined by  $J_D$ ) as possible. This is the goal of the flux pinning discussed in the next section.

## FLUX PINNING

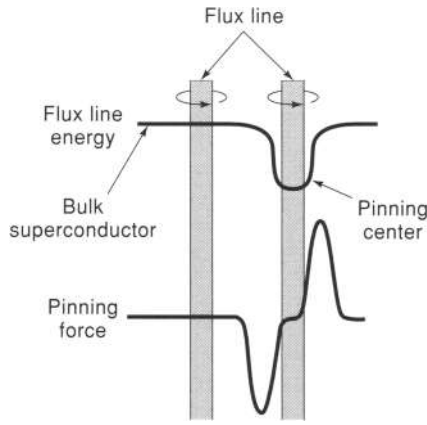
To increase the current that a superconductor may carry without power dissipation, it is necessary to restrain the FLL against the Lorentz force by "pinning" it in place. There are several mechanisms by which the FLL may be pinned, and generally these rely on developing microstructural features that interact with the individual flux lines. Examples of microstructural features that provide pinning resistance to the Lorentz force include normal conducting precipitates, inclusions, voids, and grain boundaries.

The basic theory of flux pinning in type II superconductors is conveniently broken into three sections. These are basic interactive forces, summation theory, and scaling laws (7,8).

### Basic Interactive Forces

The basic interactive forces are the forces between single, isolated flux lines and individual pinning centers. The usual model for the basic interactive force is that the pinning center must provide a spatial variation of the thermodynamic free energy of the flux line. This can be visualized as either an energy well (Fig. 10) or an energy hill. In the case shown in the upper part of Fig. 10 the flux line has a lower free energy when it sits in the energy well of the pinning center than it does in the bulk superconductor, and thus there is a pinning force holding the flux line in the well. The pinning force is related to the free energy by the first derivative with respect to position, so that the pinning force curve looks like that shown in the lower part of Fig. 10. The deeper the potential well, the steeper the energy profile, and the larger the pinning force.

If a Lorentz force is applied to a flux line trapped in this potential well, the flux line moves in the direction of the Lorentz force until it is balanced by the oppositely directed pinning force. Thus the flux line is held in place, there is no flux movement, and Eq. (14) shows that there is no dissipation. The transport current is carried without power dissipation, and the zero resistance condition is in effect. Superconducting materials that pin magnetic flux are sometimes called "hard" superconductors analogous to engineering alloys that have been



**Figure 10.** The variation in the free energy of the flux line in the vicinity of a pinning center. The energy well produces a net force on the flux line centering it in the pinning center and constraining it against the Lorentz force of the transport current.

mechanically hardened by treatments to pin the movement of dislocations.

One type of basic interactive force between a single flux line and a single pinning center is called the core interaction. To nucleate a flux line within the superconductor, the system must provide enough energy to convert the core of the flux line to the normal state. This energy (per unit length of flux line), called the condensation energy, is given by the volumetric free energy due to the magnetic field within the flux line and the cross-sectional area of the fluxon core as

$$E_{\text{COND}} = \left( \frac{\mu_0 H_C^2}{2} \right) \pi \xi^2 \quad (16)$$

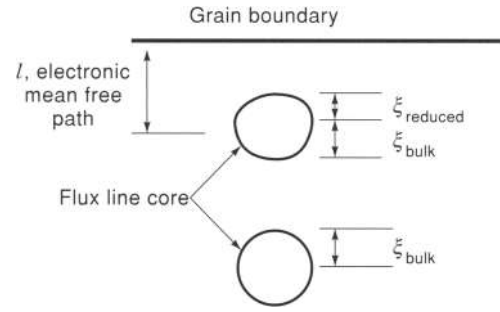
where  $H_C$  is the thermodynamic critical field and  $\xi$  is the superconducting coherence length.

Imagine that the superconductor contains a cylindrical void of diameter  $2\xi$  and its axis is oriented parallel to the flux-line axis. If the flux line were centered on this void, the condensation energy needed to produce the normal core of the flux line would be saved, and the flux line would see a lower free energy at the location of the void than it would in the bulk, similar to Fig. 10. The result of this free energy change is that the flux line requires an increase in its energy per unit length equal to the condensation energy, Eq. (16), to move away from the void. Thus the void acts as a pinning center holding the flux line in place.

As the current density is increased, the Lorentz force on the pinned flux line increases until it exceeds the maximum gradient of the free energy versus position curve (Fig. 10). At this point the flux line is free of the pinning center and moves under the Lorentz force, creating a dissipative loss due to Eq. (14).

There are many different interactions between the flux line and microstructural defects that lead to basic interactive forces and pinning. The core interaction may be applied to voids and also to normal conducting precipitates (as in the Nb–Ti system) or weakly superconducting inclusions, for which there is a spatial dependence of the superconducting condensation energy. A different interaction that is thought to be important in flux pinning in single-phase superconductors, such as Nb<sub>3</sub>Sn, is the grain boundary interaction, first proposed by Zerweck (18).

In the grain boundary interactive model, the grain boundaries are viewed as strong scattering centers for the normal



**Figure 11.** The superconducting coherence length is reduced within an electron mean free path of a scattering defect, such as a grain boundary. This causes the flux line core to distort so that the volume of the flux line changes as it approaches the grain boundary. The variation of the flux-line energy with distance from the grain boundary causes a pinning interaction between the flux line and the grain boundary.

electrons in the metal, thereby reducing the mean free path of the electrons near the grain boundary. When the electron mean free path  $l$  is less than the coherence length, the coherence length depends on the mean free path as

$$\xi_{\text{dirty}} = 0.85(\xi_0 l)^{1/2} \quad (17)$$

This is often referred to as the “dirty limit” since the mean free path of the electrons is much shorter than that of a “clean” high-purity metal (5).

From Eq. (17), the coherence length is reduced within an electron mean free path of the grain boundary. The effect of the change in the coherence length is that as the flux line core moves closer to the grain boundary, it becomes deformed (Fig. 11) and changes its volume so that the total energy (condensation energy times the core volume) changes with distance from the grain boundary. The free energy difference with position leads to a pinning force, as with the core interactive model.

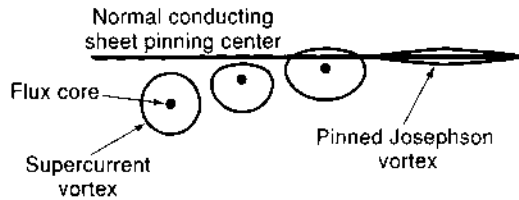
A more general approach to basic interactive forces derives from the Ginzburg–Landau theory of superconductivity, which can be written to show that the variation in the free energy of a flux line depends on spatial variations of the critical field and the Ginzburg–Landau parameter  $\kappa$ . One version of this derivation (7) is to write the variation in the free energy of the flux line due to pinning defects as

$$\delta E = \int \mu_0 H_C^2 \left[ - \left( \frac{\delta H_{C2}}{H_{C2}} \right) |\psi|^2 + \frac{1}{2} \left( \frac{\delta \kappa^2}{\kappa^2} \right) |\psi|^4 \right] dV \quad (18)$$

where  $\psi$  is the unperturbed order parameter of the superconductor.

From this perspective, any spatial variation in either critical field ( $\delta H_{C2}/H_{C2}$ ) or  $\kappa$  ( $\delta \kappa/\kappa$ ) produces a change in the free energy of the flux line that leads to a basic interactive force for pinning. Examples of pinning defects of this sort include normal and weakly superconducting inclusions and precipitates, dislocation clusters (subgrain boundaries), and chemical inhomogeneities, which produce pinning interactions through changes in the electron mean free path and therefore affect  $\kappa$  through the coherence length.

A class of basic interactive forces that can be modeled using “image” vortices to calculate the pinning forces are grouped together as magnetic interactions. In these cases the interaction between the circulating supercurrents and microstructural



**Figure 12.** In the Josephson vortex model the supercurrents surrounding a flux line are distorted as they approach a normal conducting sheet pinning center. The current spreads to the left and right of the flux line to tunnel through the sheet and complete the circuit on the opposite side. When the flux line rests on the sheet, it spreads out along the length of the pinning center, losing the normal core and distributing the flux quantum over a large area.

defects leads to pinning forces, rather than interactions involving the normal core. An example is the pinning force between a flux vortex and an electrically insulating plane. The interactive force is calculated by introducing an identical "image vortex" on the opposite side of the insulating plane. The overlapping supercurrents of the real flux line and the image vortex repel one another and produce a force between the flux line and the insulating plane.

A recent addition to the theory of basic interactive forces is that of the Josephson vortex (19,20). The Josephson vortex model accounts for the large pinning forces found in optimized Nb-Ti alloys for which the pinning centers are thin sheet-like ribbons of normal conducting  $\alpha$ -Ti. These ribbons are much thinner than the flux line core, so that the core interaction does not accurately describe the pinning interaction. At the same time, the ribbons are not insulators, so that the magnetic interaction also does not apply. The model estimates the basic interactive force by considering what happens to the circulating supercurrent around the flux line, as it approaches a normal conducting planar pinning center. Given that the supercurrent cannot readily penetrate the (normal) pinning plane, the supercurrent spreads out along the planar defect, slowly tunneling through the pinning plane as a superconducting Josephson tunneling current to complete the current loop on the other side of the pinning plane (Fig. 12). The effect is that the flux line becomes distributed over a broad area of the pinning plane and produces a pinning force due to the Josephson current interactions.

In summary, basic interactive forces can arise from many different physical mechanisms, of which only a few have been described here. By providing a spatial variation in the free energy of the individual flux lines, the pinning centers produce a pinning force that holds the flux lines stationary against the Lorentz force, thus increasing the critical current density of the superconductor.

### Summation Theory

The second key part in the discussion of flux pinning is summation theory. Given a model for the basic interactive forces between individual pinning centers and individual flux lines, now we must consider the effect of large numbers of flux lines interacting with large numbers of pinning centers. The principle complication of summation is that the flux lines interact repulsively with one another and, in the absence of a pinning force, order themselves in the flux line lattice. Thus, the flux line lattice acts as a two-dimensional, elastic, crystalline solid.

If the inter-flux-line forces are weak compared to the basic interactive forces with the pinning centers, then the individual flux lines move out of the periodic FLL and arrange themselves so that as many flux lines as possible are located on the pinning centers. If the number density of flux lines is less than or equal to the number of pinning centers (for instance, at small applied fields), then each flux line is individually pinned, and the bulk pinning force is large. This is called direct summation, and the bulk pinning force density is just the number density of pinning centers times the basic interactive force.

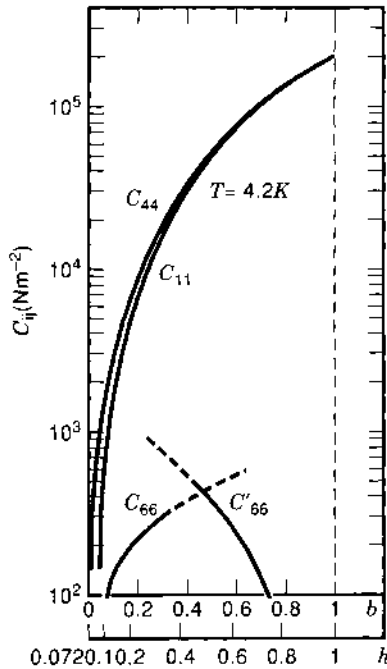
At the other extreme in which the interaction between the flux lines in the flux line lattice is infinitely strong, the FLL is completely rigid, and there can be no bulk pinning force due to a collection of randomly distributed pinning centers because, for any position of the FLL relative to the random array of pinning centers, there will be as many basic interactive forces pulling the FLL to the left as to the right, and the bulk pinning force density averages to zero. Even though the basic interactive forces are very large, if the FLL acts as a rigid solid because of interfluxon forces, there will be no bulk pinning force, and the FLL will move under the Lorentz force due to the transport current, yielding a low  $J_C$ .

The correct description of pinning certainly lies somewhere between these two extremes of direct summation and the rigid FLL lattice. There are several models proposed to account for the summation of the basic interactive forces, and all of them depend strongly on the elastic properties of the FLL as determined by the inter-flux-line forces. In essence, the FLL is a crystalline solid that is placed under an external load by the competition of the transport-current-induced Lorentz forces and the restraint of the pinning forces. As the Lorentz force loading increases, the FLL elastically distorts until either the pinning forces are exceeded, at which point the entire FLL breaks free and moves in unison through the superconductor, or the load overcomes the inter-flux-line forces, and the periodic FLL breaks apart.

This is analogous to mechanically loading a tensile specimen above its elastic limit and into the plastic deformation region. As the mechanical test specimen is plastically distorted, crystalline defects in the specimen are created (dislocations), and the mechanical properties depend strongly on the presence of these crystal defects. In the superconductor with pinning centers, the increasing Lorentz force begins to introduce crystal defects which fragment the FLL into a polycrystalline FLL. The crystalline nature of the FLL has been experimentally observed, as has the polycrystalline and defective FLL, by using both magnetic particle decoration techniques and neutron scattering (21,22). It is also the case that the presence of FLL crystal defects strongly affect its mechanical properties and response to Lorentz force loading (7).

There have been several calculations of the elastic behavior of the FLL. An example is shown in Fig. 13 (8,23) for a NbTa alloy superconductor.  $C_{11}$  is the elastic modulus in the plane normal to the flux line axes. This is a measure of the stiffness of the FLL while pushing the flux lines closer together.  $C_{44}$  is the elastic tilt modulus which describes the bending of flux lines along their axis.  $C_{66}$  is the shear modulus, which describes the resistance of the FLL to the shear of flux lines past one another. An excellent discussion of the effect of the FLL elastic constants on the deformation of the flux line lattice may be found in Ref. 8.

When the number density of flux lines is larger than the number density of pinning centers (as is often the case), there



**Figure 13.** Labusch calculation of the elastic constants of the flux line lattice in Nb-Ta. Notice that the  $C_{11}$  and  $C_{44}$  elastic constants increase with magnetic field, whereas the shear modulus  $C_{66}$  decreases with field at high fields. The  $b$  and  $h$  are the reduced fields  $B/B_{C2}$  and  $H/H_{C2}$ , respectively. Reprinted with permission from Ref. 8.

are two primary models of the summation behavior of the FLL. These are the flux line lattice shear model and the collective-pinning model.

In the FLL shear model, individual flux lines are strongly pinned on individual pinning centers, and the excess flux lines not directly pinned are held in place against the Lorentz force due to the interfluxon forces. At large enough transport currents, the Lorentz force becomes larger than the elastic shear modulus ( $C_{66}$ ) can support, and the FLL shears, allowing unpinned flux lines to flow between those strongly pinned by the pinning centers. Therefore the critical current is determined, not by the strength of the pinning forces, but by the shear stiffness of the FLL, given by  $C_{66}$ .

In the Brandt model of the flux line elastic constants, the compressive modulus  $C_{11}$  and the tilt modulus  $C_{44}$  both depend on the magnetic field roughly as

$$C_{11} \approx C_{44} \approx \frac{H^2}{4\pi} \quad (19)$$

whereas the shear modulus near  $H_{C2}$  is approximated by

$$C_{66} \approx K \left(1 - \frac{H}{H_{C2}}\right)^2 \quad (20)$$

where  $K$  is a proportionality constant. At high magnetic fields, the FLL is very stiff in bending and compression, but becomes softer and softer in shear as the field approaches  $H_{C2}$ . Therefore from this summation model one would expect that the high field critical current density would be determined by the FLL shear and would be somewhat independent of the basic pinning interaction. A detailed model of the FLL shear mechanism developed

by Kramer (24,25) predicts a different high field behavior of the bulk pinning force density compared to the predictions of the simple direct summation model. Both kinds of behavior have been experimentally observed.

An alternative model for summation that applies in the limit of a larger number density of flux lines than pinning centers is the collective-pinning model (26,27). In collective-pinning theory, the FLL is thought to consist of a polycrystalline collection of "grains" in which the periodic order of the flux line lattice is reasonably well maintained by inter-flux-line forces, but the neighboring grains are uncorrelated with one another. Within a flux lattice grain, the flux lines are strongly pinned to one or more pinning centers, and the correlated group is held in place by the combined action of the basic pinning forces and the inter-flux-line forces.

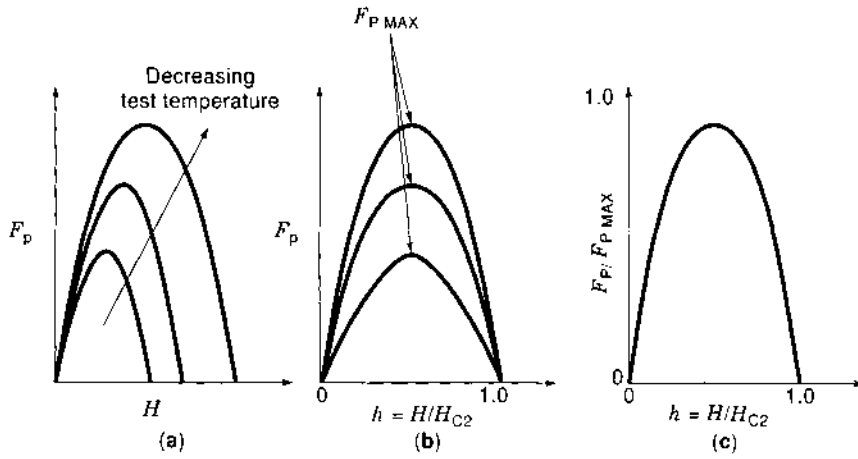
In the collective-pinning model, the size of the correlated FLL grains is determined by the number density of pinning centers and the relative strength of the pinning force and the inter-flux-line forces. In one limit, where the inter-flux-line forces are large compared to the pinning forces, the flux lattice is elastically stiff and the correlated flux grain size is large and involves many pinning centers. In this limit, the bulk pinning force is small, as in the infinitely stiff FLL limit described previously. As the pinning force increases relative to the inter-flux-line forces, the correlated grains become smaller and approach a limiting grain size equal to the mean pinning center spacing, so that each flux lattice grain interacts with only one pinning center. In this limit, the bulk pinning force is at its maximum, because each pinning center is applying a maximal constraint on the FLL. The bulk pinning force is close to the direct summation pinning force in this limit.

In summary, then, the central problem of summation theory is how one combines the effects of the individual fluxon-pinning center interactions and the interfluxon forces to produce a bulk pinning force to hold the FLL against the Lorentz force. Although the limiting cases of direct summation and the rigid FLL are well understood, the behavior of real materials is less clear. Several models for summation have been proposed, primarily the FLL shear and the collective-pinning models, and validation of them with experimental measurements shows that they all have some merit, but none are capable of a complete description of the origins of the bulk pinning force. A final tool for understanding the physical mechanisms behind the pinning force and  $J_C$  is scaling laws.

### Scaling Laws for Flux Pinning

Experimentally it has been found that as the testing temperature is varied, many superconductors exhibit scaling of the bulk pinning force density versus applied magnetic field (28). This is observed by first measuring the  $J_C$  as a function of magnetic field and developing a curve of the bulk pinning force density versus magnetic field using Eq. (15):  $F_P = J_C B$ . This curve has a characteristic shape. It is zero at zero applied field, increases through a maximum with increasing field, and drops to zero again as the  $J_C$  drops to zero at  $H = H_{C2}$  (Fig. 14).

As the test temperature changes, the  $H_{C2}$  of the sample changes, and as a result, the bulk pinning force changes at a given applied field. By scaling the ordinate using the reduced field,  $h = H/H_{C2}$ , and scaling the abscissa with the reduced pinning force density,  $f_P = F_P/F_{P\text{MAX}}$ , where  $F_{P\text{MAX}}$  is the maximum measured bulk pinning force density, the experimental



**Figure 14.** To determine whether a superconducting material displays scaling of the flux pinning curve, the pinning force density versus applied field for several different test temperatures is measured (a). The data are scaled using  $h = H/H_{C2}$  (b) and  $f_p = F_p/F_{P MAX}$  (c). If the sample displays scaling, the different temperatures collapse onto a single plot (c).

data frequently collapse onto a single line for all test temperatures (Fig. 14).

Although the majority of the experiments on scaling  $J_C$  have been performed in low temperature superconductors, a large literature has also developed for scaling behavior in high temperature superconductors. The terminology used in HTS materials has evolved along different lines than those used here, but the basic result is that temperature scaling of the pinning force is also a common feature of these materials (29,30).

It is important that the pinning force follows a scaling law with changes in magnetic field and temperature because scaling implies a single mechanism for flux pinning in the material as a function of temperature, which should be amenable to theoretical prediction. Additionally, if scaling holds for a given material, one only needs to measure the critical current at one temperature and field to estimate the performance at other temperatures, which can be useful for magnet designers.

In general terms, scaling follows an equation, such as

$$F_p = K(H_{C2})^m f(h) \quad (21)$$

where  $K$  and  $m$  are empirically determined constants,  $f(h)$  is a function only of reduced applied field, and the temperature dependence is carried in the variation of  $H_{C2}$ . Typically in LTS materials,  $m$  varies between 1.5 and 2.5. The field function  $f(h)$  may display many different kinds of scaling behavior with magnetic field (31). The two most common are the linear scaling and the quadratic scaling functions.

In linear scaling, the field function is given by

$$f(h) = h(1 - h) \quad (22)$$

This is a symmetrical function of reduced field that has a peak in the pinning force density at  $h = 0.5$ . Often, optimized Nb-Ti superconductors follow a linear scaling behavior (32).

The quadratic scaling field function is expressed by

$$f(h) = h^{1/2}(1 - h)^2 \quad (23)$$

which displays a peak pinning force density at  $h = 0.25$  and a strong quadratic curvature as the applied field nears  $H_{C2}$ . The quadratic behavior is usually associated with  $Nb_3Sn$  and other single-phase superconductors (33).

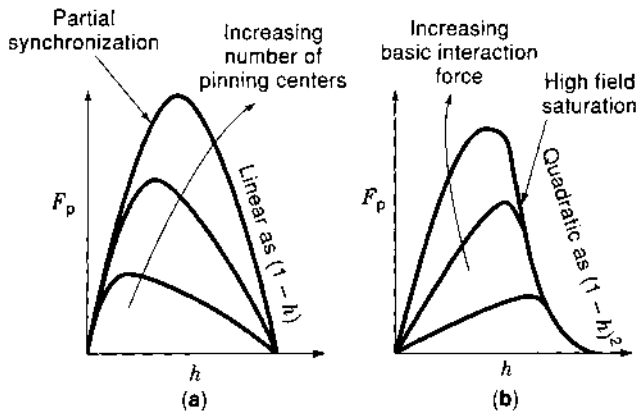
Because many superconductors exhibit scaling, it is interesting to see if a theoretical model can predict the measurements and provide some insight into flux pinning behavior. Because the theoretical picture of summation is somewhat diverse, it is not surprising that the theory of pinning force density scaling is not completely clear. However, there are qualitative models that explain some of the experimental behavior.

Campbell and Evetts (7) examined the low field region of the pinning force curve, which is roughly linear with field for nearly all of the scaling models and experiments. In this region, there are a small number of flux lines compared to the number of pinning centers. Campbell and Evetts propose that direct summation should apply because the spacing between flux lines is large enough that the interactive forces between them are weak. As the field is increased from zero, the bulk pinning force density increases linearly because of the increased number of flux lines being pinned. This is often referred to as the "partial synchronization" range of fields because the flux lines become "synchronized" with the pinning center array.

At some magnetic field, the number of flux lines is equal to the number of pinning centers, and the maximum bulk pinning force density is reached. Therefore, the field of the pinning peak depends on the number of pinning centers, and the magnitude of the peak depends on the strength of the basic interactive forces (Fig. 15).

At higher fields, there are more flux lines than pinning centers, and one expects a crossover from synchronization to where the bulk pinning force is limited by other effects. In the Campbell and Evetts model the high field pinning force falls off because of the variation of the basic interactive force with field, which falls as  $(1 - h)$  for the core interaction. Thus, Campbell and Evetts predict a linear low field region, a pinning peak at a field determined by the number of pinning centers, and a linear high field region. As the basic interactive force increases, the pinning force curve increases in all fields.

An alternative model of summation due to Kramer (25) describes the low field portion of the pinning force curve by partial synchronization, as in the Campbell and Evetts model. However, the high field behavior is determined by the shear of the flux lattice past strongly pinned individual fluxons. Therefore, the high field pinning force has a magnetic field dependence determined by the  $C_{66}$  elastic constant, which from Eq. (20) decreases as  $(1 - h)^2$ . Kramer also predicts that the high field behavior will exhibit saturation such that variations in processing



**Figure 15.** In the summation model of Campbell and Evetts, the pinning force curve is linear at both low fields and high fields, and the position of the peak pinning force density shifts depending on the number of pinning centers in the superconductor (a). The Kramer model, on the other hand, shows a quadratic high field behavior, and the pinning force density saturates at high fields, as the strength of the basic interaction increases (b).

leading to changes in the basic interactive forces will affect the magnitude and position of the peak but will not affect the high field pinning force behavior (Fig. 15).

In most flux pinning theories the basic interactive force is a function of temperature. For instance, in the core pinning model, the basic interactive force depends on the condensation energy and therefore on  $H_C$ . As the temperature changes, so does  $H_C$ . This leads to the observed temperature scaling of  $F_p$ .

However, the temperature dependence is more complicated for some basic interactions. For example, if the pinning center were a superconducting precipitate with  $T_C$  and  $H_{C2}$  below that of the bulk material, one would expect a difference in the strength of the core pinning interaction as the temperature is varied above and below the pinning center critical temperature and as the field moves above and below  $H_{C2}$  of the pinner. This effect is normally observed as a lack of scaling and commonly as a shift of the peak in the pinning force curve as a function of temperature.

A second example of a lack of scaling is a superconductor in which the pinning force on the FLL is a combination of several different basic interactive mechanisms. Such a superconductor might be a two-phase alloy in which pinning results from both core interactions with normal precipitates and grain boundary pinning. The different temperature and field dependencies of the two operating pinning mechanisms lead to a lack of scaling (34).

The main points to understand from this overview of the flux pinning mechanism are the following:

- Bulk flux pinning depends on the basic interactive forces between individual flux lines and individual pinning centers.
- Bulk flux pinning depends on the relative strength of the basic interactive forces and the fluxon-fluxon forces, which affects the summation of the individual interactions into the bulk pinning force acting on the FLL.
- Scaling, or lack of scaling, provides a tool for understanding the pinning mechanisms operating in different field

and temperature regions. This understanding can help direct modifications of the pinning microstructures by using suitable processing to optimize the pinning force and  $J_C$  of hard superconductors.

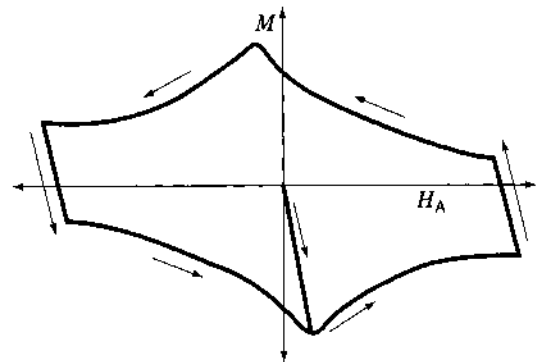
## THE CRITICAL STATE MODEL OF MAGNETIZATION

The previous discussion has centered on the electrical behavior of the superconductor, but magnetic behavior is also an important aspect of many applications. The magnetic response of the superconductor can be a valuable tool for measuring the critical current density.

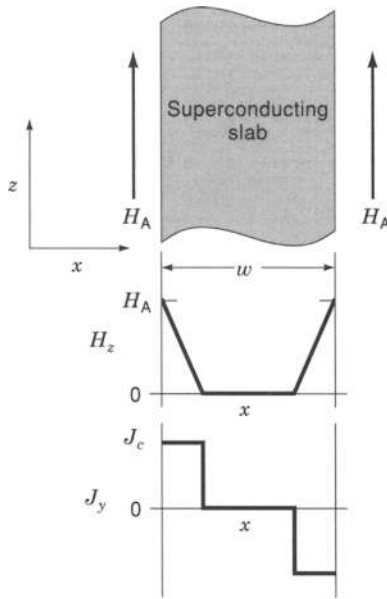
An important consequence of pinning the magnetic FLL is that the magnetic behavior of hard superconductors is strongly hysteretic (Fig. 16). To understand the development of the hysteretic magnetization curve, a simple but powerful model was proposed by Bean (35) and has since been modified and further refined (36).

To simplify discussion of the model, we use a sample geometry of an infinite superconducting plate of thickness  $W$ . The applied magnetic field will be parallel to the surfaces of the plate, as in Fig. 17. As the magnetic field is increased from zero, superconducting magnetization currents develop on the surfaces, so that they shield the interior of the superconductor from the applied field (the Meissner-Ochsenfeld effect). These shielding currents flow only within a distance  $\lambda$  (the penetration depth) of the surface and fall off exponentially into the superconductor. This condition persists until the external applied field exceeds the lower critical magnetic field  $H_{C1}$ . For fields larger than  $H_{C1}$  the superconductor is thermodynamically more stable if the applied field enters the superconductor as the flux line lattice. For simplification, the Meissner state below  $H_{C1}$  is ignored in the critical state model. This is not a bad approximation, especially for the technological superconductors which have large values of the Ginzburg-Landau parameter  $\kappa$  and therefore small values of  $H_{C1}$ .

If the sample has pinning centers to hold the entering FLL in place, a magnetic field gradient is established at the surface of the superconductor which falls off into the body of the sample. From earlier discussions of Ampere's law, we know that the current density flowing in a superconductor is directly related to the magnetic field gradient, Eqs. (9) and (10). The situation



**Figure 16.** Schematic of the hysteretic magnetization curve in strong pinning superconductors. The arrows indicate the direction of travel around the hysteresis loop during a typical magnetization measurement.



**Figure 17.** In the Bean critical state model, the semi-infinite slab develops the flux and current density profiles shown on the application of a magnetic field larger than  $H_{C1}$ .

looks schematically like that shown in Fig. 17, where the  $z$ -directed applied field falls off into the sample with a gradient in the  $x$ -direction that produces a current density flowing in the  $y$ -direction (into the page).

The current density flowing is equal to the critical current density. If the flux gradient were steeper, the current density would be larger than the  $J_C$ , and the FLL would not be fully pinned, leading to flux flow. The flux motion lowers the field gradient until the FLL is pinned by the pinning centers, leaving a critical flux gradient and a current density equal to  $J_C$ . For this reason the model is known as the “critical state model.”

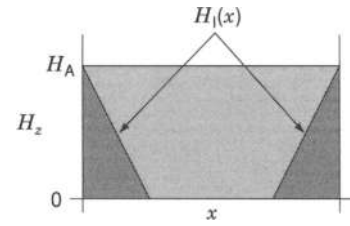
In Bean’s original version of the critical state model, the  $J_C$  is assumed to be a constant, independent of applied field from  $H_{C1} < H_A < H_{C2}$ . This assumption makes the flux profile in the sample linear such that

$$\frac{dH_z}{dx} = \frac{\Delta H_z}{\Delta x} = J_y = \text{constant} = J_C \quad (24)$$

As the applied field is increased from  $H_A = 0$ , the field penetrates the sample from both sides, and generates a circulating shielding current equal to  $J_C$  (Fig. 17). The magnetization of the slab can be found from examination of the field versus position plot. From the definition of magnetization, we know that the local magnetization response of the superconductor to the applied field can be written as the difference between the applied field and the local internal magnetic field:

$$M(x) = H_1(x) - H_A \quad (25)$$

where  $H_1(x)$  is the local internal magnetic field. To find the bulk magnetization, we must integrate the local magnetization over the sample volume. Because the sample is infinite in the  $y$ - and  $z$ -directions, we can turn this into a one-dimensional integral



**Figure 18.** The flux profile in the superconductor on applying a field  $H_A$ . This is useful for understanding the origin of the terms in the integral of Eq. (26). The magnetization response to the applied field  $H_A$  is proportional to the area of the light gray trapezoid.

over  $x$ , such that

$$\begin{aligned} M_{\text{bulk}} &= \left(\frac{1}{W}\right) \int H_1(x) dx - \left(\frac{1}{W}\right) \int H_A dx \\ &= \left(\frac{1}{W}\right) \int [H_1(x) - H_A] dx \end{aligned} \quad (26)$$

Comparing Eq. (26) with Fig. 18 shows graphically that the second term is the area of the entire rectangular region, whereas the first term is given by the area of the two darker triangular regions. The bulk magnetization is the volume averaged difference between these, or the light gray trapezoidal area of Fig. 18, divided by the sample width  $W$ .

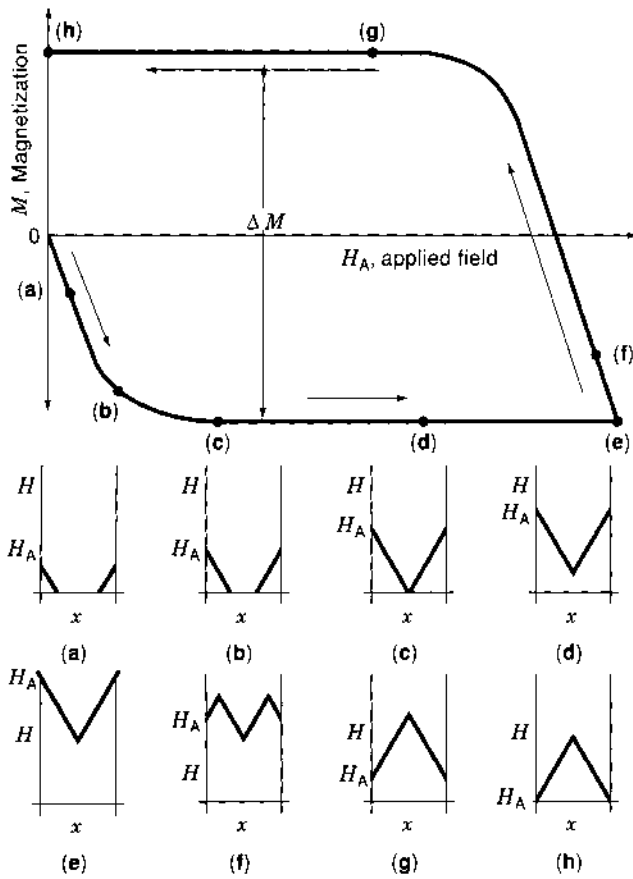
Using this simple model we can determine the behavior of the superconductor during a half magnetic field cycle used to generate a magnetization loop of  $M$  versus  $H_A$ . The process is shown schematically in Fig. 19. For small applied fields (points a, b) the field penetrates, and the magnetization increases rapidly with applied field. At point c the applied field is large enough to push the magnetic flux line lattice all the way to the center of the sample. This field is called the full penetration field  $H_P$ .

For applied fields larger than the full penetration field, the magnetization does not change, even though the internal and external fields increase. The magnetization curve remains flat (points d, e).

Now, if the applied field were to be reduced, the flux lattice, which is being pinned in place by the pinning centers, responds only near the surface region, as shown at point f. The magnetization becomes rapidly smaller with decreasing field. Now, the circulating supercurrents flowing in the sample have the spatial dependence shown in Fig. 20. Both the positive and negative flowing currents are assumed to be flowing at the critical current density. Recall that the magnitude of the critical current density in the Bean critical state model is constant with magnetic field.

As the applied field is further reduced, the current density profile eventually inverts, and the magnetization becomes positive and constant (g). Finally, at  $H_A = 0$ , the magnetization is positive because of the magnetic fields trapped in the body of the superconductor by the pinning forces acting on the FLL.

Because the flux gradient is a constant, the full penetration field  $H_P$ , varies with the width of the sample. Examination of Fig. 19 at the full penetration field (point c) shows that the

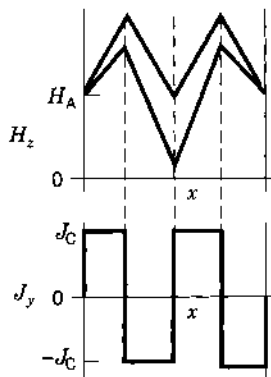


**Figure 19.** Schematic of a half hysteresis loop measurement of the magnetization in a Type II superconductor. The flux profiles at eight points around the magnetization loop are derived from the Bean critical state model and illustrate the source of the dependence of the magnetization on the pinning of magnetic flux within the superconductor.

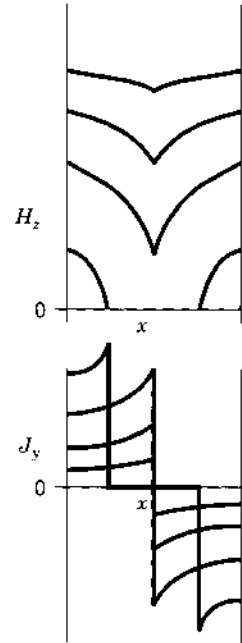
magnetic flux gradient

$$\frac{dH}{dx} = -\frac{H_p}{(W/2)} = J_c \quad (27)$$

We can also see that the magnitude of the magnetization at this



**Figure 20.** The flux profile and the accompanying current density profile for point (f) of Fig. 19. The current density profile matches the profile of the magnetic flux at all points in the superconductor.



**Figure 21.** The Bean critical state model assumes that the critical current density is constant with the magnetic field. In modified versions of the critical state model, the critical current density is allowed to vary with the magnetic field. The effects on the magnetic flux profiles and accompanying current density profiles are shown here for four different applied fields. As the applied field becomes larger, the critical current decreases. The slope of the flux profile changes with the magnetic field, and it is also no longer linear with position in the superconductor.

field is given by

$$M(H_p) = \frac{2}{W} \int \left[ \left( H_p + x \frac{dH}{dx} \right) - H_p \right] dx = \frac{J_c}{\left( \frac{W}{2} \right)} \int x dx \quad (28)$$

integrated from  $x = 0$  to  $W/2$ . This gives the constant value of the magnetization with applied field (points c, d, e) as

$$M = J_c \left( \frac{W}{2} \right) \quad (29)$$

Figure 19 also illustrates that, when one considers the entire hysteresis loop of the magnetization measurement, the distance between the increasing field and decreasing field magnetization at any applied field is twice the result of Eq. (29), or the more usual result from the critical state model,

$$\Delta M = J_c W \quad (30)$$

The Bean critical state model has been modified to account for the fact that the critical current density is not a constant with applied field (36). These modifications lead to curved magnetic field and current density profiles (Fig. 21) in addition to more realistic magnetization loops (Fig. 16). It is still the case, however, that these models predict a direct relationship between the height of the magnetization loop at a given field and the critical current density multiplied by the sample dimension.

Other modifications to the Bean Critical State Model have incorporated the change in magnetization due to finite sample



sizes (37) and demagnetization factors for non-spherical samples (38).

The most important result of the Bean model is that the magnetization behavior can be used as a probe of the critical current density within the superconductor by using a technique different from the four-point resistive measurement. For many emerging superconducting materials, magnetization measurements have allowed measuring the critical current density before long lengths of wire were available for resistive testing. It is also possible to measure  $J_C$ s that would be difficult to measure with a conventional resistive technique. An example is a cabled conductor with an  $I_C$  of thousands of amperes. Multikiloamp power supplies are expensive to purchase and operate, and the high current significantly complicates the experimental design. Magnetization measurements of  $J_C$  are not limited by the need for high current power supplies. The magnetization measurement of  $J_C$  continues to be an important tool for the materials engineer in optimizing the flux pinning process.

### DISSIPATION EFFECTS IN HIGH CURRENT DENSITY SUPERCONDUCTORS

The movement of the FLL within the superconductor has many consequences for the applications of superconductors. Examples include flux jumping (the rapid movement of magnetic flux within the superconductor which leads to localized heating effects and the loss of the superconducting state), flux flow near the  $J_C$ , flux creep (the slow movement of the FLL caused by random thermal jumping of flux lines out of the pinning potentials), and magnetic hysteresis (which causes an additional heating effect and resistive loss in ac applications of superconductors). Because of their importance to applications, the dissipative effects have been carefully studied and have led to some useful insights into the flux pinning process.

#### Flux Flow and Resistive Transition Analysis

As we have seen previously, the transition from the flux pinning to the flux flow state is generally not sharp but occurs over a range of currents during a resistive critical current measurement. Several models for the shape of the resistive transition have been developed to account for this behavior (39–41), but they all assume a distribution of pinning center strengths within the wire. The idea was first proposed by Baixeras and Fournet (42) but was not fully developed and applied to technological superconductors until the 1980s. If one assumes that the superconducting wire is made of an assortment of independent, current-carrying segments in series, each with its own value of critical current (as determined by the flux pinning defects within each segment), then the equation for the  $V(I)$  curve can be written as

$$V(I) = A \int (I - I') f(I') dI' \quad (31)$$

where  $I'$  is the local critical current of an individual wire segment,  $A$  is a constant,  $f(I')$  is the critical current distribution of the segments in the wire, and the integral is performed from zero current to  $I$ .

In essence this states that the voltage at any transport current  $I$  results from the segments with  $I_C < I$  and are therefore in flux flow. Variations in the critical current of each segment may occur because of variations in processing, chemical

inhomogeneities, or changes in geometry. For wires that are very homogeneous along their lengths, the distribution in critical currents  $f(I)$  is very narrow, and the  $V(I)$  characteristic is quite sharp and steep, leading to large  $n$ -values and the associated measure of “high-quality” in technological superconductors (43). At the other extreme, a large variation in the critical current distribution leads to a broad  $V(I)$  transition and a low  $n$ -value.

The importance of the model for helping the development of high  $J_C$  superconductors is that it can be used in reverse to determine the critical current distribution function from a measurement of the  $V(I)$  curve. By taking the second derivative of both sides of Eq. (31),

$$\frac{d^2V}{dI^2} = Af(I) \quad (32)$$

The second derivative of the experimentally determined  $V(I)$  curve yields the distribution  $f(I)$  of critical currents within the sample. In practice it is found that the resistive critical current measurement is always at a current well below the peak of the critical current distribution  $f(I)$ . Technological superconductors are limited by the weakest flux pinning segment along the sample length.

#### Temperature Dependence: Flux Creep and FLL Melting

As the temperature of a strongly pinned superconductor is increased, the thermal energy available to the pinned flux lines increases, allowing the possibility of thermally induced flux depinning. The higher temperatures lead to thermal excitation of the flux lines within the pinning potential wells. In low temperature superconductors where the thermal energies are low, the thermal energy is typically much smaller than the depth of the pinning potential. This led early investigators to name the effect “flux creep” because of the similarity to mechanical creep of crystalline materials at stresses much lower than the yield stress.

The theory and model for flux creep was developed by Anderson (44) and elaborated by Kim and Anderson (45). The basic premise is that flux lines at any finite temperature, experience a thermal excitation due to the thermal energy  $kT$ . The excitation energy has a statistical probability of causing a flux line pinned by a pinning center to become unpinned. The probability of this thermal depinning occurring is given by the Arrhenius factor

$$p = \exp\left(-\frac{U}{kT}\right) \quad (33)$$

where  $U$  is the depth of the energy well of the pinning center (Fig. 10).

For an isolated flux line, assuming no temperature gradient, there is no preferred jump direction because both the pinning potential and the thermal excitation are spatially symmetrical. Thus the flux line would execute a random walk as the thermal excitation allowed it to leave the pinning potentials, and no net change in flux would occur.

For a flux line in a magnetic field gradient, the flux motion associated with the thermal activation of the flux line derived

from a diffusion argument is

$$\frac{dB}{dt} = Aw_0 \exp\left(-\frac{U}{kT}\right) \quad (34)$$

where  $A$  is a factor that depends on the local magnetic field and field gradient and  $w_0$  is the frequency with which the flux line tries to jump out of the well.

The flux gradient provides a driving force that leads to diffusion of the flux line preferentially down the flux gradient. This produces a changing magnetic flux density according to Eq. (34) which can be experimentally measured. The flux gradient can exist because of the presence of either a transport current or a magnetization current in the sample.

In the case of a magnetization-induced flux gradient, the flux gradient would decay, eventually disappearing, even though the flux gradient (or magnetization current density) is less than the critical current density. This is in contrast to the assumptions of the Bean critical state model, in which the flux gradient adjusts itself to match the critical current density at all points in the sample. In the original experimental work on flux creep (46), the measured decay of the magnetization currents translated into a decay time of  $10^{92}$  years, effectively infinite time, so that, even with flux creep, the persistent magnetization supercurrents that flow as a result of the Meissner–Ochsenfeld effect are truly persistent.

The situation is somewhat different when the flux gradient driving force is produced by an externally applied transport current. In this case, the flux gradient crosses the superconductor, and the effect of the flux creep is to move flux lines across the sample in a manner identical to the flux depinning process at  $J_C$ . In the same way that the FLL motion causes a resistive-like dissipation in the superconductor, the flux creep motion also contributes to the dissipation. This results in a power loss in the superconductor that must be supplied by the external power supply and a consequent heating of the superconductor caused by the flux motion. If the flux motion is too great, the temperature increases, which increases the probability of thermally activated flux motion, and a thermal runaway to the normal state ensues.

At higher temperatures, the thermal excitation of the flux lines becomes comparable to the pinning potential, and the FLL is expected to move more easily. This, in fact, was one early argument against the possibility of technologically important superconductors at high temperatures. The argument was that no known basic pinning interactions are strong enough to prevent thermal excitation of the FLL at high temperatures (e.g., at 77 K, liquid nitrogen), even at very small transport currents. Fortunately for high temperature superconductors, this has not been the case. However, the concerns about the higher thermal energy available for excitation of the FLL out of the HTS pinning centers led along a path different from the original flux creep models into the theories of flux lattice melting.

Much of the theoretical picture of thermal effects on the FLL in HTS materials is at heart the same as in low temperature materials. The basic interests are in the possibility of scaling the pinning force with temperature, magnetic field dependence of the  $J_C$ , and the importance of the higher thermal energy in the depinning and motion of the FLL. The terminology has developed differently, but the physical mechanisms of flux pinning and depinning do not differ substantially.

## SUMMARY AND CONCLUSIONS

The critical current density of superconducting materials is a technologically important parameter, often more important in engineering design than the  $T_C$  or the  $H_{C2}$ . The theory of the  $J_C$  is broad and has provided useful insights into the physical mechanisms responsible for the large values of critical current density found in technological superconductors. This article has described the standard resistive and magnetic measurements of critical current. The theories of superconducting critical current range from the ultimate limits to the  $J_C$ , as described by the depairing current density, to the limits imposed by the thermal excitation of flux lines out of their pinning centers. The development of a bulk pinning force from the summation of individual pinning center interactions leads to the empirically observed scaling laws for the  $J_C$  that are important both as a tool for understanding the flux pinning processes and as an extrapolating technique useful for technological design.

Ultimately, this article has attempted to show the connection between the measured values of the  $J_C$  and the microstructural defects responsible for pinning the magnetic flux-line lattice in place. Understanding the physical processes underlying the development of the  $J_C$  has led materials engineers to design fabrication and processing that produces the desired microstructural features necessary for high  $J_C$ . In the past 10 to 15 years, this has led to massive increases in the performance of superconducting materials for applications requiring high current density and promises additional improvements in the future.

## BIBLIOGRAPHY

1. M. N. Wilson, *Superconducting Magnets*, Oxford, UK: Oxford University Press, 1983.
2. D. C. Larbalestier, in S. Foner and B. B. Schwartz (eds.), *Superconductor Materials Science: Metallurgy, Fabrication and Applications*, New York: Plenum Press, pp. 133–200, 1981.
3. M. Suenaga, in S. Foner and B. B. Schwartz (eds.), *Superconductor Materials Science: Metallurgy, Fabrication and Applications*, New York: Plenum Press, pp. 201–274, 1981.
4. A. C. Rose-Innes and F. H. Rhoderick, *Introduction to Superconductivity*, Oxford, UK: Pergamon, 1969.
5. M. Tinkham, *Introduction to Superconductivity*, New York: McGraw-Hill, 1975.
6. T. P. Orlando and K. A. Delin, *Foundations of Applied Superconductivity*, Reading, MA: Addison-Wesley, 1991.
7. A. M. Campbell and J. E. Evetts, *Adv. Phys.*, **21**: 199–428, 1972.
8. H. Ullmaier, *Irreversible Properties of Type II Superconductors*, Berlin: Springer-Verlag, 1975.
9. Y. B. Kim, C. F. Hempstead, and A. R. Strnad, *Phys. Rev.*, **139** (4A): 1163–1172, 1965.
10. F. Volker, *Particle Accelerators*, **1**: 205–207, 1970.
11. H. K. Olsson et al., *Phys. Rev. Lett.*, **66** (20): 2661–2664, 1991.
12. N.-C. Yeh et al., *Phys. Rev. B*, **45** (10): 5710–5713, 1992.
13. T. S. Kreilick, Niobium-Titanium Superconductors, in *Metals Handbook*, Vol. 2, 10th edition, Materials Park, OH: ASM International, pp. 1043–1059, 1990.
14. F. B. Silsbee, *J. Wash. Acad. Sci.*, **6**: 597, 1916.
15. A. A. Abrikosov, *Sov. Phys. JETP*, **5**: 1174, 1957.
16. U. Essman and H. Trauble, *Phys. Lett.*, **24A**: 526–527, 1967.
17. J. D. Cribier et al., *Progress in Low Temperature Physics*, Vol. V, C. J. Gorter (ed.), Amsterdam: North-Holland, 1967, p. 161.

18. G. Zerweck, *J. Low Temp. Phys.*, **42** (1): 1–9, 1981.
19. A. Gurevitch and L. D. Cooley, *Phys. Rev. B*, **50** (18): 13563–13576, 1994.
20. L. D. Cooley, P. J. Lee, and D. C. Larbalestier, *Phys. Rev. B*, **53** (10): 6638–6652, 1996.
21. H. Trauble and U. Essmann, *Phys. Stat. Sol.*, **25**: 395–402, 1968.
22. J. Schelten, H. Ullmaier, and G. Lippmann, *Phys. Rev. B*, **12** (5): 1772–1777, 1975.
23. E. H. Brandt, *J. Low Temp. Phys.*, **21**: 709, 1977.
24. E. J. Kramer, *J. Appl. Phys.*, **44** (3): 1360–1370, 1973.
25. E. J. Kramer, *J. Electron. Mater.*, **4** (5): 839–881, 1975.
26. A. I. Larkin and Yu. N. Ovchinnikov, *J. Low Temp. Phys.*, **34**: 409–428, 1979.
27. R. Labusch, *Cryst. Lattice Defects*, **1**: 1–16, 1969.
28. W. A. Feitz and W. W. Webb, *Phys. Rev.*, **178** (2): 657–667, 1969.
29. M. P. A. Fisher, *Phys. Rev. Lett.*, **62** (12): 1415–1418, 1989.
30. B. Brown et al., *Phys. Rev. B*, **55** (14): 8713–8716, 1997.
31. D. Dew-Hughes, *Philos. Mag.*, **30**: 293–305, 1974.
32. D. G. Hawkesworth and D. C. Larbalestier, *Proc. 8th Symp. Eng. Probl. Fusion Res.*, 1979, Vol. 1, p. 245.
33. R. J. Hampshire and M. T. Taylor, *J. Phys. F*, **2**: 89–106, 1972.
34. C. Meingast, P. J. Lee, and D. C. Larbalestier, *J. Appl. Phys.*, **66** (12): 5962–5970, 1989.
35. C. P. Bean, *Phys. Rev. Lett.*, **8**: 250–253, 1962.
36. D.-X. Chen and R. B. Goldfarb, *J. Appl. Phys.*, **66** (6): 2489–2500, 1989.
37. R. B. Goldfarb and J. V. Minervini, *Rev. Sci. Instrum.*, **55** (5): 761–764, 1984.
38. M. Daumling and D. C. Larbalestier, *Phys. Rev. B*, **40** (13): 9350–9353, 1989.
39. J. E. Evetts and C. J. G. Plummer, *Proc. Intl. Symp. Flux Pinning Electromagnetic Properties Superconductors*, T. Matsushita, (ed.), Fukuoka, Japan: Matsukama Press, 1986, p. 150.
40. D. P. Hampshire and H. Jones, *Proc. 9th Intl. Conf. Magnet Technology*, C. Marinucci (ed.), Swiss Institute for Nuclear Research, Geneva, Switzerland, 1985, p. 531.
41. W. H. Warnes and D. C. Larbalestier, *Appl. Phys. Letts.*, **48** (20): 1403–1405, 1986.
42. J. Baixeras and G. Fournet, *J. Phys. Chem. Solids*, **28**: 1541, 1967.
43. W. H. Warnes and D. C. Larbalestier, *Cryogenics*, **26** (12): 643–653, 1986.
44. P. W. Anderson, *Phys. Rev. Letts.*, **9** (7): 309–311, 1962.
45. P. W. Anderson and Y. B. Kim, *Rev. Mod. Phys.*, **36**: 39, 1964.
46. Y. B. Kim, C. F. Hempstead, and A. R. Strnad, *Phys. Rev. Lett.*, **9** (7): 306–311, 1962.

WILLIAM H. WARNES  
Oregon State University

## SUPERCONDUCTIVITY: ELECTROMAGNETS

The design and construction of superconducting magnets has been made possible by the development of technical superconductors. There are three principal materials: the alloy niobium-titanium (NbTi), the intermetallic compound niobium-tin (Nb<sub>3</sub>Sn), and the collective high-temperature superconductors (HTS) based on copper oxide layers in a perovskite structure. Nb<sub>3</sub>Sn and NbTi were developed in the early 1960s. A small solenoid made of Nb<sub>3</sub>Sn achieved a field of 10 T in 1963,

and the alloy NbTi was also first examined as a wire for a superconducting magnet in that year. In the nearly four decades that have elapsed since those first steps, superconducting magnets weighing up to many hundreds of tonnes have been built and operated.

Technical superconductors are a class having special properties that allow superconductive operation in high magnetic fields and with useful current densities. Superconductors are of two types; both can support the flow of electrical current without resistance only below a combination of maximum temperature, field, and current density, the critical parameters being  $T_c$ ,  $B_c$ , and  $J_c$  (1). For Type I superconductors, typical values of  $T_c$  and  $B_c$  are 9 K and 0.1 T, respectively. Flux is excluded from the bulk of a Type I superconductor, and current flows only in a surface layer, about  $10^{-4}$  mm thick. The current is carried by electron pairs of equal and opposite momentum. (This momentum includes terms arising from both velocity and linked flux.) Since flux is excluded from the interior of a Type I superconductor, and current flows only on the surface, a useful overall current density can only be obtained in wires of a diameter unacceptably small for magnet construction. Consequently a Type I superconductor cannot be used to construct a useful magnet, even were the critical field to be higher than 0.1 T. Examples of Type I superconductors are lead, indium, mercury, and pure annealed niobium.

By contrast, Type II superconductors allow flux to penetrate into the bulk of the lattice in the form of an array of flux quanta. A circulating supercurrent surrounds each flux quantum ( $2.07 \times 10^{-15}$  V · s) (2). If a net transport current flows along a Type II superconducting wire, a density gradient of flux quanta exists across the wire. The flux density gradient is given by

$$\frac{dB}{dx} = -\mu_0 J_c \quad (1)$$

where  $J_c$  is a bulk critical current density and  $\mu_0$  is the permeability of free space.

However, a Type II superconductor is not necessarily a technical superconductor. The interaction of the net transport current flow with the flux quanta creates a Lorentz force that causes a migration of the flux across the superconductor. Such a migration appears as a voltage gradient along the superconductor, which causes a dissipation; that is, the superconductor appears to be resistive. For these reasons pure annealed Type II superconductors are resistive with a transport current in the presence of a field. Because of the penetration of flux, the maximum temperatures and fields below which superconductivity arises in them are up to 18 K and 25 T, respectively. The latter is the upper critical field,  $B_{c2}$ , at which the lattice is fully packed with flux quanta. The penetration of flux into the lattice begins at a lower critical field,  $B_{c1}$ .

The technical superconductor is a Type II superconductor in which this migration of flux quanta is inhibited. This is achieved by the introduction of lattice defects, in the form of crystal boundaries, impurities, vacancies, or dislocations (3). At these defects flux quanta tend to become pinned against further movement, at least until the Lorentz force increases.

To the extent that mechanical models assist in the understanding of quantum effects, a typical pinning center may be imagined as a long hole through the lattice of the superconductor. If a flux quantum, or an integral number of flux quanta,

were located in this hole, the associated circulating current would lie in the material around the hole. If now the flux quantum were to move away from the hole, the current would have to become elongated so as to continue to flow in solid material. That would increase the interactive force between flux and current, so tending to force the flux back into the hole. The movement of flux quanta through an array of pinning centers is also influenced by thermal vibration. Since the temperature at which electron pairs can form is low, thermally induced lattice vibration is comparable in energy density with that of the pinning strength of the centers. Thus a flux quantum trapped at a pinning center may be dislodged by a large enough random thermal movement. That flux quantum would then migrate down the Lorentz force gradient to another pinning center. On the macroscopic scale this effect will reduce the flux density gradient and hence the Lorentz force. So, under conditions of constant transport current and external field, the random movement of flux quanta will die away until the current, field, or temperature changes. This process is known as flux creep. It is central to the manner in which technical superconductors can be used in magnets (4).

The continuous movement of flux across an annealed Type II superconductor appears as a voltage gradient along it. Similarly the process of flux creep through a field of pinning centers creates a voltage gradient, and through the interaction with the transport current, it gives rise to dissipation. The faster the flux creep, the greater is the local dissipation. Thus with flux creep is associated a local temperature rise. Since the specific heat of materials is small at low temperatures, the dissipation associated with flux creep can cause significant temperature rise. (At room temperature the energy of dissipation through flux creep would generate little temperature rise.) Thus the dissipative effect of flux creep can lead to a runaway flux movement. As the temperature rises, the critical current density decreases. According to Eq. (1), that in turn leads to an increased flux movement and thence to more dissipation. In the extreme case the temperature rises above the critical value, and superconductivity ceases. This process is known as a flux jump (5). In a magnet wound with superconducting wire, the sudden reversion to the normal state of a small length of the wire will result in the dissipation of the stored magnetic energy at high density in a volume that expands through thermal diffusion. If the stored energy is large enough and the thermal diffusion slow enough, the temperature rise at the point of initiation may exceed the melting point of the wire. The essence of the design of superconducting magnets is the suppression or control of this fundamental thermal instability.

## TECHNICAL SUPERCONDUCTORS

The major properties of the three technical superconductors are typically as shown in Table 1.  $T_c$  is the highest temperature below which superconductivity can exist (electron pairs can form) in zero field and with zero current.  $B_{c2}$  is the upper critical field at 0 K at which the lattice is fully packed with flux quanta.  $J_{c\text{typ}}$  is a typical operating current density at useful values of field and temperature. (Below the lower critical field,  $B_{c1}$ , the technical superconductor behaves as a Type I superconductor; that field is typically 0.01 T.) The two leading HTS materials are based on yttrium or on bismuth. The critical parameters refer to typical bismuth-based HTS.

**Table 1. The Critical Properties of the Principal Technical Superconductors**

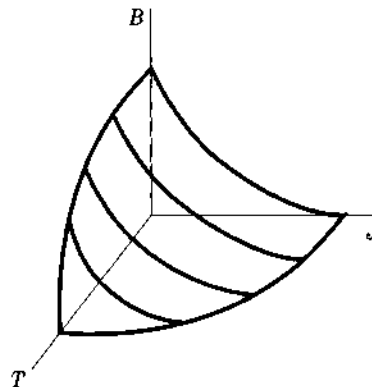
|   | NbTi | Nb <sub>3</sub> Sn | HTS  |
|---|------|--------------------|------|
| $T_c$ (K)                                 | 10   | 18                 | 90   |
| $B_{c2}$ (T)                              | 13   | 27                 | >100 |
| $J_{c\text{typ}}$ (A · mm <sup>-2</sup> ) | 1000 | 5000               | 1000 |

Figure 1 illustrates the typical variation of critical current density with field and temperature for the two low-temperature technical superconductors. Depending on the local values of any two of the parameters, the third has a local critical value. In a magnet conductor,  $B$  and  $J$  are usually fixed, so  $T$  is the critical parameter. The distance between the operating values  $B$ ,  $J$ , and  $T$  and the critical surface is the margin. In fact, since current and field are locally fixed, only the temperature can vary. Thus the temperature margin determines the susceptibility of the conductor to disturbance. The higher the margin, the more stable it will be. Since NbTi must operate at the lowest temperatures, it is the least stable. The HTS, by contrast, can operate with such high-temperature margins that they are very stable. However, that also leads to a problem of protection which must be addressed carefully in the future design of large HTS magnets.

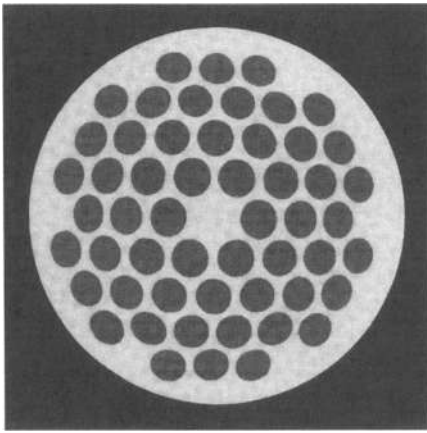
Despite its low-temperature margin NbTi is the most widely used technical superconductor. It is a ductile alloy, compatible in metallurgical processing with copper and cupro-nickel. However, Nb<sub>3</sub>Sn must be used in magnets or sections of magnets generating fields above about 12 T. HTS materials are now also beginning to be used in small magnets. Almost all magnets are wound with wire or a large cross-sectional conductor. (In principle, a magnet might be formed from a block of superconductor into which a field is frozen, e.g., by applying a field and then lowering the temperature. Instability makes that infeasible in low-temperature superconductors, but HTS may make such monolithic magnets possible.)

## Composite Superconductors

The suppression of instability and the protection of the conductor in the event of an instability are the predominant problems in the design of superconducting magnets. If a magnet is to operate reliably, the tendency for a runaway magnetothermal effect to trigger a full transition from the superconductive to



**Figure 1.** Variation of critical current density of a technical superconductor as a function of field and temperature.



**Figure 2.** Cross section of a typical niobium-titanium composite superconductor. (Courtesy of Vacuumschmelze GMBH.)

the normal conducting state must be suppressed. Two methods have been developed to achieve that: the elimination of flux jumping and the limiting of its consequences through cryogenic stabilization (5).

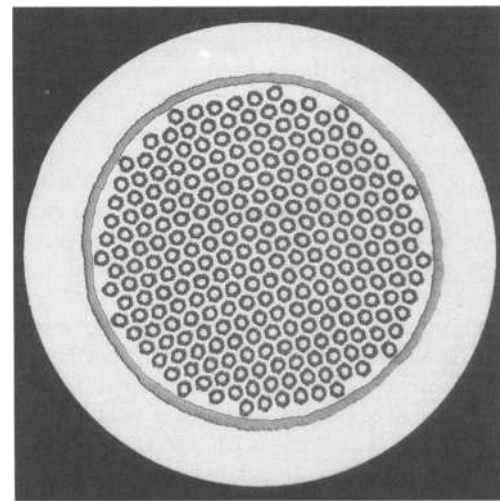
The essence of the stabilization of a technical superconductor is its subdivision into fine filaments. In the composite conductor this is achieved by embedding a large number of fine filaments in a matrix. The matrix is usually high-conductivity copper, although aluminum has also been used. The copper serves several functions. It rapidly conducts heat away from the surface of the fine filaments; since its resistivity is much lower than that of the superconductor in its normal state, it allows current to flow with relatively low dissipation when the adjacent superconductor has become normal through an instability.

#### Niobium-Titanium

Figure 2 shows a typical cross section through a NbTi composite superconductor. It is manufactured by a co-drawing process (6). Typical wire diameters lie in the range of 0.1 mm to 3 mm with critical currents at 4.2 K and 8 T of between 5 and 500 A. The insulation is typically Formvar (poly-vinyl-formal). The diameter of the superconducting filament is determined largely by stability but also by hysteresis loss. The latter is the integrated effect of dissipation by flux creep during the rise of current and field in a magnet. Hysteresis loss is minimized by decreasing the diameter of the filaments. Typically the number of filaments in a wire may lie between 10 and 100,000.

#### Niobium-Tin

$\text{Nb}_3\text{Sn}$  is a brittle compound, and its use in a magnet winding is considerably more complex than NbTi. The predominant method of forming the compound is the diffusion of tin into niobium at a high temperature (6). The tin is contained either internally within filamentary niobium tubes or in a tin-rich bronze surrounding solid niobium filaments. The niobium has typically up to 5% of tantalum for improved critical current density. The stock wire is unreacted for winding and is insulated with a glass braid, such as silica glass. After winding, the coil is reacted at about 700°C for up to 200 h. During this reaction the tin diffuses into the niobium and forms the compound  $\text{Nb}_3\text{Sn}$ . For adequate diffusion the diameter of the



**Figure 3.** Cross section of a typical niobium-tin superconductor. (Courtesy of Vacuumschmelze GMBH.)

niobium filament must be no greater than 0.005 mm. After reaction the wire is brittle and cannot tolerate bending. Therefore the winding must be consolidated by vacuum impregnation with epoxy resin. Figure 3 shows a typical cross section through a bronze route superconducting wire before reaction.

As in the NbTi composite, copper is required as a stabilizer and for protection. In order to preserve the low resistivity of the copper, tin must not be allowed to diffuse into it during the reaction. (That would, in any case, dilute the tin available for diffusion into the niobium.) A barrier is inserted between the bronze and the copper matrix to prevent such unwanted diffusion. This barrier is in the form of a thin wrap of tantalum or niobium.

#### High-Current Conductors

To form high-current conductors in either NbTi or  $\text{Nb}_3\text{Sn}$ , cables of small wires are usually embedded in a copper or aluminum matrix. In this way NbTi composite conductors with critical currents of up to 100,000 A can be constructed. Similarly high-current  $\text{Nb}_3\text{Sn}$  composites can be constructed by carefully cabling pre-reacted wires and soldering the cable into a copper matrix. Handling such pre-reacted  $\text{Nb}_3\text{Sn}$  conductors requires exacting quality control procedures (7).

An important form of high-current conductor is the so-called cable-in-conduit (CiC). A cable of composite strands is inserted into a closed sheath, usually of stainless steel. Liquid helium (sometimes in the superfluid state at a temperature below 2.17 K) is circulated through the gaps between the strands of the cable (8). This arrangement combines good support of the strands against the Lorentz forces with a low shielding loss and high current (see Section 6).

A recent development in  $\text{Nb}_3\text{Sn}$  conductors is a pre-reacted cable of very fine wires. The strands are so thin that bending radii appropriate to normal winding techniques can be tolerated (9).

#### HIGH-TEMPERATURE SUPERCONDUCTORS

At present only a very few superconducting magnets use these materials. These superconductors all contain copper-oxide

layers in a perovskite crystal structure (10). Two major types are being actively developed for application to magnets: yttrium based and bismuth based.

### Yttrium HTS

This material has the chemical composition  $\text{YBa}_2\text{Cu}_3\text{O}_7$ , abbreviated to YBCO. It has a critical temperature of 80 K. Its critical field at 4.2 K has not been directly measured but is believed to be about 100 T. Wire is manufactured in the form of a tape, samples of which have achieved a critical current density of  $7000 \text{ A}\cdot\text{mm}^{-2}$  at 5 T and 77 K. Lengths of wire suitable for the construction of magnets have not yet been produced. Their manufacture suffers from the disadvantage of a highly anisotropic critical current density. That is, the critical current depends strongly on the orientation of the local field relative to the crystal orientation. Nevertheless, magnets operating at liquid nitrogen temperatures will probably use this material in the near future.

### Bismuth HTS

This material is manufactured in two forms, with chemical compositions of  $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_7$ , abbreviated to 2212, or  $(\text{BiPb})_2\text{Sr}_2\text{Ca}_2\text{Cu}_3\text{O}_7$ , 2223. Both of these are being manufactured in lengths of up to 1000 m, and small magnets have been wound from them. The 2212 requires lower temperatures for high-field operation than the 2223. However, it is more easily fabricated and coils of the 2212 have been tested in external fields to show that useful currents can be carried at 30 T and 4.2 K. Either of these BSCCO materials may soon be the choice for inserts in very high field solenoids.

## STABILITY

### Adiabatic Stability

Regrettably, the flux jump is not the only source of heating sufficient to raise the temperature above the local critical value. Movement of a wire under the influence of the Lorentz forces in the body of a winding can generate local frictional heat sufficient to raise the temperature of a wire above the local critical value, especially where the temperature margin is small. A further precaution that must therefore be exercised in a magnet designed to have no flux jumping is the consolidation of the winding so that no sudden frictional movement can occur to generate heat. This is achieved by impregnating the winding with epoxy resin, or another agent, in order to prevent sudden movement between wires, or between wires and stationary structure such as coil forms. This is the form of winding of the adiabatic magnet, which is so termed because no thermal sink exists within the winding.

Despite the use of fine filaments and winding techniques to avoid mechanical disturbances, the adiabatic winding is susceptible to sudden local transition from the superconductive to the normal state, if only through failure of the cryogenic environment. The resulting spreading normal zone is a process known as quenching. It may be precipitated by the relief of local stresses, which may merely be the repositioning of a wire, or cracking in a resin impregnant, and it may result in quenching at a current and field below design values. Upon recooling and recharging, the next premature occurrence of quenching

can be at a higher current. This sequential improvement in the operating current is known as training. It is a common characteristic of large adiabatic magnets, especially of magnets with complex topology, such as particle-beam-handling magnets.

Despite the disadvantages associated with instability, quenching, and training, adiabatic magnets are the most common embodiment because they achieve the highest overall current densities in the windings. They enable the most compact magnets and the highest fields.

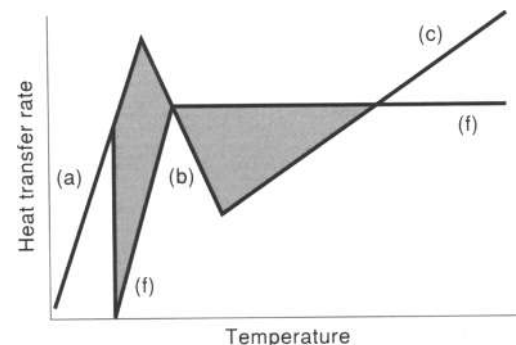
### Cryostability

A second method of stabilization exploits the powerful cooling effect of liquid helium to maintain the temperature of a conductor below the local critical value even in the presence of flux jumping or other severe perturbations. This is achieved by paralleling the superconductor with sufficient copper or aluminum cooled by liquid helium (11). Now, even though all the transport current were to flow in the copper, the temperature would be kept below the critical value of the superconductor if sufficient surface were presented to the helium. This is the principle of cryostability. The governing expression for this mode of operation of a superconductor is

$$\frac{I^2 \rho}{A} = hP \quad (2)$$

where  $\rho$  is the resistivity of the copper,  $A$  is the copper cross section,  $h$  is the heat transfer rate, and  $P$  is the wetted perimeter of the copper. If the value for  $h$  is chosen to be about  $1000 \text{ W}\cdot\text{m}^{-2}$ , this becomes a conservative stability criterion but leads to a low overall current density in a winding because the cross section of the stabilizing copper is large.

The earliest large magnets (mainly for bubble chambers built in the late 1960s) were designed in accordance with this principle. However, the boiling heat transfer curve for liquid helium has two branches, as shown in Fig. 4: nucleate boiling, line (a), with high heat transfer coefficient and film boiling; line (c) with lower coefficient ( $1000 \text{ W}\cdot\text{m}^{-2}$  is typically the lowest value of heat transfer rate in the film boiling regime). Intermediate between these is a region of unstable boiling, line (b). During an instability some or all of the current may transfer locally from the superconductor to the copper. The local generation of heat in such an extended region is line (f), rising from zero where no current flows in the copper to the full heating where all the current flows in the copper. The areas (d) and (e) then represent



**Figure 4.** Boiling heat transfer curve for liquid helium and the heat generation curve for a composite superconductor.

respectively regions where more heat is removed than generated and where more heat is generated than removed.

The sharing of current between the superconductor and copper results in the following expression for the local temperature:

$$\frac{I(I - I_s)\rho}{A} = q(\theta - \theta_b)P \quad (3)$$

where  $I$  is the total current,  $I_s$  is the current still in the superconductor,  $\rho$  is the resistivity of the copper,  $q$  is the heat transfer coefficient at the temperature difference  $\theta - \theta_b$  between the copper and the helium bath. A good approximation for the current in the superconductor at a temperature  $\theta$  is

$$I_s = \frac{I_c(\theta_c - \theta)}{(\theta_c - \theta_b)} \quad (4)$$

where  $I_c$  is the critical current at  $\theta_b$ , the bath temperature. Equations (2), (3), and (4) are represented together with the boiling heat transfer in Fig. 4. If the cooling exceeds the heating, area (d) greater than area (e), then recovery of superconductivity will occur. The normal region will collapse from the cold ends. This type of stability, called cold end recovery, allows higher values of  $h$  to be assumed in Eq. (2),  $3000 \text{ W}\cdot\text{m}^{-2}$  being commonly accepted. This in turn allows a smaller cross section of copper and a more compact winding (12).

## PROTECTION

The other crucial issue in the design of superconducting magnets is protection. In both the adiabatic and cryostable types, the consequences of quenching must be anticipated. The methods by which this is achieved are quite different in the two types.

### Adiabatic Protection

In this winding the conductors are in a compact array without an internal heat sink but with close thermal coupling between conductors. Therefore, when a conductor becomes normal through a perturbation, not only does the normal zone travel along the wire but the local heating rapidly causes adjacent conductors to heat up and become normal. Quenching then expands as a three-dimensional zone of increasing volume and surface area. The volume of normal conductor at any instant is approximated by

$$V_n \propto v_1 v_2 v_3 t^3 \quad (5)$$

where  $v_{1,2,3}$  are effective velocities of quench propagation along the wire and transverse to it and  $t$  is the elapsed time. Thus the magnetic energy stored in the field is dissipated in a rapidly expanding volume of conductor. The velocities of propagation depend, among other things, on the local temperature margin: The greater the margin, the slower are the velocities. Although not as high as the velocity along the wire, the dominant velocities are those between turns and between layers. Rapid propagation of quenching between turns is desired in order to spread the temperature rise over a large volume. For that reason HTS winding may be more difficult to protect than those with the smaller temperature margins of  $\text{Nb}_3\text{Sn}$  or  $\text{NbTi}$ .

Two criteria are applied to this process to gauge the vulnerability of the magnet winding to quenching: the final temperature of the wire at the point of initiation of quenching, and the voltage appearing between layers or turns of the winding. The latter criterion involves complex computation of the spatial distribution of normal and superconducting regions during the quench. However, the peak temperature generally dominates the criteria. This must clearly not exceed the allowable temperature for the insulation or impregnant nor produce unacceptable local stresses due to differential thermal expansion. A rule of thumb for acceptable peak temperature is 100 K. To achieve such limit (and simultaneously limit the voltage), sufficient copper must parallel the superconducting filaments so that the integrated heating in the copper does not exceed the acceptable peak value. Because of the rather complex time dependency of the current in a quenching adiabatic magnet, this criterion is best illustrated by reference to the protection of a cryostable magnet.

### Cryostable Protection

In this type of magnet no propagation of a quench zone occurs between adjacent conductors. Furthermore the one-dimensional propagation of quench along a conductor is slow. If, for example, the level of liquid helium is low so that a length of conductor is exposed to helium vapor and has quenched, no propagation will occur, and the normal zone will be stationary. In such a case protection of the conductor against excessive temperature rise must be by external discharge of the stored magnetic energy. Interturn voltage is hardly ever a dominant effect in the protection of a cryostable magnet. The limiting criterion is therefore the temperature rise of the conductor at the point of quench initiation. Just as for the adiabatic magnet, a generally accepted criterion is a rise of not more than 100 K. The prediction of the temperature rise is now rather simple. If the magnet is discharged exponentially, the governing equation is

$$J_{\text{Cu}}^2 \rho(\theta) dt = c(\theta) d\theta \quad (6)$$

The left-hand side of Eq. (6) represents the heat being generated in the resistance of the copper, and the right-hand side represents the same heat being absorbed by its enthalpy. The quantities  $\rho(\theta)$  and  $c(\theta)$  are the temperature-dependent resistivity and volumetric specific heat of the copper.

Equation (6) may be rewritten as

$$\int_0^\infty J_{\text{Cu}}^2 dt = \int_{4.2}^{\theta_{\text{max}}} \left\{ \frac{c(\theta)}{\rho(\theta)} \right\} d\theta \quad (7)$$

In this expression the right-hand side is a function of the copper, or some other conducting material, between a low temperature, generally assumed to be that of liquid helium boiling at atmospheric pressure, 4.2 K, and the allowable upper temperature, generally 100 K (13). (It is sometimes called the G-function but not in the original reference.) If the discharge of the magnet is exponential, Eq. (7) becomes

$$\frac{1}{2} J_{0\text{Cu}}^2 \tau = G(\theta_{\text{max}} - 4.2) \quad (8)$$

where  $J_{0\text{Cu}}$  is the initial maximum current density in the copper. For copper at 100 K, the value of the G function is typically

$6 \cdot 10^{16} \text{ A}^2 \cdot \text{m}^{-4} \cdot \text{s}$ . In order to achieve the required time constant of discharge, the voltage across the magnet must be as large as the inductance demands. As an example, a large magnet might have a stored energy of 100 MJ, an operating current of 10,000 A, and a copper cross section of  $500 \text{ mm}^2$ . The inductance would be 2 henries and  $J_{0 \text{ cu}}$  would be  $2 \cdot 10^7 \text{ A} \cdot \text{m}^{-2}$ . The time constant of discharge would have to be 17.32 s, and the maximum initial discharge voltage 1.154 kV.

### Winding Topology

The design of a winding array is dictated by the desired field properties. The basic expression governing the field produced by a winding is the Biot-Savart law, expressed as

$$dB = \frac{I \cdot dl \sin(\theta)}{r^2} \quad (9)$$

where  $I \cdot dl$  is a current element,  $\theta$  is the angle between that element and the radius vector  $r$  to the field point. The direction of the field vector is perpendicular both to the current element and the radius vector. This expression is integrated in closed form for the fields of solenoids and of long linear windings. The forward computation of field produced by a given current array is straightforward. It can be performed for any configuration of currents, although the closed form calculation is easy only for circular or straight arrays. The reverse computation, of a current array to produce a given field topology, is possible only for a few simple configurations, specifically for the axially symmetric field of a current ring and the circular field of an infinite straight current. The off-axis fields of a circular current array involve elliptic integrals. Numerical methods for the solution of the off-axis field of a current arc are routinely available as commercial codes. Most field geometries are produced by combinations of straight currents and current arcs. Generally the design problem proceeds by the iterative recomputation of the fields generated by incrementally modified straight or arc current arrays.

### OPERATION IN TIME-VARYING FIELDS

Although most superconducting magnets constructed to date have been designed for steady state operation, a few have been constructed for use either under ac conditions or with a rapidly changing field (14). The design of a conductor for such operation involves the reduction of dissipative mechanisms in the superconductor. One source of dissipation under conditions of changing field has been described, that is, flux creep. That loss is generally called hysteresis loss and is approximated per unit volume by the expression

$$Q \approx \frac{1}{4} \Delta B d_w J_c \quad (10)$$

where  $\Delta B$  is the field change,  $d_w$  is the wire diameter, and  $J_c$  is the critical current density. It is seen that the dissipation is smallest for thin wires.

Other sources of dissipation are found in composite superconductors. The most important of those is what is called shielding loss. It arises as follows: To reduce hysteresis loss and to improve the stability of a superconductor against flux jumping, the superconductor is finely divided into filaments in a (typically) copper matrix. However, it may be seen that the collection

of filaments will still tend to exclude flux unless the filaments are transposed. Full transposition of any but a single layer of filaments is impossible within the matrix. However, a compromise is the twisting of the conductor. A short twist pitch or high resistivity of the matrix will allow flux to penetrate through the layers of filaments (15). The time constant of penetration is given roughly by

$$\tau = \left( \frac{\mu_0}{2\rho} \right) \left( \frac{L}{2\pi} \right)^2 \quad (11)$$

where  $\rho$  is the resistivity of the matrix and  $L$  is the twist pitch. External field change will generate dissipation roughly given by

$$Q = \left( \frac{B^2}{2\mu_0} \right) \left( \frac{\tau}{\tau + t} \right) \quad (12)$$

where  $t$  is the time constant of the field change  $B$ . If the change is very rapid compared with the time constant of penetration of the conductor, the dissipation approaches the energy density of the magnetic field. Some magnets have been made with superconductors that have a matrix of cupro-nickel. The resistivity of cupro-nickel is high, which leads to a short time constant for flux penetration. However, the conductor is not well stabilized, and in the event of a quench it is not well protected against local hot spot. The CiC conductor, described above, also minimizes the shielding losses.

### FORCES AND STRESSES

The force on a conductor in a field is given by the expression,

$$F = BI \quad (13)$$

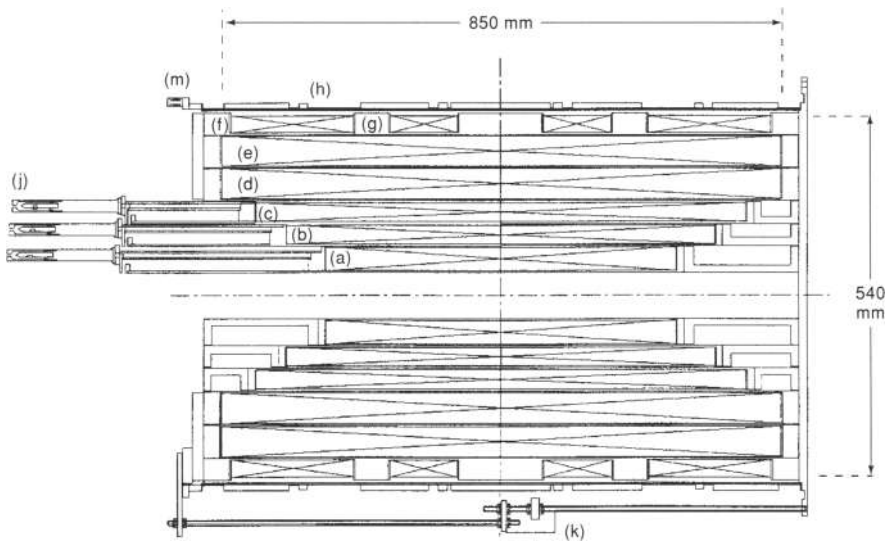
where  $I$  is the current in amperes,  $B$  is the component of the field in tesla perpendicular to the current vector, and  $F$  is the force per unit length in newtons which is directed perpendicular to the current and field vectors. The calculation of stresses in the conductors of a winding can be performed in closed form only for a few simple cases. For an isolated circular loop the hoop stress averaged over a section of wire (whose cross section may include superconducting filaments, copper, depleted bronze, etc.) is given by

$$\sigma = BJr \quad (14)$$

where  $B$  is the average field at the wire perpendicular to the plane of the loop,  $J$  is the current density averaged over the wire section, and  $r$  is the mean radius of the loop. If  $J$  is in  $\text{A} \cdot \text{m}^{-2}$ , and  $r$  is in meters, then  $\sigma$  is in pascals. In a multilayered solenoidal winding the stresses in the wire are influenced by the transmission of radial forces between layers and by axial forces generated by radial components of field at the ends of the solenoid. The analysis of these stresses usually is performed by standard programs. However, Eqs. (13) and (14) provide a useful rough guide to the stresses that will be encountered in a solenoidal winding.

The forces in long straight windings, such as extended dipoles, may be similarly estimated by the summation of the forces found from Eq. (13). The stresses in solenoidal windings





**Figure 5.** The winding profile of a typical commercial NMR magnet. (Courtesy of Magnex Scientific PLC.)

are often supported by the conductor alone. Since the stresses are predominantly tangential, they produce tension in the components of the wires. In NbTi composites the stresses tend to be divided roughly equally between the NbTi superconductor and the copper unless the latter is annealed. In Nb<sub>3</sub>Sn conductors most of the tensile load appears in the Nb<sub>3</sub>Sn superconductor itself because the Young's modulus is high, 172 GPa. The other components of a Nb<sub>3</sub>Sn composite tend to be annealed by the heat treatment.

### Examples of Superconducting Magnets

Common applications of superconducting magnets include the following: magnetic resonance imaging (MRI), laboratory magnets (for NMR analysis, susceptibility measurement), particle beam handling, subatomic particle analysis, power conditioning, energy storage, open gradient magnetic separation, magnetohydrodynamic power generation (MHD). NMR and MRI magnets are described in more detail in *MAGNETS FOR MAGNETIC RESONANCE ANALYSIS AND IMAGING*. All superconducting magnet are wound. The most common winding topologies are solenoids and extended linear windings. The latter are in general of the form of armature windings, having straight sections with end turns of complex shape. Solenoids are used in magnets for MRI, most laboratory magnets, particle analysis, and energy storage. Extended linear windings are used in particle-beam-handling magnets, in rotating machines (rotor windings), for particle focusing (quadrupoles, hexapoles), and in MHD. Occasionally a superconducting winding has an associated ferromagnetic flux path. (In such so-called superferric magnets, the iron is usually included in the magnet's cryogenic environment, to avoid the transmission of high forces between low temperature and room temperature.)

Although most superconducting magnets are solenoids or extended dipoles and multipoles, a few magnets of more complex shape have been manufactured. These include very large "Ying-Yang" coils for the mirror fusion test facility (MFTF) (16), saddle coils for experiments on magnetohydrodynamic power generation (MHD) (17) and toroidally shaped coils for several Tokamak fusion experiments (18). (Toroidal field magnets for fusion devices are D-shaped coils which are arrayed around a

circular plasma path; in the absence of orthogonal—poloidal—fields, the conductors of these coils experience pure tension as in a solenoid.)

All are large magnets, storing energy in the range of 100 MJ. The stored magnetic energy is a rough classification of the size and cost of a magnet, and not necessarily of the difficulty of design and construction. Very high field, compact magnets present more difficulty because of small operating margins, high stresses, and high-energy densities during quenching.

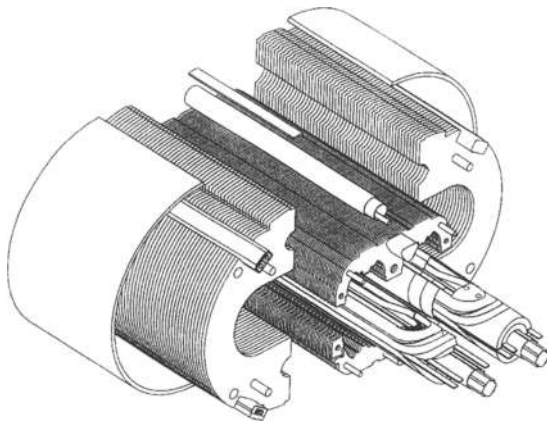
One of the largest superconducting magnets under construction is the model central solenoid for the International Thermonuclear Experimental Reactor (ITER). It stores 650 MJ and weighs 100 tonnes (19). The most precisely engineered have been the high-field NMR magnets. They now have reached fields of 18.8 T with homogeneity of  $10^{-9}$  in a 10 mm spherical volume.

NMR and MRI magnets operate in persistent mode. After energizing, a superconducting switch is closed across the winding so that the current continues to flow in an essentially resistanceless circuit. These are the most commonly produced commercial magnets (20).

Superconducting particle-beam-handling magnets have made possible the very high-energy accelerators in operation at BNL, CERN, DESY, and the Fermi Laboratory. Superconducting dipole magnets are typically 10 m long or greater, generating transverse fields of up to 8 T to provide the deflection required to bend the particle beam to a circular orbit. The requirement for uniformity of field is about  $10^{-4}$  over 25 mm width. Superconducting quadrupole magnets are used for focusing of the particle beam and typically provide field gradients of up to 25 T/m. >

**Table 2.** Major Characteristics of a 750 MHz NMR Solenoid

| Type           | Solenoid                    |
|----------------|-----------------------------|
| Materials      | NbTi and Nb <sub>3</sub> Sn |
| Energy         | 15 MJ                       |
| Center field   | 17.6 T                      |
| Bore at 4.2 K  | 74 mm                       |
| Typical weight | 250 kg                      |



**Figure 6.** An illustrative view of the LHC dipole construction, showing the windings and the structural yokes. (Courtesy of CERN.)

### Specific Examples

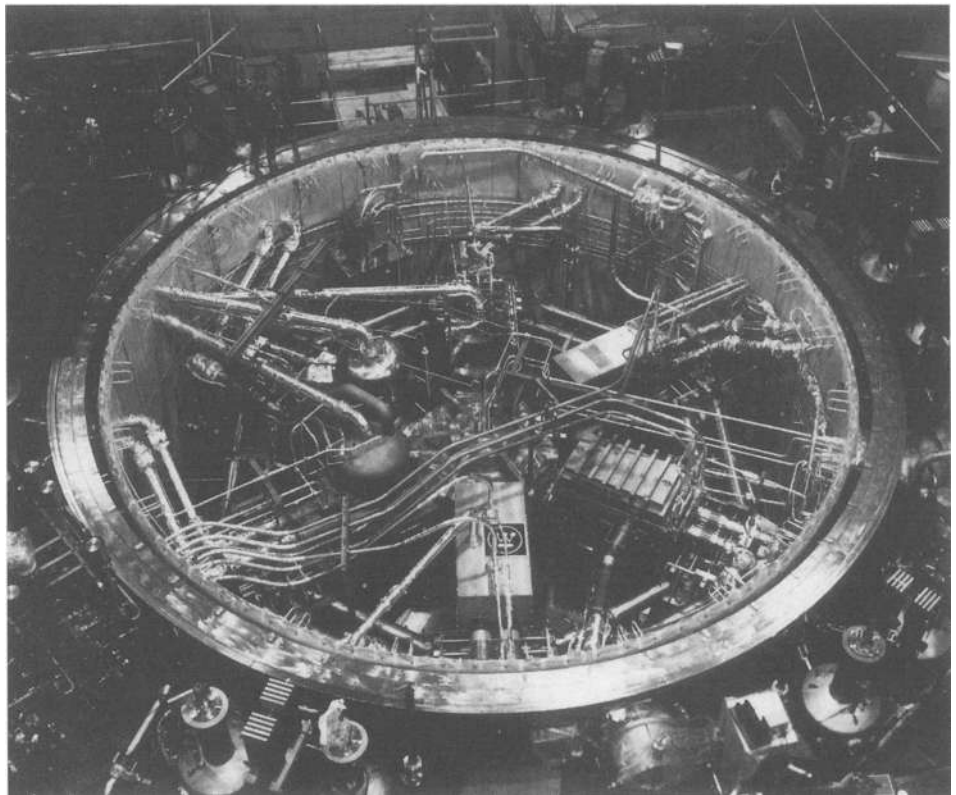
**NMR Magnet for 750 MHz Proton Frequency (21).** Superconducting magnets for NMR and MRI are described separately within this encyclopedia. However, as they typify solenoidal magnets the 750 MHz magnet is described here in outline. Figure 5 gives the profile of the windings of a commercial magnet along with some of the structural details. The global characteristics are listed in Table 2.

In the figure the windings are labeled (a) through (g). Windings (a), (b), and (c) are of  $\text{Nb}_3\text{Sn}$ . The wire size is largest in section a where the field is highest and the critical current density lowest. In fact section (a) consists of three grades of wire

in order to maximize the local winding current density. Sections (d) and (e) are the main  $\text{NbTi}$  winding, again graded to maximize overall current density. Sections (f) and (g) compensate the axial gradients introduced by the full-length sections. All magnets used for NMR are operated in persistent mode. The windings are connected in series through superconducting joints. After energizing, the windings are closed through a superconducting switch (k). The joints (j) between the  $\text{Nb}_3\text{Sn}$  sections provide the best performance (highest critical current) in low field. Therefore they are mounted away from the windings on posts that rigidly support the brittle conductor. Surrounding the windings is an array of shim coils (h) which provides for fine adjustments. Each shim set has its own switch (m).

**Particle Accelerators (22,23).** The Fermilab energy doubler/saver is typical of the use of superconducting beam-bending magnets used in particle accelerators. The field generated by the magnet is oriented vertically and extends over a length of 6 m. High-energy protons are slightly deflected by the field-length product of the magnet. The transverse force on the particle beam is as predicted by the Lorentz equation (13). A large number of these dipoles (774), quadrupoles (216), and sextupoles are arranged to produce a total deflection of  $360^\circ$ . In the Fermilab energy doubler each dipole magnet produces a mere  $0.47^\circ$  deflection at maximum beam energy.

A more recent project involving superconducting beam-handling magnets is the Large Hadron Collider (LHC) at CERN. This also uses a sequence of long dipoles, quadrupoles, and sextupoles in two adjacent rings in which protons, or other heavy particles, will circulate in opposite directions to collide at energies up to 14 TeV. The cross section of the winding and associated structure is shown in Fig. 6. The conductors are located



**Figure 7.** The six dee coils of the Large Coil Task installed in the test cryostat. (Courtesy of the Oak Ridge National Laboratory.)

**Table 3. Major Characteristics of the Coils in the Large Coil Task**

| <i>Global</i>     |                           |                |                |              |             |
|-------------------|---------------------------|----------------|----------------|--------------|-------------|
|                   | Design field              |                |                |              | 8 T         |
|                   | Total stored energy       |                |                |              | 600 MJ      |
|                   | Operating temperature     |                |                |              | 4.5 K       |
|                   | Inside height (each coil) |                |                |              | 3.5 m       |
|                   | Inside span (each coil)   |                |                |              | 2.5 m       |
|                   | Inside radius             |                |                |              | 1.25 m      |
| <i>Individual</i> |                           |                |                |              |             |
| Origin            | Conductor Type            | Inductance (H) | Peak Field (T) | Current (kA) | Energy (MJ) |
| EU                | NbTi FF                   | 1.57           | 9.01           | 11.40        | 200         |
| WH                | Nb <sub>3</sub> Sn FF     | 0.75           | 8.23           | 17.76        | 202         |
| GE/ORNL           | NbTi PB                   | 1.81           | 8.89           | 10.50        | 193         |
| GD                | NbTi PB                   | 1.81           | 7.72           | 10.20        | 136         |
| JAERI             | NbTi PB                   | 2.00           | 8.81           | 10.22        | 198         |
| CH                | NbTi FF                   | 1.00           | 7.86           | 13.00        | 123         |

so as to approximate a cosine distribution of current density. That arrangement generates a field uniform throughout the center region of each beam. The outward forces on the conductors are constrained by a stainless steel collar. This cold assembly is supported within an iron yoke which provides a return path for the flux. The yoke is assembled from transformer iron laminations, welded in a set of long steel girders.

**Large Coil Task (24).** Six coils were built, one each by the following groups: EURATOM (EU), The Swiss Institute for Nuclear Science (CH), The Japanese Atomic Energy Research Institute (JAERI), Westinghouse Electric, (WH), General Dynamics (GD), and General Electric/Oak Ridge National Laboratory (GE/ORNL). Each coil was a constant tension dee coil (18), but each used a different form of conductor. The Westinghouse coil alone used Nb<sub>3</sub>Sn, in a CiC configuration; all the others used NbTi in a cryostable (helium-cooled) configuration. The cooling method was either forced-flow (FF) or pool-boiling (PB). The purpose of the program was to demonstrate stability of the coils under conditions of combined toroidal and pulsed poloidal fields. This international program has resulted in the construction of a model toroidal coil array as shown in Fig. 7. The main specifications were as listed in Table 3.

## BIBLIOGRAPHY

1. A. C. Rose-Innes, *Introduction to Superconductivity*, 2nd ed., Amsterdam, The Netherlands: Elsevier Science, 1994.
2. H. Trauble and U. Essmann, Flux-line arrangement in superconductors as revealed by direct observation, *J. Appl. Phys.*, **39**: 4052-4057, 1968.
3. E. J. Kramer, Scaling laws for flux pinning in hard superconductors, *J. Appl. Phys.*, **44**: 1360-1370, 1973.
4. P. A. Anderson and Y. B. Kim, Hard superconductivity: Theory of the motion of Abrikosov flux lines, *Rev. Mod. Phys.*, **36**: 39-45, 1964.
5. M. N. Wilson, *Superconducting magnets, Monographs on Cryogenics*, Oxford: Clarendon Press, 1983, p. 131.
6. E. Gregory, Conventional wire and cable technology, *Proc. IEEE*, **77**: 1110-1123, 1989.
7. M. J. Leupold, A 42 cm bore superconducting coil using pre-reacted Nb<sub>3</sub>Sn, *IEEE Trans. Magn.*, **24**: 1413-1416, 1988.
8. M. O. Hoenig, Internally cooled cabled superconductors, *Cryogenics*, **22**: part 1, 373-389; part 2, 427-434, 1980.
9. S. Pourrahimi and K. Demoranville, Development of flexible Nb<sub>3</sub>Sn CiCC suitable for the react-then-wind approach, *IEEE Trans. Appl. Supercond.*, **7**: 816-819, 1997.
10. D. C. Larbalestier, The road to conductors of high temperature superconductors: 10 years does make a difference, *IEEE Trans. Appl. Supercond.*, **7**: 90-97, 1997.
11. A. R. Kantrowitz and Z. J. J. Stekly, A new principle for the construction of stabilized superconducting coils, *Appl. Phys. Lett.*, **6**: 65-67, 1965.
12. B. J. Maddock, G. B. James, and W. T. Norris, Superconducting composite: Heat transfer and steady state stabilisation, *Cryogenics*, **9**: 261, 1969.
13. B. J. Maddock and G. B. James, Protection and stabilisation of large superconducting coils, *Proc. Inst. Electr. Eng.*, **115**: 543, 1968.
14. O. Miura et al., Development of high-field ac superconducting magnet using ultrafine multifilamentary Nb-Ti superconducting wire with designed artificial pins, *Cryogenics*, **35**: 181-188, 1995.
15. M. N. Wilson, *Superconducting magnets, Monographs on Cryogenics*, Oxford: Clarendon Press, 1983, p. 159.
16. T. Kozman et al., Construction and testing of the mirror fusion test facility magnets, *IEEE Trans. Magn.*, **23**: 1448-1463, 1986.
17. S.-T. Wong et al., Design and construction of a large superconducting MHD magnet for the coal-fired flow facility at the University of Tennessee Space Institute, *Proc. Int. Cryogenic Eng. Conf.*, June 1980, Guildford: IPC Science Technology Press, pp. 785-789.
18. J. File, R. G. Mills, and G. V. Sheffield, Large superconducting magnet designs for fusion reactors, *IEEE Trans. Nucl. Sci.*, **4**: 277-282, 1971.
19. R. Jayakumar et al., Fabrication of ITER central solenoid model coil inner module, *IEEE Trans. Appl. Supercond.*, **7**: part 1, 981-984, 1997.
20. D. G. Hawsworth, Superconducting magnets systems for MRI, *Int. Symp. New Developments Appl. Supercond.*, Singapore: World Scientific, 1989, pp. 731-744.
21. A. Zhukovsky et al., 750 MHz NMR magnet development, *IEEE Trans. Magn.*, **28**: 644-647, 1992.
22. P. V. Livdahl et al., Status of the Fermilab energy doubler/saver project, *IEEE Trans. Nucl. Sci.*, **24**: 1218-1221, 1977.
23. J. Ahlback et al., Electromagnetic and mechanical design of a 56 mm aperture model dipole for the LHC, *IEEE Trans. Magn.*, **30**: 1746-1749, 1994.
24. S. S. Shen et al., First results of the full array LCT coil tests, *IEEE Trans. Magn.*, **23**: 1678-1682, 1987.

**Further Reading**

More information may be found in the following books and articles:

- A. C. Rose-Innes, *Introduction to Superconductivity*, 2nd. ed., Amsterdam, The Netherlands: Elsevier Science, 1994. This book describes the fundamental physics of superconductivity and its application to conductors for superconducting magnets.
- Proc. IEEE*, Special suppl. on superconductivity, **77**: 1110–1287, 1989. This IEEE review contains comprehensive contributions on magnets, conductors, power devices, and electronics. Despite the date of publication, the contents are still appropriate.
- H. Desportes, Three decades of superconducting magnet development, *Cryogenics*, ICEC suppl., **34**: 46–56, 1994. This is one of the more recent reviews of superconducting magnet development but contains less detail than the *Proc. IEEE* issue above. Reference 10 above is a concise review of the status of HTS conductors and their applicability to superconducting magnetic construction.
- P. Schmuser, in *The Physics of Particle Accelerators: Superconducting Magnets for Particle Accelerators*, Melvin Month and Margaret Diennes (eds.), AIP Conf. Proc., **249**: 1992, pp. 1099–1103.

JOHN E. C. WILLIAMS  
Massachusetts Institute of  
Technology

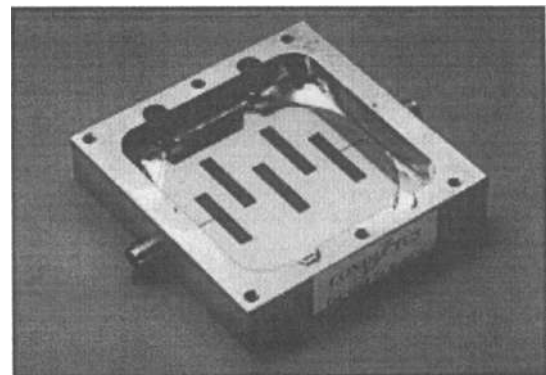
**SUPERCONDUCTIVITY: ELECTRONICS**

Since the late 1980s, rapid advances have been made in superconducting electronics. For example, it is now possible to fabricate low-temperature superconductor (LTS) digital circuits with thousands of active devices. These circuits operate at tens of gigahertz clock speeds and with several orders of magnitude less power consumption than conventional room-temperature electronics. Furthermore, high-temperature superconductor (HTS) devices are gradually entering the commercial market. One example is the demonstration HTS filter systems that are operational in a number of cellular phone base stations in the United States. This section gives the reader an overview of the broad field of superconducting electronics (1–3) with a description of applications and provides a historical perspective of technological development. A more in-depth description of a few specific aspects of superconducting electronics are given in the following sections.

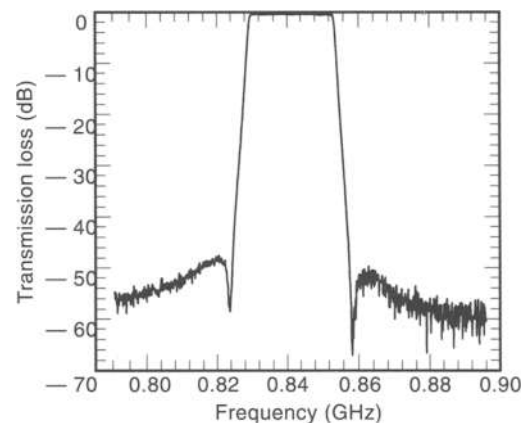
Superconductivity was discovered in 1911 in the laboratory of Kamerlingh Onnes, as a direct consequence of Onnes' invention of a method to liquefy helium. Helium becomes a liquid below 4.2 K (–268.8°C) so that with liquid helium one could study the properties of materials at very low temperatures. Researchers were measuring the resistance of mercury, lead, and tin and found that below a characteristic transition temperature  $T_c$  for each material, the resistance drops to zero. Subsequent experiments revealed that many metals become superconducting at very low temperatures. More recently, in 1986 Bednorz and Müller (4) discovered that certain ceramic materials are superconducting at relatively high temperatures. The most developed of these materials is  $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$  (YBCO) which undergoes a superconducting transition at approximately 90 K (–183°C).

Two fundamental characteristics of a superconductor are zero dc electrical resistance and the exclusion of internal magnetic fields, called the Meissner effect (5–7). A direct consequence of the second property, the Meissner effect, is that a magnet floats above the surface of a superconductor. The Railway Technical Research Institute in Tokyo, Japan, is using this property to develop a very fast magnetically-levitated train (8). Zero electrical resistance means that one can construct ideal lossless cables, wires, and transmission lines. Note that the resistance is really zero, and experiments have shown that for most cases a dc current flowing in a superconducting ring will persist for a million or more years. One application of superconducting cables is the coils of high-power electromagnets for magnetic resonant imaging (MRI).

As an example of a superconducting electronic application (1), see the analog filter shown in Fig. 1 (9). The resistance of the superconductor is zero at dc and very low at high frequencies. This property makes superconductors ideal for very sharp microwave frequency filters. Figure 1(a) shows a high-temperature superconductor (HTS) filter manufactured by Conductus Inc. (10), and Fig. 1(b) shows the frequency response of the circuit (9). The filter consists of a single layer of YBCO patterned into strips. These strips, coupled together by the electrical and magnetic fields from currents flowing in the superconductor, are designed to resonate at specific frequencies. These resonant frequencies define the frequency response

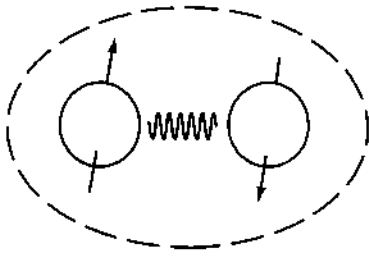


(a)



(b)

**Figure 1.** (a) A high-temperature superconducting filter and (b) the frequency response of a typical filter. Data courtesy of D. Zhang and co-workers at Conductus Inc..



**Figure 2.** Paired electrons in a superconductor. The arrows represent the spin of the electrons, and the oscillating line represents the interaction force. In each electron pair the spins of the electrons align opposite one another, so that the total Cooper pair spin is zero.

of the filter. This is due to the low insertion loss of the filter when incorporated into the front end of the receiver. Furthermore, the sharp edges of the filter characteristic allow frequency bands to be more closely spaced, which in turn allows more channels in a given bandwidth. Note that the HTS circuit in Fig. 1 requires a refrigerator to operate in the field, which is approximately the size of a large coffee can.

Other electronic applications include the superconducting mixer and the bolometer used in radio astronomy (1). In a typical bolometer, millimeter wave or infrared radiation is incident on a thin superconducting film, or microbridge. The high frequency radiation causes the superconductor to go normal, and the resistance of the film is proportional to the amplitude of the radiation. The bolometer is used in radio astronomy as a sensitive broadband detector.

The reader may be wondering how materials such as lead, niobium, or mercury become superconductors when cooled to very low temperatures? The mechanism for low-temperature superconductivity was not completely understood until 1957 when Bardeen, Cooper, and Schrieffer (BCS) (11) proposed a microscopic theory. The basic idea of their theory is that electrons pair up in a superconductor, see Fig. 2, to form so-called Cooper pairs. This electron pairing takes place throughout the superconductor so that all of the electrons are correlated with one another, as components of Cooper pairs. This correlation means that all of the electron pairs move together and do not experience collisions with other particles which would cause resistance. The BCS theory also showed that the mechanism, or force, keeping the electron pairs together is the vibrations of the crystal lattice. This explains why a definite transition temperature  $T_c$  exists for each material where the binding action of the lattice vibrations are strong enough, compared to the Coulomb electrical repulsion, to make the material superconducting. The BCS theory accurately describes low-temperature superconductors. However, now the mechanism for high-temperature superconductivity is not completely understood.

Prior to the BCS theory, London in 1935 proposed a phenomenological theory of superconductivity which derived from Maxwell's equations and Newton's laws (7). Later, in 1950, Ginzburg and Landau developed a more sophisticated phenomenological theory incorporating quantum mechanics (7). The BCS theory focuses on the microscopic electron-to-electron physics of a superconductor, whereas the London and Ginzburg-Landau theories give a very useful description of macroscopic phenomena that result from the electron pairing. Today both the London and the Ginzburg-Landau theories are often used to understand the dynamics of superconductors for

electronic applications (2). Furthermore, these phenomenological theories are also useful in that they are applicable to high-temperature superconductors.

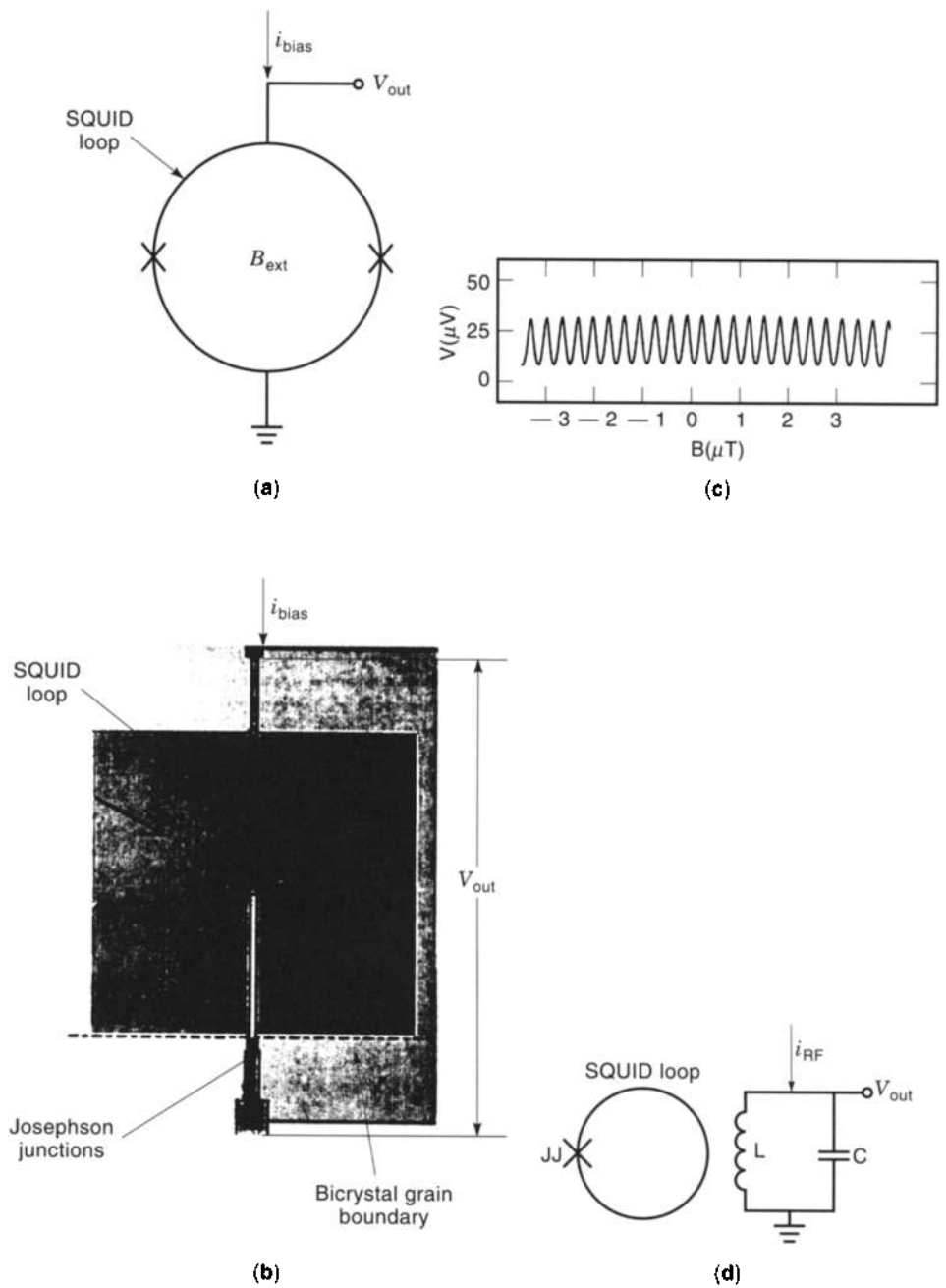
In 1962 B. D. Josephson predicted two dynamic properties of superconductivity (12). He analyzed two superconductors separated by a thin barrier through which quantum tunneling is possible. This device, called a Josephson junction today, can consist of a superconductor-insulator-superconductor sandwich, of which the insulator is approximately as thick as the Cooper pair diameter. Josephson's first prediction was that current can flow between the superconductors with zero applied dc voltage. Today this phenomenon is called the dc Josephson effect. Josephson's second prediction was that if a dc voltage  $V_{dc}$  is applied to the junction, then an alternating (ac) current will flow between the two superconductors with a very high frequency  $f = (2e/h)V_{dc}$ , where  $e$  is the magnitude of the charge of the electron and  $h$  is Planck's constant. This phenomenon is called the ac Josephson effect. Both the dc and the ac Josephson effect are a consequence of tunneling of Cooper pairs from one superconductor through the insulating barrier to the other superconductor. The Josephson junction has many applications in superconducting electronics. One of the first applications was to make a very sensitive magnetic field detector called a superconducting quantum interference device (SQUID) (13).

There are two types of SQUIDS: the direct-current (dc) SQUID and the resonant-frequency (RF) SQUID.

A dc SQUID consists of a parallel combination of two Josephson junctions in a superconducting (inductive) loop; see Fig. 3(a). The most sensitive SQUIDS are made with low-temperature metal superconductors. Figure 3(b) is a photograph of an YBCO HTS dc SQUID fabricated on a strontium titanate ( $\text{SrTiO}_3$ ) bicrystal substrate. When a dc current is applied, the output voltage measured across the two junctions is modulated by the magnetic field through the loop enclosed by the two junctions, see Fig. 3(c). This modulation is relatively large and is a function of the applied magnetic field. Using feedback electronics, one can measure magnetic fields extremely accurately and with a high dynamic range. The operating principle of the dc SQUID is described in more detail in the following section.

The RF SQUID is perhaps the simplest Josephson circuit and consists of a single Josephson junction in a superconducting loop, shown schematically in Fig. 3(d). The superconducting loop has an inductance and is magnetically coupled to a resonant circuit with an oscillating RF current input. Magnetic field coupled into the SQUID loop has the effect of detuning the resonant circuit, and this modulates the output voltage. The RF SQUID readout electronics is similar to the dc SQUID, and consists of feedback and lock-in amplification of the tuned circuit output voltage. The magnitude of the magnetic field input to the RF SQUID can be measured to an accuracy similar to the dc SQUID (14).

SQUIDS are the most sensitive magnetometers and can even be used to measure the magnetic fields generated by neurons firing in the human brain. Typical magnetic field sensitivities for LTS and HTS dc SQUIDS in the white-noise limit are  $2 \text{ fT}/\sqrt{\text{Hz}}$  and  $10 \text{ fT}/\sqrt{\text{Hz}}$  respectively ( $1 \text{ fT} = 10^{-15} \text{ T} = 10^{-11} \text{ G}$ ). Figure 4 shows one application of the dc SQUID as a scanning microscope (15). The image was generated by a HTS SQUID microscope developed by Clarke and co-workers at the University of California at Berkeley (UC Berkeley) (16). The image shown in Fig. 4 is a magnetic field map of the magnetic ink



**Figure 3.** (a) Circuit diagram of a SQUID where the X's represent Josephson junctions. (b) An HTS SQUID fabricated on a SrTiO<sub>3</sub> bicrystal. The SQUID is patterned from a thin film deposited on the bicrystal, and the Josephson junctions are formed on the grain boundary of the crystal. (c) Typical modulation data for a dc SQUID. The output voltage  $V_{out}$  in (b) is a periodic function of the magnetic field  $B$  coupled into the SQUID loop. Data provided courtesy of J. Clarke and E. Dantsker, University of California, Berkeley. (d) Schematic diagram of an RF SQUID.

for a small portion of a one-dollar bill, and was first observed by Welstood and coworkers at the University of Maryland (15). The image was obtained by scanning the SQUID at a distance of 150  $\mu\text{m}$  from the sample surface. The microscope resolution in Fig. 4 is approximately 130  $\mu\text{m}$ , but decreasing the size of the SQUID pickup loop can increase the resolution. For example, the IBM scanning SQUID microscope with a 4  $\mu\text{m}$  pickup loop has a resolution of approximately 4.5  $\mu\text{m}$  (15). The HTS SQUID in Fig. 4 operates at 77 K and is separated from the room temperature sample by a small window.

In addition to SQUID magnetic field sensors, the Josephson junction has other important applications, such as the international volt standard (17). The volt standard uses the ac Josephson effect. When tunnel junctions are irradiated by microwaves they produce constant voltage steps  $n(\hbar/2e)f$ , where  $n$  is an integer, and  $f$  is the frequency of applied microwave

radiation. A large number of junctions in series irradiated by a microwave source produces a large constant voltage step. The present standard for 1.2 V is maintained at the United States National Institute of Standards and Technology (NIST) by an array of approximately 2000 Josephson junctions irradiated by a 94 GHz microwave source. Figure 5(a) shows the commercial volt-standard system manufactured by HYPRES, Inc. (18). Note that the system incorporates a 5 K closed-cycle refrigerator. In recent years, Josephson fabrication technology has improved so that much larger arrays with a voltage of 10 V can be used as a standard. The 19-mm-long 20,208 junction array chip for the 10 V voltage standard is shown in the photograph in Fig. 5(b).

As another application example, researchers at TRW have recently demonstrated a dc to 10 GHz phase-shifter based upon a nonlinear transmission line (19). The circuit resembles the



**Figure 4.** Inset shows the magnetic image of a portion of George Washington's face on a \$1 bill observed with an HTS SQUID microscope. Data provided courtesy of T. S. Lee, G. Dantsker, and J. Clarke University of California, Berkeley.

voltage standard, and consists of many RF SQUIDS weakly coupled along the length of a superconducting transmission line. The RF SQUIDS add inductance along the transmission line, with the amount controlled by the current along the transmission line. The phase shifter has been successfully fabricated and tested using both low temperature (180,000 junction circuit) and high temperature (30,000 junction circuit) superconductor fabrication processes (19).

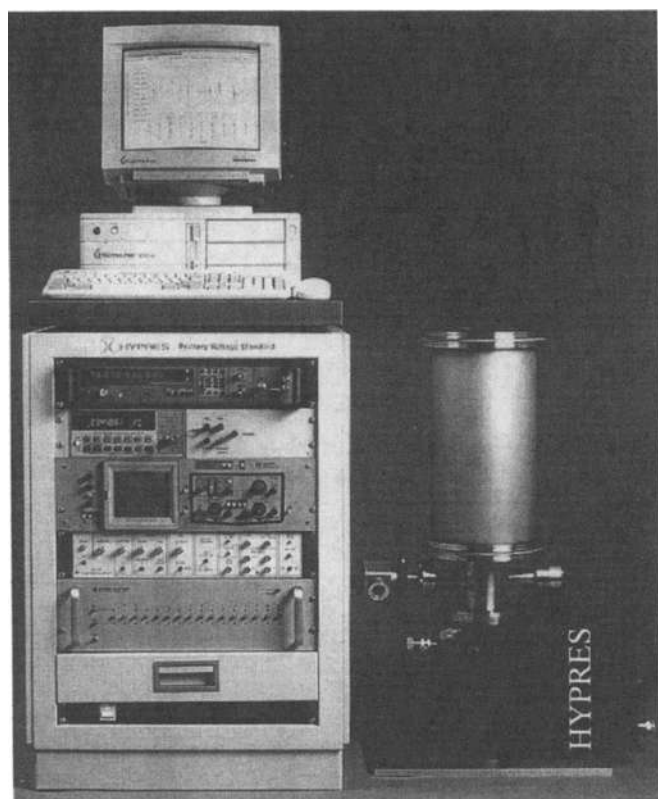
The near-instantaneous phase response of the phase shifter enables an additional new device, the superconducting upconverting parametric amplifier. Using a superconducting phase-shift transmission line, designers propagate a pure microwave tone (the "carrier") through the phase shifter. Simultaneously, low frequency signal current is also sent along the phase shifter. The low frequency signal varies the phase shift of the carrier, producing a phase-modulated carrier at the output. In addition to converting the low frequency waveform up to the higher frequency range, the phase-modulation mechanism produces power gain, with as much as 10 times (20 dB) gain demonstrated in some experiments (19).

The transmission line parametric amplifier is related to the single junction parametric amplifier, demonstrated in the mid-1980s by TRW and Bell Labs. Applying 36 GHz as a pump, an RF SQUID produced reflection gain for signals near 18 GHz. While producing as much as 32 times (30 dB) reflection gain, the amplifier successfully achieved noise levels at the quantum limit. Parametric amplifiers can therefore have extremely low noise, and degenerate parametric amplifiers in particular can exhibit the novel property of actually "squeezing" the internal noise to less than the amplifier quantum limit (20).

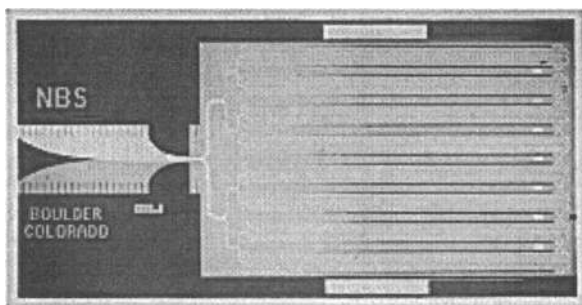
However, by far the most promising electronic application of the Josephson junction is digital circuits. The Josephson junction is the fundamental component of all modern digital superconducting electronics, analogous to a transistor in conventional digital circuits (although the operating principle is completely different). The Josephson junction has the property that it can switch on in less than a picosecond ( $1 \text{ ps} = 10^{-12} \text{ s}$ ), and the switching involves very little power dissipation. These properties make Josephson junction circuits ideal candidates for ultrahigh-speed computing applications.

In the last decade low-temperature superconductor microfabrication technology has progressed to the point that superconducting circuits with thousands of Josephson junctions are routinely manufactured. With this level of integration, one can make complex digital circuits, such as analog-to-digital converters (ADC) and small microprocessors that operate at multi-gigahertz clock speeds ( $1 \text{ GHz} = 10^9 \text{ cycles/s}$ ). As an example Fig. 6(a) shows a digital synthesizer designed by Spooner and co-workers at TRW (21). The circuit consists of approximately 700 logic gates ( $\sim 3000$  Josephson junctions), a read only memory (ROM) that stores a sine wave, and a digital-to-analog converter (DAC).

The circuit in Fig. 6(a) is designed to synthesize waveforms at high speed given a digital program input. Figure 6(b) shows synthesized output for the circuit operating with a 2 GHz internal clock. The power dissipation of the synthesizer is 6 mW, which is approximately 1000 times less power than an equivalent GaAs semiconductor circuit. The low on-chip power consumption of superconducting digital circuits is a fundamental advantage of the technology compared to semiconductor



(a)



(b)

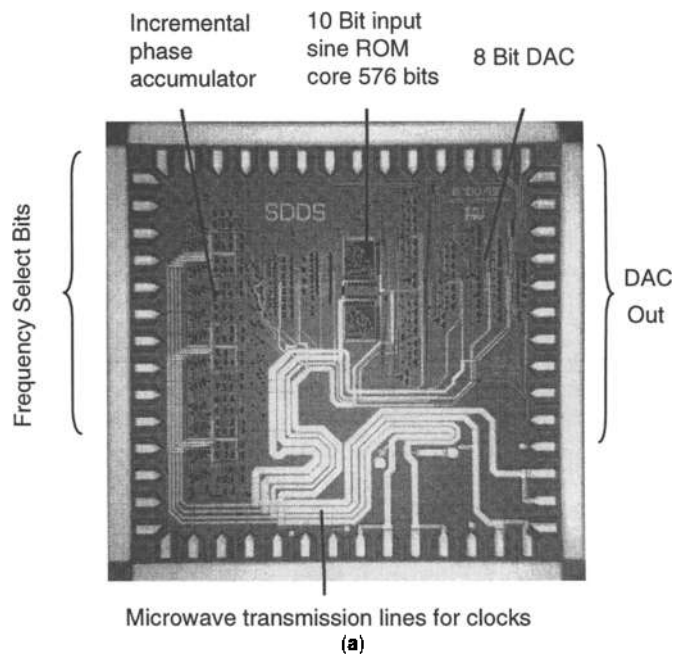
**Figure 5.** (a) The HYPRES commercial voltage standard system, including the 4 K Boreas cryocooler cold head, bottom right. The chip is mounted inside the evacuated cylinder on top of the cold head. The compressor is not shown. Photograph courtesy of HYPRES (18). (b) The 19 mm long 10 V voltage-standard chip consisting of over 20,000 Josephson junctions. Photograph courtesy of C. Hamilton and C. Burroughs at the National Institute of Standards and Technology, Boulder, CO (17). Original figure (b) © 1997 IEEE.

circuits. Therefore, superconducting digital circuits are ideal for low-power applications where large amounts of data must be processed at very high speed. One such application is digital on-board processing in a communications satellite.

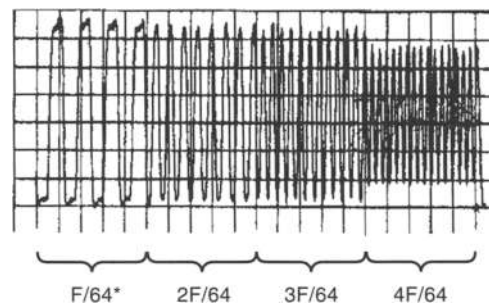
Digital superconducting electronic circuits are grouped in two main types: voltage-state logic (22) and rapid single-flux quantum (RSFQ) logic (23). The operating principles and differences between these two types of logic are described in detail in the following sections. Historically, voltage-state logic circuits were the digital circuits used in the IBM computer project of the 1970s and early 1980s (24). This project aimed

to build a mainframe computer using superconducting voltage-state logic. The project ended in 1983, and IBM chose not to pursue superconducting digital electronics further. The main reason was that IBM felt that, at that time, superconducting logic was not a significant enough improvement over projected future semiconductor circuits to warrant such a drastic shift in its core mainframe computer technology. Furthermore, the IBM Josephson circuits used a lead alloy technology that degraded over time (especially during thermal cycling), and, while the digital logic met the project specification goals, considerable difficulty was encountered in making high-speed memory suitable for a mainframe computer (24).

However, research in superconducting microprocessors continued in Japan using voltage-state logic. Several successful projects were completed under Ministry of International Trade and Industry (MITI) sponsorship. A series of processor chips were demonstrated at Fujitsu and Hitachi, and a four-chip model computer at the Electrotechnical Laboratory (22). In 1988 Kotani and co-workers at Fujitsu demonstrated a completely operational 4-bit microprocessor consisting of 1841 logic



(a)



(b)

**Figure 6.** (a) A superconducting digital synthesizer consisting of an incremental phase accumulator, 10 bit sine ROM, and an 8 bit DAC on a 1 cm × 1 cm chip. (b) Typical synthesized output for clock frequency  $f = 2$  GHz. Data and chip photograph courtesy of A. Spooner and co-workers at TRW (21). Original figure © 1997 IEEE.



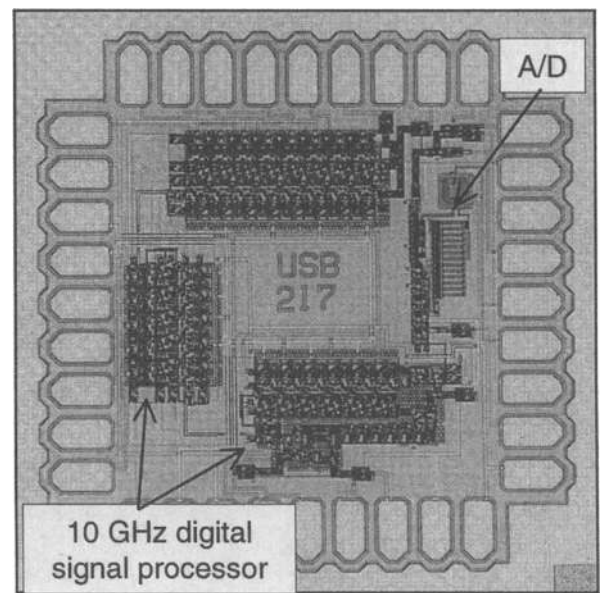
gates operating at 1 GHz with 6 mW of power dissipation (22). This processor used superconducting voltage-state modified variable threshold logic (MVTL) developed by Fujimaki and co-workers (25). MVTL circuits have much higher operating margins than the previous IBM logic circuits, and the MVTL microprocessor chip was fabricated using a more stable niobium trilayer technology. The circuit in Fig. 6 was designed using MVTL voltage-state logic. In 1990 the Fujitsu team designed and fabricated an 8-bit microprocessor consisting of 6300 MVTL gates (23,000 Josephson junctions) on a 5 mm × 5 mm chip (22). All of the component circuits operated at 1 GHz, and the multiplier had an average loaded gate delay of 5.3 ps. More recently the team led by Tahara at NEC in Japan has demonstrated a 4 kbit superconducting random access memory (RAM) with subnanosecond access times (26).

Voltage-state logic circuits have traditionally been used only for 1 GHz to 2 GHz applications. However Jeffery, Perold, and Vanduzer at UC Berkeley demonstrated a new type of voltage-state logic (27) in simple circuits at 10 GHz to 18 GHz. Therefore, much higher speed operation should be possible with this technology. One reason for the perceived limited speed of voltage-state logic is that the gates are latching. This means that circuits have to be reset to "zero" after a logical "one" operation, and before the next cycle of operation. Furthermore, the information (the logical ones and zeros) cannot be moved from one gate to the next without more than one clock. Typically 2 to 4 overlapping clocks power the gates, and these clocks have to be distributed throughout the circuit. The microwave transmission lines to power the circuit are clearly seen in Fig. 6. Without careful design, the large number of clock lines can cause excessive cross talk that can degrade circuit performance.

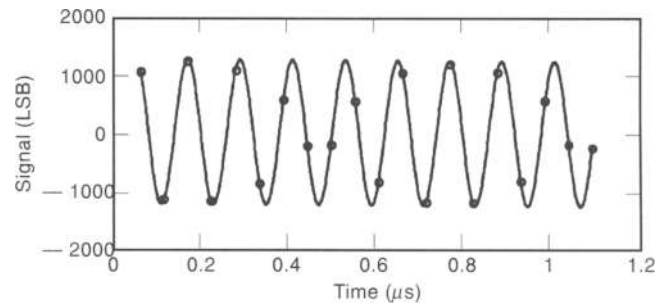
In 1985 Likharev, Mukhanov, and Semenov proposed a new type of superconducting logic based upon picosecond voltage pulses (23,28). They called this type of logic rapid single-flux quantum (RSFQ) logic and showed that, in principle, one could make RSFQ circuits that operate at clock frequencies approaching terahertz ( $10^{12}$  cycles/s). The operating principle of RSFQ, described in the following section, is fundamentally different from voltage-state logic. Because of the potential ultrahigh-speed operation (greater than 100 GHz), RSFQ is becoming the logic of choice for future digital superconducting applications.

Figure 7 shows a state-of-the-art RSFQ oversampling analog-to-digital converter (ADC) designed and tested by Semenov and co-workers (29) at the State University of New York at Stony Brook, and fabricated using the HYPRES Inc. 3.0 μm niobium process (18). The circuit consists of a single bit sampler and an integrated digital signal processor, called a decimation filter. The circuit is completely operational with a 9 GHz clock and a 10 MHz analog input bandwidth. The circuit comprises 1777 Josephson devices, consumes 0.5 mW of power, and has 11 effective bits. A similar ADC chip designed and fabricated at HYPRES operated in excess of 11 GHz with 11.5 effective bits. Note that these circuits have 10,000 times less on-chip power dissipation than an equivalent GaAs ADC.

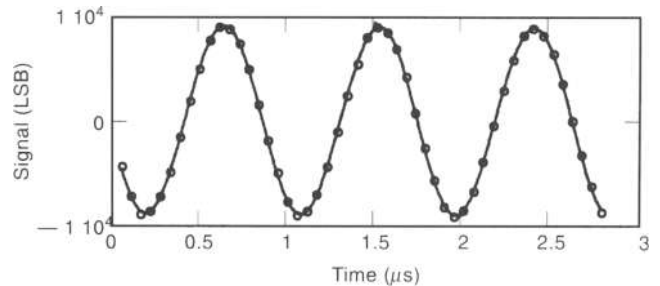
The circuit in Fig. 7 converts lower speed ( $\leq 10$  MHz) analog input signals into binary information that can be processed by a computer. Figure 7(b) and (c) show reconstructed sine wave data for the ADC for two different frequency sine wave inputs. The circuit is programmable so that, as the frequency of the input is reduced, higher resolution digital outputs can



(a)



(b)



(c)

**Figure 7.** RSFQ oversampling analog-to-digital converter consisting of a 1 bit sampler and an integrated digital signal processor on a 0.5 mm × 0.5 mm chip. The circuit consists of 1,777 Josephson junctions and is completely operational with a 9 GHz internal clock. (b) 11 bit reconstructed sine wave data for an 8.13 MHz input signal. (c) Reconstructed sine wave for a 1.13 MHz input signal. The circuit is programmable so that for the lower frequency input there are over 14 bits of resolution. Circuit photograph and data provided by V. Semenov, State University of New York at Stony Brook.

be obtained. This circuit performance is as good as the best semiconductor ADC presently available.

The key advantages of RSFQ circuits are their speed compared to all other digital circuits and that they use a dc power supply. However, the main circuit design challenge is that the gates require clocking. The circuit clock consists of picosecond very low energy voltage pulses with quantized area. The clock

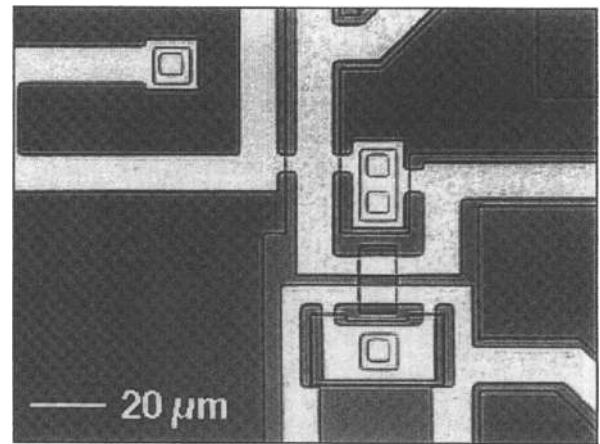
pulses are difficult to distribute in complex circuits operating at ultrahigh speeds. Furthermore, the picosecond RSFQ pulses cannot be directly interfaced with room temperature electronics because room-temperature electronics are not fast enough or sensitive enough to detect the low-energy picosecond pulses. Furthermore, RSFQ pulses are easily attenuated by nonsuperconducting cables and by impedance mismatches within superconducting circuits. Researchers are working hard to overcome these engineering design challenges to unlock the full technological potential.

RSFQ and voltage-state are not the only types of superconducting logic, although they are the most widely used today. For example, a type of superconducting logic circuit has been developed in Japan called the quantum flux parametron (QFP) (30), also called the parametric quantron (3). The QFP consists of a two junction loop with parameters chosen so that when an external clock is applied, an input to the loop will cause a quantized current to circulate in either the clockwise or the counterclockwise direction, depending upon the input. The direction of the current flow signifies the logical "1" or "0" in a QFP computer. A team directed by Goto at the RIKEN laboratory has developed innovative pipeline architectures for supercomputers based upon the QFP (30). Unfortunately, QFP circuits are sensitive to noise coupled from the clock lines, and are today not used for complex computing applications. However, the QFP is extremely sensitive to input signals, and is sometimes used as a comparator in analog-to-digital converter applications.

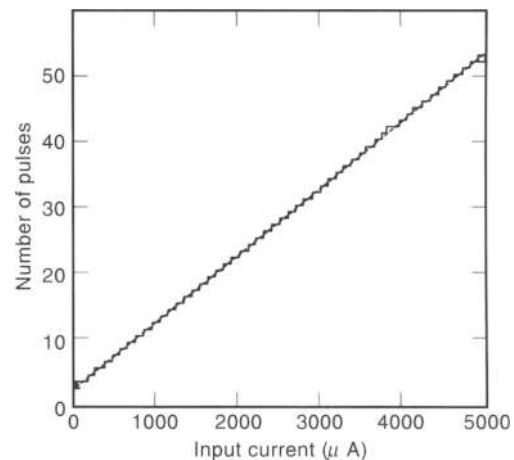
At present all practical superconducting digital circuits are fabricated from low-temperature superconducting materials. For example the circuits in Figs. 6 and 7 were fabricated by a process consisting of layers of patterned and etched thin film niobium and a resistive layer, separated by  $\text{SiO}_2$  insulator layers. Niobium is a superconducting metal at temperatures below 9 K. For digital circuits metallic layers are relatively straightforward to pattern and etch into integrated circuits.

Single-flux quantum logic circuits can also be implemented using high-temperature superconductor (HTS) materials. The ultimate advantage of HTS circuits is that they require much smaller refrigerators compared to low-temperature superconducting circuits. In the last few years there has been considerable interest in implementing ultra-high-speed RSFQ circuits using HTS materials. Figure 8 shows one such HTS circuit demonstrated at Northrop Grumman. The circuit uses fundamental properties of superconductivity to quantize a linear input signal and is the basic component of an analog-to-digital converter (ADC). The circuit was cooled to 65 K for the experiment.

High temperature superconductors, such as YBCO, are complex multilayered materials called perovskites, which are related to ceramics. When fabricating integrated circuits, if each layer of YBCO is not planarized, discontinuities of the crystal structure, or grain boundaries, occur at the wire crossovers. These discontinuities can form natural Josephson junctions and open circuits, which can cause circuit failure because they were not originally part of the circuit design. To date only small RSFQ digital circuits, consisting of fewer than 20 Josephson junctions and operating at low speed, have been demonstrated using HTS materials. There are various ways to make HTS Josephson junctions for digital circuits. However, none have thus far demonstrated sufficient process control to allow making junctions in quantity whose characteristics are adequately similar. (See HTS JOSEPHSON JUNCTION DEVELOPMENT.)



(a)



(b)

**Figure 8.** (a) High-temperature superconducting digital quantizer circuit for an analog-to-digital converter fabricated by a multilayer YBCO process. (c) The measured linearity of the quantizer is better than 5.7 effective bits at 65 K. Photograph and data provided courtesy of M. G. Forrester, D. L. Miller, and J. Przybysz at Northrop Grumman.

Recent Josephson junction fabrication research has also focused on using niobium nitride (NbN) for Josephson circuits (31). NbN is a low-temperature metallic superconductor, and it is much easier to fabricate digital circuits using this material than using HTS YBCO. The advantage of NbN is that its superconducting transition temperature is 16 K so that circuits can operate at 10 K: a 10 K cryocooler is significantly smaller than a 4 K cryocooler.

Since all superconducting electronic components operate at cryogenic temperatures, it is important to briefly discuss cryocooler technology (32). Table 1 lists representative cryocoolers, and temperatures that are important for superconducting electronics applications. The operating temperature of the cryocooler is dependent upon the superconducting electronic application, and in particular on the materials technology being used and the thermal load of the system. For example, present Nb LTS digital circuits require a 5 K cryocooler, whereas NbN digital circuits need a 10 K cooler. HTS circuits can operate at higher temperatures; for example, the wireless filter in Fig. 1 operates at 65 K, and some HTS SQUID systems can operate at

**Table 1. Representative Cryocoolers and Operating Temperatures for some Superconductive Electronic Applications**

| Temperature | SC Application                                | Operating Principle  | Cooling Power | Manufacturer                      |
|-------------|---|--|---------------|-----------------------------------|
| 5 K         | Nb LTS Digital                                | 2-stage GM with rare earth regenerator<br>or with JT stage | 0.5–1.5 W     | Sumitomo, Chesapeake/Boreas       |
| 10 K        | NbN Digital                                   | 2-stage GM <sup>a</sup>                                    | 1–7 W         | Sumitomo, CTI, Leybold            |
| 35 K        | HTS SQUIDS and HTS Digital                    | Single-stage Stirling                                      | 0.5 W         | CTI, Leybold                      |
|             |   | Single-stage GM  | 4 W           | CTI, Leybold                      |
|             |   | Single-stage pulse-tube                                    | 1–10 W        | TRW, Lockheed Martin <sup>b</sup> |
| 65 K        | HTS lters for cellular phone<br>base stations | Single-stage Stirling                                      | 2.5 W         | CTI, Leybold                      |
|             |   | Single-stage GM  | 6–90 W        | CTI, Leybold                      |
|             |   | Single-stage pulse-tube                                    | 5–60 W        | TRW, Lockheed Martin <sup>b</sup> |
|             |   | Single-stage pulse-tube                                    | 4.5 W         | Iwatani                           |
| 75 K        | Some HTS SQUID systems                        | Mixed gas cooler   | 1 W           | APD CryoTiger                     |

GM = Gifford–McMahon closed cycle, JT = Joule–Thomson.

<sup>a</sup>10 K 2-stage pulse-tube and 2-stage Stirling cycle cryocoolers are presently in development.

<sup>b</sup>Not commercially available, although presently used for satellite applications.

75 K. Note that the power dissipation of superconducting components, except for HTS wireless filter systems, is extremely small. In typical applications the leads connecting from room temperature to the superconducting component are the dominant heat load. If this is minimized, cooling powers of 100 mW to 1 W are adequate for many superconducting electronics applications, and hence the cryocoolers can operate near to their base temperature.

Most cryocoolers use the principle of compressing and then expanding a gas to produce cooling. For a given thermal load, the cryocooler size decreases dramatically as the temperature is increased. Applications at temperatures less than 20 K typically require a Gifford–McMahon or Boreas type cryocooler, shown for the voltage-standard system in Fig. 5(a). The superconducting circuits are attached to the end of a cold head which consists of a capped metal cylinder separating the circuit from the cold helium gas. In Fig. 5(a) the voltage standard circuit is sealed inside an evacuated container around the cold head. Gifford–McMahon/Boreas cryocoolers use an external compressor, and additional pistons behind the cold head. The piston expands the gas from the external compressor, and oscillates at approximately 1 Hz. The cryocooler cold head plus the vacuum jacket is roughly the size of a personal computer tower (~36 kg), and the compressor is the size of a small filing cabinet (~90 kg), not shown in the figure. Two stages of cold head and piston configurations are required to get to 5 K temperatures, with a rare earth regenerator used to extract additional heat from the gas, or with an additional Joule–Thomson cooling stage.

Stirling and pulse-tube cryocoolers generally operate at higher temperatures (although a pulse tube has been demonstrated below 2.5 K), and are often much smaller than the Gifford–McMahon/Boreas. They can be approximately the size of a small coffee can (5 kg to 10 kg total weight) for superconducting electronic applications. The pulse-tube cryocooler, in particular, incorporates a small diaphragm and piston that is oscillated at approximately 60 Hz to compress and expand the gas coolant; this rapid oscillation gives a relatively large cooling power for a small mechanical volume. The focus of commercial R&D for Stirling and pulse-tube refrigerators is now to increase the operating lifetime, for example through the use of compressors with gas bearings to avoid any mechanical wear.

Cryocooler reliability has long been an issue for commercial superconducting electronics applications. However, in recent years vast improvements have been made. For example the 5 K Sumitomo cryocooler requires servicing after 10,000 h of

continuous operation, and servicing consists of swapping the regenerator filter. At higher temperatures, the pulse-tube type cryocoolers have impressive reliability records; two TRW pulse-tube cryocoolers are presently in orbit in space satellites, and the reliability of these coolers is estimated at 10 years with zero maintenance. However, it should be noted that such coolers for space applications, including the Stirling machines pioneered at Oxford, are built at enormous cost, many hundreds of thousands of dollars. The challenge is to achieve the same reliability in commercial coolers sold for less than \$10,000, that is, manufactured for a few thousand dollars.

In the following sections superconducting electronics is described in more detail. Examples of applications are limited to SQUIDS and digital electronics. Basic design and circuit testing principles are described. Finally, possible future directions for the field of digital superconducting electronics are discussed. For a more in-depth introduction to superconducting electronics, see Refs. 1, 2, and 3.

## BASIC PRINCIPLES OF SUPERCONDUCTING ELECTRONICS

### Josephson Junctions and the Josephson Effect

At temperatures below the superconducting transition, the electron pairs in a superconductor are correlated with each other. Quantum mechanically this means that all of the Cooper pairs are in the same macroscopic quantum state, so that we can describe the superconductor mathematically by a macroscopic wave function  $\Psi$ :

$$\Psi = \sqrt{\rho} \exp(i\phi) \quad (1)$$

where  $\rho$  is the density of the superconducting paired electrons and  $\phi$  is the phase of the wave function. Similar to the single-particle Schrödinger equation, the phase of the wave function Eq. (1) is related to the external magnetic field by  $\phi = 2e/\hbar \int \mathbf{A} \cdot d\mathbf{l}$  where  $\mathbf{A}$  is the magnetic vector potential ( $\mathbf{B} = \nabla \times \mathbf{A}$ ).

A Josephson junction is shown schematically in Fig. 9(a). The device is rather simple and consists of two superconductors separated by a thin barrier (such as an insulator, normal metal, or semiconductor) through which quantum tunneling of electron pairs can take place. Mathematically, this is equivalent to coupling the wave functions of the two superconductors through the barrier. Following Feynman (5), one can assume linear coupling across the barrier and that the wavefunctions on both sides obey the single-particle Schrödinger equation.

Using Eq. (1) and these coupled equations, one can derive the following famous Josephson equations:

$$I_s = i_c \sin \phi \tag{2}$$

and

$$V = \frac{2e}{\hbar} \frac{d\phi}{dt} \tag{3}$$

where  $I_s$  is the supercurrent through the junction,  $V$  is the voltage across the junction,  $\phi$  is the phase difference between the wave functions on both sides of the barrier, and  $i_c$  is the critical current of the junction.

Equation (2) describes the supercurrent flow through the junction, and Eq. (3) gives the voltage across the junction. For real applications the Josephson junction is modeled by the equivalent circuit in Fig. 9(b). The superconducting Josephson junction is denoted by the cross in the circuit and is in parallel with a resistance and a capacitor. The resistor represents the flow of nonsuperconducting electrons through the junction, and the capacitance exists because the device is a “sandwich” of two superconducting layers separated by a thin dielectric insulator.

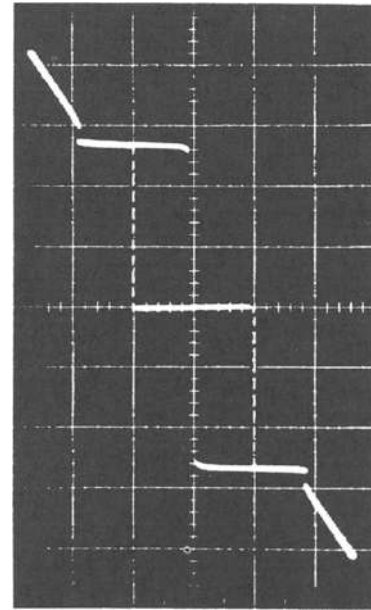
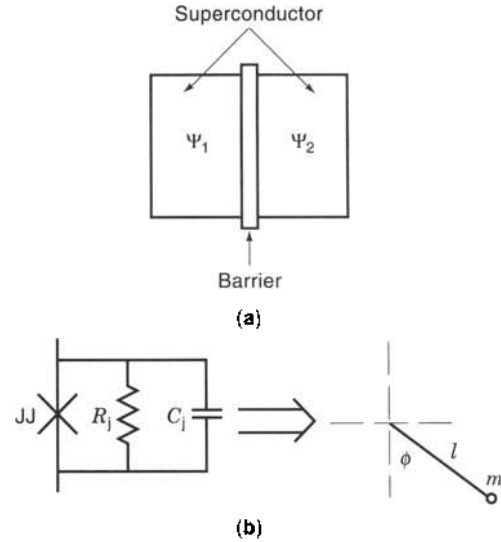
From Eqs. (2) and (3) and Kirchhoff’s laws (that is, set the bias current equal to the sum of the currents through each component) it is straightforward to show that the general equation governing the Josephson junction in Fig. 9(b) is

$$C_j \frac{d^2\phi}{dt^2} + \frac{1}{R_j} \frac{d\phi}{dt} + \frac{2\pi i_c}{\Phi_0} \sin \phi = \frac{2\pi}{\Phi_0} i_{\text{bias}} \tag{4}$$

where  $R_j$  is the junction resistance,  $C_j$  is the junction capacitance,  $i_c$  is the critical current of the junction,  $i_{\text{bias}}$  is the current applied to the junction, and  $\Phi_0$  is the constant  $h/2e$ . The nonlinear term with  $\sin \phi$ , if linearized, has the same form as an inductor in place of the junction. Therefore the nonlinear term is sometimes called the kinetic inductance of the Josephson junction.

Equation (4) is identical to the equation for a pendulum where  $\phi$  is the angle of the pendulum, shown schematically in Fig. 9(b). Therefore one can visualize the pendulum to obtain an intuitive feeling for the dynamic behavior of the Josephson junction. Comparing Eq. (4) to the pendulum mechanical analog, the junction capacitor  $C_j$  corresponds to the pendulum moment of inertia (mass  $m$  times length  $l$  squared),  $1/R_j$  to the damping coefficient,  $(2\pi/\Phi_0)i_c$  corresponds to  $mg/l$  (where  $g$  is the acceleration due to gravity), and  $(2\pi/\Phi_0)i_{\text{bias}}$  to an externally applied torque.

Intuitively, an increasing dc current applied to the junction is equivalent to increasing the torque on the pendulum Fig. 9(b). When the current is less than  $i_c$ , the pendulum does not spin, and the voltage across the device is zero. However, when the current exceeds a threshold corresponding to the critical current  $i_c$ , the pendulum begins to spin because of the torque of the applied current. As the pendulum spins, the angle (junction phase) changes with time, so that by Eq. (3) a voltage appears across the device. The pendulum is nonlinear, and once it starts spinning it has an “angular momentum.” Hence, as the bias current (externally applied torque) is removed, the pendulum continues to spin, so that the junction voltage is nonzero. This corresponds to the fact that the  $I$ - $V$  curve of the lightly damped Josephson junction is hysteretic. Adding a resistance of a few



**Figure 9.** Schematic diagram of a Josephson junction consisting of two superconductors separated by a thin barrier. (b) The equivalent circuit model for a Josephson junction with a parallel normal resistance and a capacitance, sometimes called the *RSJ* model. The circuit model is equivalent to a pendulum mechanical analog where the angle of the pendulum is equal to the phase difference of the wave functions across the barrier. (c) Experimental  $I$ - $V$  curve for a typical  $\text{NbAlO}_x/\text{Nb}$  Josephson junction photographed from an oscilloscope. Note the 2.5 mV gap voltage for the junction and that the  $I$ - $V$  curve is hysteretic. The dashed line denotes switching that is too fast to see on the oscilloscope. Data provided courtesy of X. Meng, University of California, Berkeley.

ohms in parallel with the junction removes the hysteretic behavior and corresponds to adding damping to the pendulum analog.

The pendulum analog described so far assumes the junction resistance is constant, which is useful for analytic calculations, and describes the general dynamics of the junction. However, the junction resistance is actually a nonlinear function so that when the junction begins to spin (switches on) the resistance sharply increases. The nonlinear resistance of the Josephson

junction is included in Josephson SPICE circuit simulators (35). For the specific case of an underdamped superconductor-insulator-superconductor (SIS) tunnel junction, the voltage that appears across the device when it switches on is called the gap voltage of the superconductor. For Josephson junctions in general, the gap voltage physically corresponds to the energy required to break apart Cooper pairs into individual electrons on one side of the Josephson junction, and recombine them as Cooper pairs on the other side. The gap voltage is dependent on the superconducting material and is 2.4 mV to 2.8 mV for Nb, and 4 mV to 5 mV for NbN. The hysteretic  $I$ - $V$  curve for an underdamped SIS niobium Josephson junction with a nonlinear resistance is shown in Fig. 9(c).

Superconducting circuits consist of Josephson junctions and inductive loops for SQUIDS, where the SQUIDS and junctions are wired together using superconducting transmission lines. Resistors are used for damping, and to feed currents into the SQUIDS and junctions. In order to make a superconducting circuit one therefore needs multiple layers of superconductor for wires, a layer of resistor, and a superconducting groundplane.

Low-temperature superconductive niobium Josephson circuits are fabricated by using a trilayer process. First, a layer of niobium a few tenths of a micrometer thick is deposited on a silicon wafer. The surface of this layer is coated with a layer of aluminum about 4 nm to 10 nm thick. Then the aluminum is oxidized to a thickness of about 1 nm, and a second layer of niobium (Nb) is deposited on top. The thin aluminum oxide layer is as the tunnel barrier for the NbAlO<sub>x</sub>Nb Josephson junction. Note that the barrier thickness is approximate, and is only within a factor of two or three of 1 nm. Using photolithography and reactive ion etching, one can pattern circuits from the trilayer with many thousands of Josephson junctions each a few micrometers in size. Additional superconducting wiring and resistor layers are deposited and etched, separated by SiO<sub>2</sub>, to make complex circuits. The wiring and resistor layers are connected to each other and the Josephson junctions by vias, or vertical contacts, that are patterned and etched into the SiO<sub>2</sub> insulating layers. For high-speed circuits, a ground plane is usually deposited either as the first or the last step in the process. Note that the actual process is rather simple, that is, all deposition is at 300 K and there is no doping, ion implantation, or high frequency diffusion.

By using a trilayer all of the Josephson junctions are effectively fabricated simultaneously across the wafer, and no processing occurs between depositions of the layers. This means that there are only small variations among the characteristics of each Josephson junction. To make complex circuits the critical current variations across a chip should have a standard deviation less than 3%. At present no trilayer process exists for HTS circuits, and hence the parameter spreads on  $i_c$  are rather large (approximately 20 to 100% standard deviation). These large process spreads are the main reason that only small HTS Josephson circuits have been demonstrated thus far. (See HTS JOSEPHSON JUNCTION DEVELOPMENT.)

Finally, note that the switching-on time of the Josephson junction is very fast and can be less than a picosecond. This makes the device useful for ultra-high-speed computing applications. Intuitively, decreasing the junction capacitance corresponds to decreasing the mass of the pendulum, which in turn makes the junction switch (rotate) faster. Mathematically the ultra-fast switching speed is understood by noting that the linearized Eq. (4) (i.e., assuming  $\sin \phi \approx \phi$ ) has a natural angular

frequency

$$\omega_c = \sqrt{\frac{2\pi i_c}{\Phi_0 C_j}} \tag{5}$$

where  $\omega_c$  is called the plasma frequency of the Josephson junction. Similar to the pendulum angular frequency, the plasma frequency relates to how fast the Josephson junction phase (angle) can change. For example, a 300  $\mu$ A junction with a 0.002 pF capacitance has a Josephson frequency  $f_c = \omega_c/2\pi = 1.1$  THz. Since the junction capacitance depends upon the area  $a$  of the junction, the switching speed of Josephson circuits can be increased by increasing the critical current density  $J_c = i_c/a$ , and decreasing the area of the junction to keep  $i_c$  constant. With today's fabrication process, typical  $J_c$  values are 1 kA/cm<sup>2</sup> to 2.5 kA/cm<sup>2</sup>, where the smallest junction linear dimension is 3  $\mu$ m. The junction switching time with a 3  $\mu$ m process is a few picoseconds. However, processes with  $J_c$  as high as 50 kA/cm<sup>2</sup> have been demonstrated for small circuits. With this high critical current density, and a submicron Junction fabrication process, subpicosecond switching times have been demonstrated (see the section on RSFQ logic).

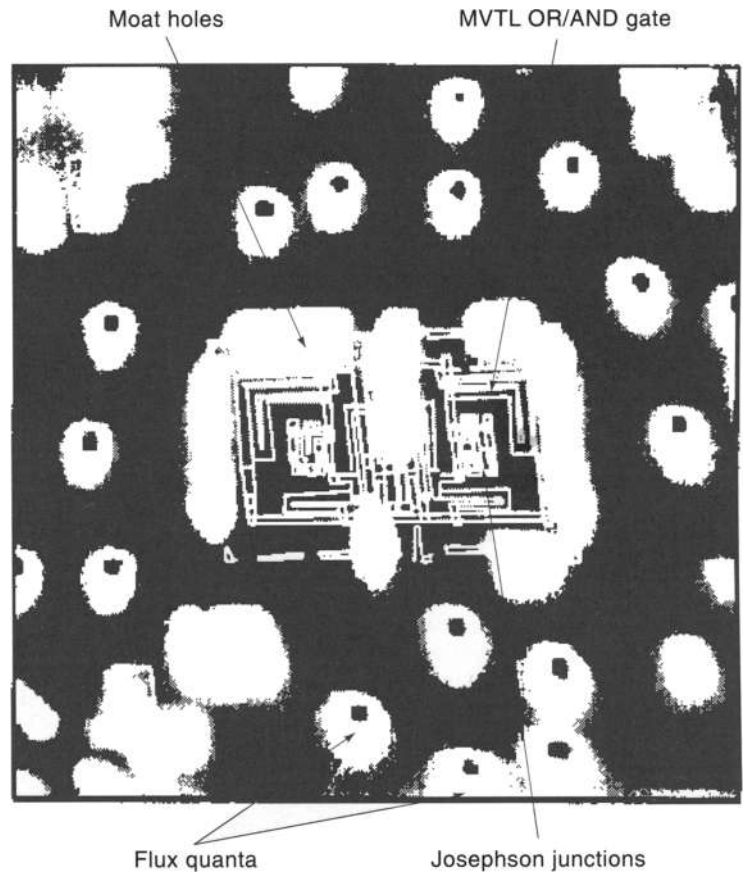
### Flux Quantization

The response of superconducting materials to external magnetic fields divides materials into Type I and Type II superconductors. Type I superconductors remain in the superconducting state and exclude all external magnetic fields until they reach a critical field. At the critical field, the magnetic field enters the material, and it returns completely to the normal nonsuperconducting state. However, this is strictly true only for long thin samples parallel to the magnetic field. Type II superconductors, on the other hand, allow magnetic fields to enter while the material remains in the superconducting state.

The magnetic field that enters in a Type II superconductor has the interesting property that it is quantized, so that the magnetic flux ( $\Phi = B \times A$ ) equals  $h/2e$ . The quantized magnetic flux corresponds to a circulating current loop in the superconductor. Mathematically, the flux quantization phenomenon is a consequence of single valuedness applied to the macroscopic wave function Eq. (1), that is, the phase change  $\Delta\phi$  around any closed superconducting loop must be an integer  $n$  times  $2\pi$ , which is equal to the magnetic action  $(2e/\hbar)A$  integrated around the loop:

$$\Delta\phi = \frac{2e}{\hbar} \oint \mathbf{A} \cdot d\mathbf{l} = \frac{2e}{\hbar} \Phi = 2\pi n \Rightarrow \Phi = n\Phi_0 \tag{6}$$

where  $\Phi$  is magnetic flux and the flux quantum  $\Phi_0 = h/2e = 20.7$  G  $\cdot$   $\mu$ m<sup>2</sup>. The third term is derived by Stokes's theorem because the magnetic field is related to the vector potential  $\mathbf{A}$  by  $\mathbf{B} = \nabla \times \mathbf{A}$ . On a microscopic level, the actual structure of a flux quantum in a superconducting film is complicated. However, to a good approximation it corresponds to a small supercurrent loop in the plane of the film, where the film is normal on the inside of the loop. The magnetic field generated by the current loop integrated over the area surrounding the flux quantum corresponds exactly to Eq. (6). The superconductors used for circuit fabrication, YBCO and niobium when deposited as a thin film, are Type II superconductors.



**Figure 10.** A  $400\ \mu\text{m} \times 400\ \mu\text{m}$  scanning SQUID microscope image of an MVTL OR-AND logic gate. The magnetic field scale is from 0 (white) to 400 mG (black), and the dots are the flux quanta trapped in the thin film of the superconducting ground plane. Data provided courtesy of M. Jeffery and T. Vanduzee, University of California, Berkeley, and J. Kirtley and M. B. Ketchen at IBM.

Because the earth's magnetic field is approximately 650 mG ( $10^4\ \text{G} = 1\ \text{T}$ ), many flux quanta will be trapped in a superconducting thin film if it is cooled in the earth's magnetic field. From Faraday's law  $V = d\Phi/dt$ , so that thermally induced motion of the flux quanta generate low-frequency voltage noise in superconducting circuits. This noise can limit performance of SQUID magnetometers, and researchers have spent considerable effort devising methods to reduce this noise (see the following section on SQUIDs). Furthermore, if flux quanta trap close to a Josephson junction they can change the junction critical current, and this prevents digital circuits from operating correctly.

Figure 10 is an image of the magnetic field above a superconducting MVTL logic gate. The sample was cooled inside a mumetal magnetic shield in a field of 4.5 mG. Mumetal is a very high permeability material commonly used for magnetic shields. A line drawing of the logic circuit has been superimposed on the image to show its location. The image was obtained by scanning a small ( $4\ \mu\text{m}$  diameter) SQUID loop across the area of the circuit and measuring the magnetic field. The image measures  $400\ \mu\text{m} \times 400\ \mu\text{m}$  and the dots clearly show the flux quanta trapped in the film. The Josephson junctions in the circuit can be protected from trapped flux by cutting small "moat" holes in the superconducting ground plane surrounding the circuits. Then the magnetic field traps in the moat holes and not in the circuits (33). The use of moats to protect superconducting digital circuits enables designing complex digital circuits that operate in magnetic fields of several milligauss. Magnetic field values less than 5 mG are easily obtained with mumetal magnetic shields.

### The dc Superconducting Quantum Interference Device (SQUID) Magnetometer

The basic SQUID circuit is shown schematically in Fig. 3(a). The device consists of two Josephson junctions connected in parallel forming an inductive superconducting loop. For low-temperature superconductor (LTS) SQUIDs normal resistors  $R$  are typically connected in parallel with the Josephson junctions to give a nonhysteretic  $I$ - $V$  characteristic for the combination. When a dc bias current is applied, the voltage across the resistor is modulated by the magnetic flux ( $\Phi = B_{\text{ext}} \times A$ ) coupled into the SQUID loop. This modulation phenomenon is understood as follows (7).

The total current flowing through the SQUID loop is the sum of currents from the two Josephson junctions given by Eq. (2). Each junction has a phase, and the phase change around the SQUID loop is related to the magnetic field. Specifically, if  $\phi_1$  and  $\phi_2$  denote the phase of each junction, then the total phase change around the SQUID loop is given by

$$\phi_1 - \phi_2 = \frac{2e}{\hbar} \oint \mathbf{A} \cdot d\mathbf{l} = 2\pi \frac{\Phi}{\Phi_0} \quad (7)$$

where the last term is derived similarly to Eq. (6).

If a dc current  $I > 2i_c$  is applied to the SQUID loop, then the excess current must be carried by the parallel normal resistance, so that  $I = I_s + V/2R$  or

$$V = R \left[ \frac{I}{2} - i_c(\sin \phi_1 + \sin \phi_2) \right] \quad (8)$$

Using Eq. (7) and a trigonometric identity,

$$V = R \left[ \frac{I}{2} - i_c \cos \left( \pi \frac{\Phi}{\Phi_0} \right) \sin \gamma \right] \quad (9)$$

where  $\gamma = (\phi_1 + \phi_2)/2$ . The phase  $\gamma$  evolves rapidly in time by the Josephson Eq. (3), and because  $\sin \gamma$  is oscillating, it may appear that the voltage contribution from the Josephson junctions averages to zero. However, the junctions obey the nonlinear Eq. (4), and the oscillation is not a true sinusoid. Therefore, the average of  $\sin \gamma$  is nonzero.

Using Eqs. (3) and (4), one can integrate  $\sin \gamma$  to find its average value (34). The corresponding average value of  $V$  is given by

$$\bar{V} = R \left[ \left( \frac{I}{2} \right)^2 - i_c^2 \cos^2 \left( \pi \frac{\Phi}{\Phi_0} \right) \right]^{1/2} \quad (10)$$

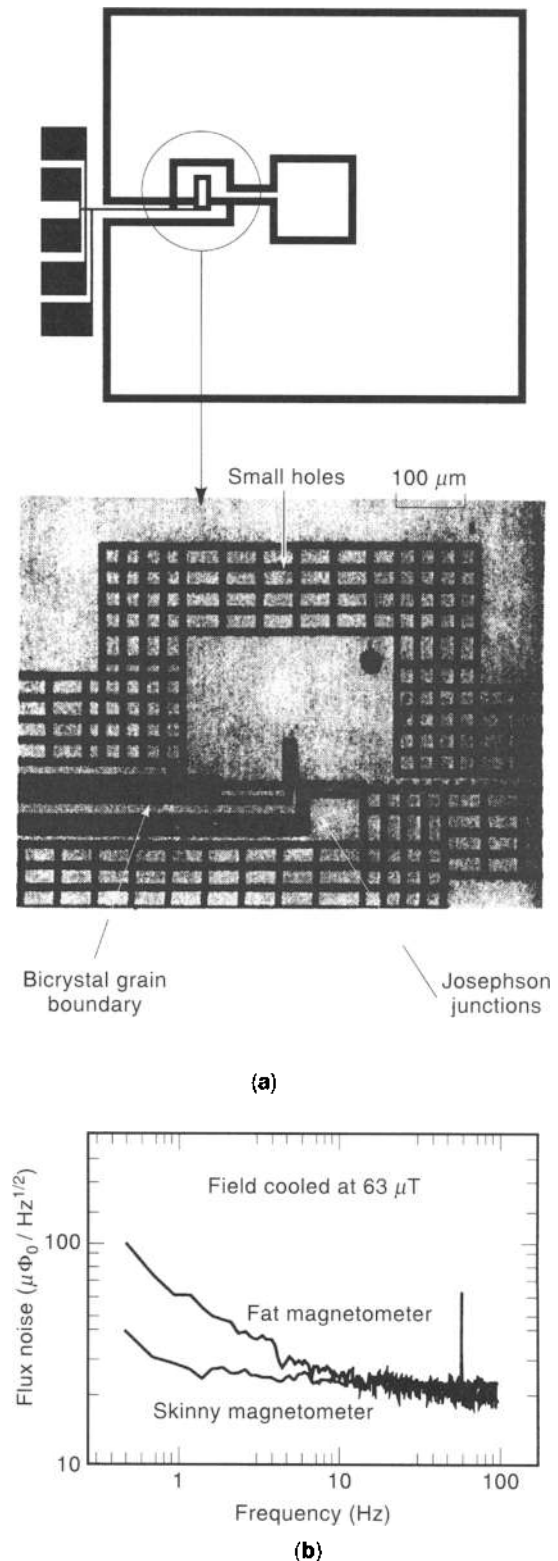
Therefore the average voltage across the dc SQUID, which can be measured with a dc voltmeter, oscillates as a function of  $\Phi$  with a period  $2\Phi_0$ . These oscillations are shown for a typical HTS SQUID in Fig. 3(c). One can think of the SQUID as a flux-to-voltage transducer. Using a kilohertz modulator and feedback electronics, called a flux-locked loop, one can use the SQUID characteristic Fig. 3(c) to measure magnetic fields very accurately and with a high dynamic range (13).

For an HTS SQUID, shown in Fig. 3(b), the Josephson junctions are naturally resistively damped by a low normal resistance (of order  $1 \Omega$ ), so that no external resistors are necessary. In this case the resistance  $R$  in Eq. (10) becomes  $R_j$ , where  $R_j$  is the normal resistance of the HTS junction.

At present, practical YBCO HTS SQUID magnetometers are fabricated using a strontium titanate ( $\text{SrTiO}_3$ ) bicrystal. The crystal lattice structure and size of  $\text{SrTiO}_3$  are similar to those of YBCO. Therefore, when a film of YBCO is deposited on  $\text{SrTiO}_3$  the crystal growth matches the  $\text{SrTiO}_3$  substrate lattice.  $\text{SrTiO}_3$  bicrystals consist of two crystals fabricated and polished with a mismatched grain boundary. When YBCO is deposited on this bicrystal, the YBCO film aligns with the  $\text{SrTiO}_3$  crystals on each side of the boundary, and the discontinuity forms a long uniform Josephson junction along the boundary. Then the thin film is patterned to form a SQUID structure, such as Fig. 3(b).

The dashed line in Fig. 3(b) shows the location of the bicrystal grain boundary. The two Josephson junctions of the SQUID are too small to be seen on the scale of the photograph. However, the SQUID loop is clearly seen. External magnetic fields are coupled into the SQUID loop by a flux transformer (not shown in the figure). The transformer consists of a thin film of superconductor patterned into a square spiral of several large loops connected to a large pickup loop. It can be fabricated on a separate chip and "flipped" on top of the SQUID. However, today most flux transformers are fabricated on the same chip and direct-coupled to the SQUID loop, similar to Fig. 11.

As described in the previous section, noise from flux trapping is a major problem that limits the performance of HTS SQUID magnetometers operating in the earth's magnetic field. Clarke and co-workers at UC Berkeley have spent considerable effort working on this problem. Figure 11 shows a photograph of a new type of magnetometer consisting of a lattice of small ( $4 \mu\text{m}$  wide) superconducting lines patterned in YBCO on a  $\text{SrTiO}_3$  bicrystal. The idea of using a lattice for the body of the



**Figure 11.** (a) A new type of SQUID consisting of a lattice of thin,  $4 \mu\text{m}$  YBCO wires patterned in a lattice on a  $\text{SrTiO}_3$  bicrystal substrate. The top picture shows the overall geometry of the device, and the lower photograph shows a portion of the microlattice with the two SQUID Josephson junctions. (b) Flux noise in the SQUID for a conventional "fat" magnetometer and the new "skinny" microlattice type. The flux noise in the skinny magnetometer is significantly less at low frequencies than the conventional magnetometer. Data and photograph provided by R. McDermott, H. M. Cho, B. Oh, K. A. Kouznetsov, A. Kittel, and J. Clarke, University of California, Berkeley.

magnetometer is similar to using moats to protect superconducting logic circuits (see Fig. 10).

Flux in a superconductor is quantized in units  $B \times A = \Phi_0$ . For the microlattice design of Fig. 11, the line width is chosen at  $4 \mu\text{m}$  so that  $\Phi_0/16 \mu\text{m}^2 \sim 1 \text{G}$ , which is approximately twice the value of the earth's magnetic field. Therefore, at the transition temperature, flux quanta are unlikely to trap directly in the thin superconducting wires. Instead, they trap in the honeycomb of holes comprising the SQUID body, where they are effectively pinned and cannot easily move around.

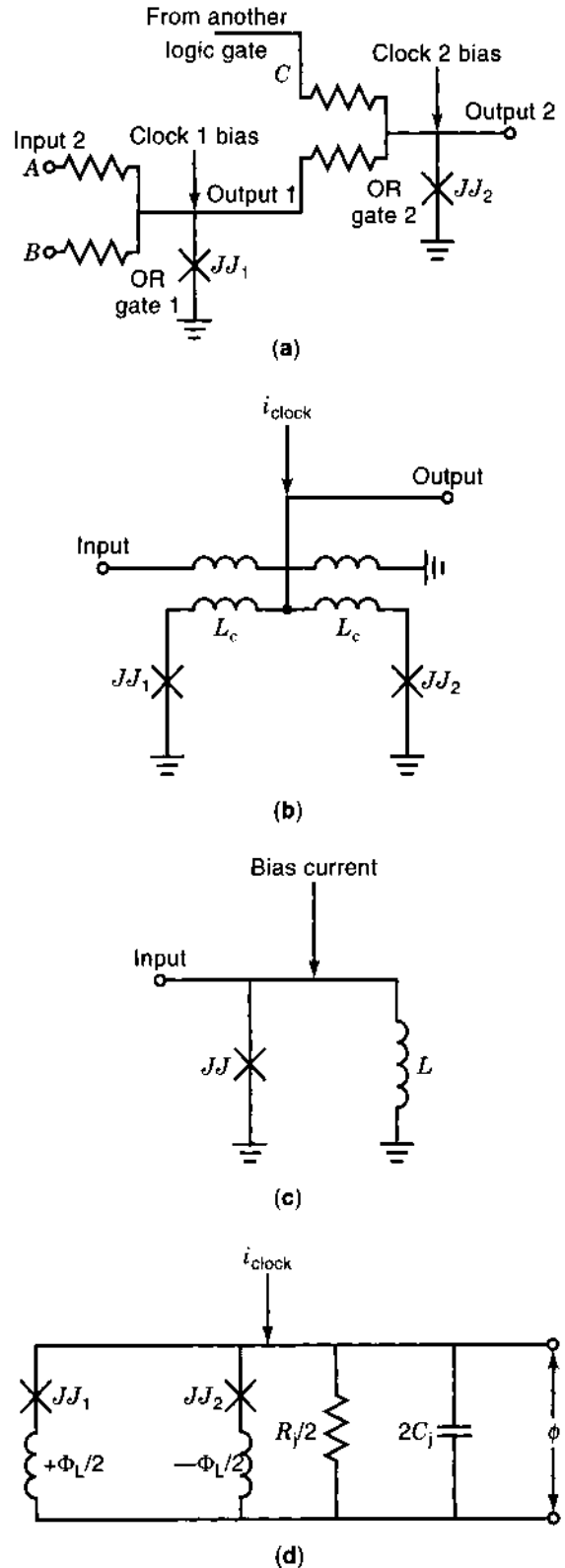
Because flux motion is a significant cause of voltage noise in the SQUID, the net result is that the noise floor of the SQUID is considerably reduced. Figure 11(c) shows the flux noise versus frequency of the thin lattice type SQUID and a conventional SQUID magnetometer cooled in 630 mG ( $63 \mu\text{T}$ ) of applied magnetic field (approximately the earth's magnetic field). Notice that the low frequency  $1/f$  noise of the thin magnetometer is much lower than the conventional (fat) dc SQUID.

**DIGITAL SUPERCONDUCTING ELECTRONICS**

**Voltage-State Logic**

The basic logical "switch" for voltage-state logic is the hysteretic Josephson junction; see Fig. 9. We can model the hysteretic Josephson junction by the mechanical analog of a pendulum that is free to rotate. The junction phase is analogous to the pendulum angle, and a current bias applied to the junction is equivalent to an external torque applied to the pendulum. When the external torque on the pendulum is increased beyond a critical value, the pendulum begins to spin rapidly. This situation is analogous to the dynamics of the Josephson junction. For a Josephson junction, when the bias current is increased to a value greater than the critical current, the junction phase begins to change rapidly. From Eq. 3 the voltage across the junction is proportional to the rate of change of the phase, so that a voltage appears across the device. The junction is said to have switched to the voltage state. This output voltage is the logical "one" in a computer, and no voltage corresponds to a logical "zero." In order to reset the junction after a logical "one," the bias must be turned off, and for this reason logic using hysteretic junctions is sometimes termed "latching" logic. The switching-on time of the junction can be less than a picosecond. However, because the junction is hysteretic (the pendulum has an angular momentum), it takes a much longer time to switch off when the bias is removed. Therefore, the practical limitation on how fast voltage-state circuits can be clocked in real applications is set by the switch-off time.

As an example of the basic voltage-state logic principle, one could use the switching properties of a single Josephson junction as a logic gate. For an OR gate, the critical current of the junction is chosen so that when the clocked bias and either of two inputs A or B is applied, the critical current of the junction is exceeded, and the device switches on; see Gate 1 in Fig. 12(a). For Gate 1 in the figure, note that before the junction  $JJ_1$  switches to the voltage state the Clock 1 bias current has a superconducting path to ground through the junction, so there is no output current from the gate. However, after the junction  $JJ_1$  has switched it becomes a relatively high impedance (since there is now a voltage across the device), and some of the Clock 1 bias current is shunted to the output of the gate. This current



**Figure 12.** (a) Schematic diagram of a simple voltage-state logic circuit constructed from hysteretic Josephson junctions. (b) Schematic diagram of the two-junction SQUID and (c) the one-junction SQUID used in digital superconducting electronics. (d) Equivalent circuit to (b) assuming the junctions are identical and the SQUID loop inductances  $L_c$  are small so their corresponding voltages can be neglected. The Josephson junctions' normal resistances and capacitances are not shown in (a)–(c) and are shown schematically in (d) for the equivalent circuit.



will switch Gate 2 when Clock 2 is applied. A logical AND can be constructed similarly by choosing the critical current so that it takes a clocked bias plus both inputs *A* and *B* to switch the junction. This is exactly the logical AND of two input signals. The inverse logic function is more complicated but can also be constructed.

The logic gate is on only when the clock is applied. Therefore, for data to propagate between two gates, Gate 1 and Gate 2 in Fig. 12(a), the first gate must be clocked on before the second. Furthermore, Gate 1 must remain on when the clock for Gate 2 is applied. This means that different clocks must be used to move data from one gate to the next. Overlapping clock phases are required for data to propagate through a chain of arbitrary logic gates. Voltage-state logic is often powered by a three-phase clock, whose phases are 120° apart. However, three phases are not always used for voltage-state logic. For example, the original IBM project used a single-phase clock with dc latches after each logical operation, and the circuit Fig. 6 uses a four-phase clock.

The single Josephson junction logic gates described previously are not practical for real applications, since they are very sensitive to parameter variations. Furthermore, there is no isolation between junctions, so that output current from a switched gate can feed back through the gate input and erroneously trigger previous gates. This is clearly seen in Fig. 12(a). Given that Clock 1 is high and there is an input at *C*, when Clock 2 is applied the Gate 2 junction  $JJ_2$  switches to the voltage state and becomes a relatively high impedance. The current from Clock 2 is then shunted to both the output and the input of the gate. This feedback of current through the input may cause the junction  $JJ_1$  in Gate 1 to also switch to the voltage state. Several different voltage-state logic families have been demonstrated that are robust to these problems. Most notably Fujimaki and coworkers have developed modified variable threshold logic (MVTL) (25). MVTL circuits have been demonstrated with thousands of gates at multigigahertz clock rates (see Fig. 5).

The basic building blocks of voltage-state logic circuits are the two- and one-junction SQUIDs shown schematically in Fig. 12(b) and (c). These circuits are similar in structure to the dc SQUID and RF SQUID magnetometers described previously; however, the parameters and modes of operation can be quite different. In order to analyze the underdamped two-junction SQUID in Fig. 12(b) it is useful to make the approximation that the coupling inductors  $L_c$  in the SQUID loop are small, so that only the flux  $\Phi_L$  coupled into the SQUID loop ( $\pm\Phi_L/2$  in each SQUID inductor  $L_c$ ) is important and the voltage across the coupling inductors can be neglected.

Assuming that the Josephson junctions in the two-junction SQUID have identical critical currents, capacitances, and linear resistors, the equation governing the circuit is straightforward to derive from the equivalent circuit Fig. 12(d). This is an ideal two-junction SQUID. Summing the currents in the branches and setting them equal to the input clock current, and noting that the phase difference across  $JJ_1$  is  $\phi - (\pi/\Phi_0)\Phi_L$  and  $JJ_2$  is  $\phi + (\pi/\Phi_0)\Phi_L$ , where  $\phi$  is the total phase across the SQUID, and  $V = (\Phi_0/2\pi)d\phi/dt$  is the total voltage, we obtain

$$2C_J \frac{d^2\phi}{dt^2} + \frac{2}{R_J} \frac{d\phi}{dt} + \frac{4\pi i_c}{\Phi_0} \cos\left(\frac{\pi\Phi_L}{\Phi_0}\right) \sin\phi = \frac{2\pi}{\Phi_0} i_{\text{clock}} \quad (11)$$

where a trigonometric identity has been used to combine the

two sine functions. This equation is similar to the single junction pendulum (4) where the critical current of the junction is  $2i_c \cos(\pi\Phi_L/\Phi_0)$ , and is modulated by the coupled magnetic flux  $\Phi_L$ . Hence, a current in the control line varies the switching point of the device.

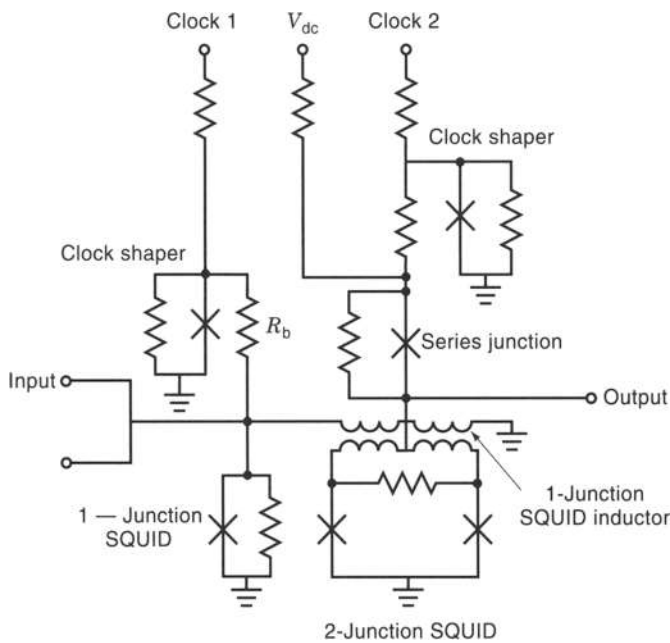
We can think of the two-junction SQUID circuit as two coupled pendulums with a difference in angle  $\Delta\phi = \pi\Phi_L/\Phi_0$ . When the applied bias current, or torque on the pendulums, exceeds the critical value  $2i_c \cos(\pi\Phi_L/\Phi_0)$ , the two pendulums spin and a voltage  $V = (\Phi_0/2\pi)d\phi/dt$  appears across the device; the two-junction SQUID has switched to the voltage state. Similar to the single junction, the two-junction SQUID switching is hysteretic. The switching-on time is very fast, but the coupled pendulums have an angular momentum so that they continue to spin when the clock bias is removed. The advantage of this circuit is that the output voltage is isolated from the input control line.

The operation of the one-junction SQUID is understood as follows. For the one-junction SQUID in Fig. 12(b), the input currents (which can be a clock bias and an external bias) must equal the current through the Josephson junction, junction resistance, junction capacitance, and inductor. Because the flux  $\Phi = Li$ , the current  $i_L$  through the inductor adds an additional linear term  $i_L = (\Phi_0/2\pi)\dot{\phi}/L$  to Eq. (4). This linear term adds an effective quadratic term to the potential of the nonlinear pendulum, so that two different modes of operation are possible. Analysis of the one-junction SQUID equation shows that when the dimensionless parameter  $\beta_L = 2\pi Li_c/\Phi_0 \leq 1$  the current in the inductor is a single-valued oscillating function for a linear input current. However, when  $\beta_L > 1$  the current in the inductor is no longer a single valued function, and as a linear input current is applied abrupt switching is observed. For  $\beta_L > 1$  the circuit operation is hysteretic, so that after switching, if the applied current is reduced the current in the inductor will continue to flow until a different switching point is reached. This hysteretic behavior is shown in Fig. 13(b) for the one-junction SQUID used in the complementary output switching logic (COSL) gate.

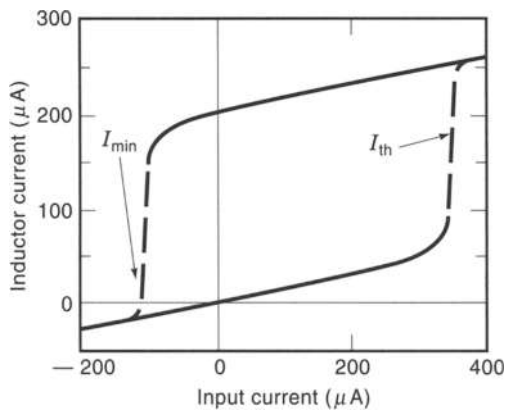
As an example of voltage-state logic circuit design, the basic ideas of the COSL family are briefly reviewed (27). COSL was developed for applications from 5 to 20 GHz and was optimized by using a Monte Carlo method in HSPICE (35), so that the basic gates and logic circuits have a high probability of operating at ultrahigh speed despite the process variations of critical current, resistance, and inductance.

Figure 13(a) is a schematic diagram of the COSL OR/AND gate. The XOR function is derived from the OR gate by including a 300  $\mu\text{A}$  Josephson junction in series with the inputs. All of the COSL family of gates consist of a one-junction SQUID input stage and a two-junction SQUID output stage. The two-junction SQUID in the output stage is connected in series with a Josephson junction. The COSL circuits are designed to use a three-phase sinusoidal clocking scheme, and the input and output stages of the gates use two of the clock phases applied through the clock-shaping junctions. These junctions have the effect of clamping the SQUID biases at approximately 2.5 mV when the clocks are applied, independent of the process variations.

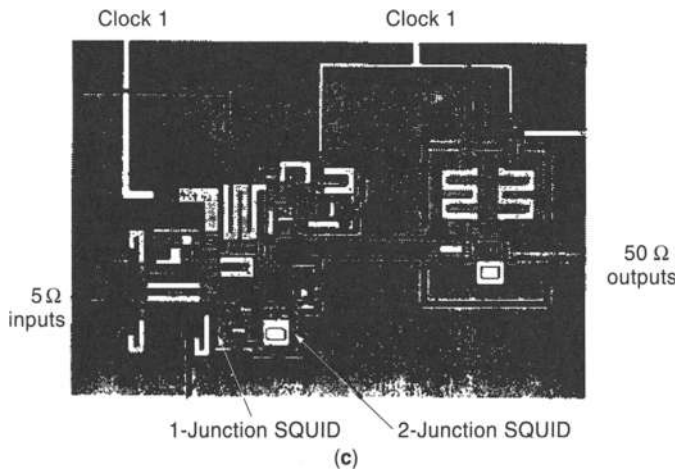
The input circuit for the COSL gate is a one-junction SQUID, similar to Fig. 12(c). For this circuit  $\beta_L > 1$ , so that the SQUID is hysteretic, and when it switches, a relatively large current flows through the inductor. Figure 13(b) shows the relationship



(a)

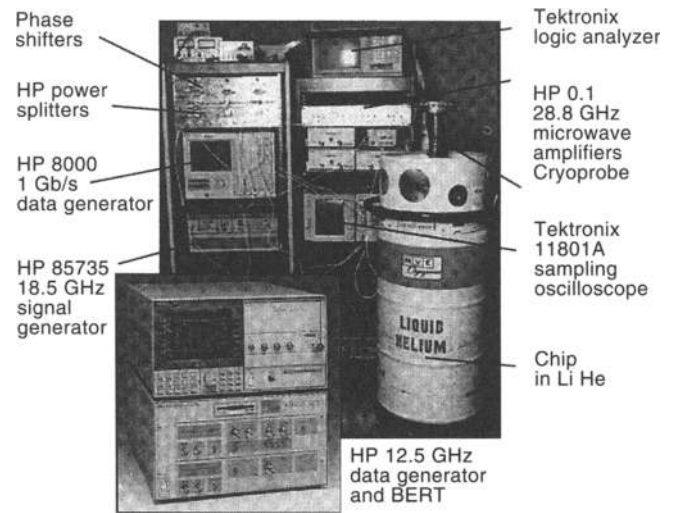


(b)



(c)

**Figure 13.** (a) The COSL voltage-state logic gate consists of a one-junction SQUID input stage, and a two-junction SQUID output. (b) The current transfer curve for the COSL one-junction SQUID. (c) Micrograph of the COSL OR/AND gate fabricated using the HYPRES 2.5 μm process. The input and outputs of the circuits are carefully impedance-matched.



**Figure 14.** A typical laboratory high-speed test setup. The superconducting chip is mounted on the end of a high-bandwidth probe and is immersed in a liquid helium dewar. Inputs and outputs are via high-speed SMA cables, and the output is detected on the sampling oscilloscope.

between the input current and the inductor current for the one-junction SQUID. For the COSL gate parameters, with an input current of 350 μA, the current in the inductor switches to 250 μA. The output circuit of the COSL gate consists of a hysteretic two-junction SQUID, and is similar to Fig. 12(b).

The operation of the COSL OR gate in Fig. 13 is understood intuitively as follows. When clock 1 is applied, an input to the gate greater than 60 μA is sufficient to fire the one-junction SQUID. Switching the one-junction SQUID causes a relatively large current to flow in the inductor, which is coupled to the output two-junction SQUID loop. By Eq. (11), the one-junction SQUID current suppresses the critical current of the two-junction SQUID so that when clock 2 is applied, the two-junction SQUID switches, giving 1 mV at the output, which produces 200 μA in a 5 Ω load. An AND gate is constructed by increasing the resistor  $R_b$  in Fig. 13(a); this reduces the current from the clock so that two inputs are required to switch the gate.

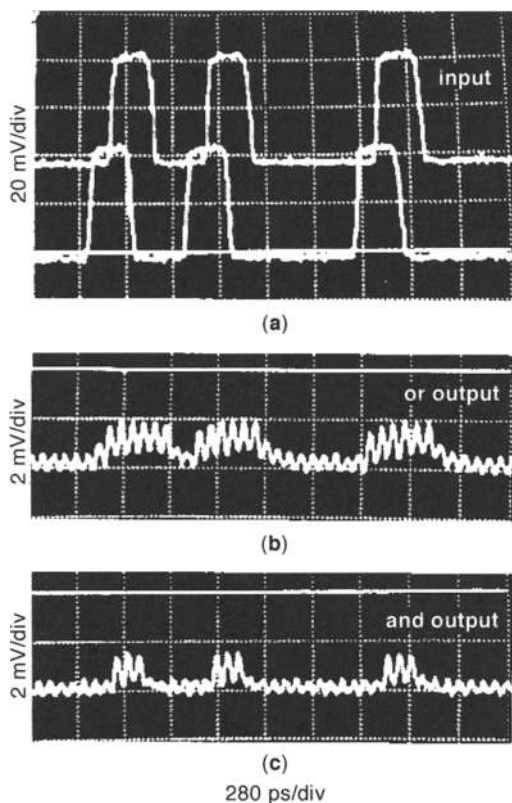
Figure 13(c) is a micrograph of a COSL gate fabricated by the HYPRES Inc. (18) 2.5 kA/cm<sup>2</sup> niobium process. Special care is taken to impedance match all inputs, outputs, and clock lines of the circuit. Specifically, the inputs and outputs to the gate are 5 Ω superconducting transmission lines, and the off-chip driver is a large single Josephson junction designed to switch into a 50 Ω load. The inputs to the circuit from the room temperature electronics are impedance-matched from 50 Ω to 5 Ω using a simple resistive matching network (not shown in the figure). Impedance matching is essential for circuit operation above a few gigahertz because reflections within the circuit can cause erroneous switching and can make signal detection difficult.

A typical laboratory test setup is shown in Fig. 14. The chip is mounted on the end of a high-bandwidth probe and immersed in a liquid helium dewar. The end of the probe is surrounded by two mumetal magnetic shields. Inputs to the circuit are generated by room temperature multigigabit per second data generators, and the circuit outputs are observed on the sampling oscilloscope. The sinusoidal clock is generated by a signal generator and then split into three phases and amplified

by microwave amplifiers. Phase shifters and programmable attenuators vary the phase and amplitudes of the clocks independently. The clock, input signals, and sampling scope must be phase-locked to observe the output of the circuit at high speed.

Figure 15 shows typical test data for the COSL gate in Fig. 13(c) clocked at 15 GHz. The data inputs (top trace) are overlapping low-speed 4 GHz signals. The output, however, is much faster than the input data because the gate is clocked at 15 GHz. For OR operation the COSL gate switches "on" when either of the input signals are high, and switching is observed for AND operation when the two input signals overlap. The oscillation on the background in Fig. 15(b) and (c) is due to feedthrough of the unbalanced clocks, and is sometimes called "ground bounce." The basic COSL gates have been demonstrated at 18 GHz, the maximum speed of the test equipment, and complex COSL encoder circuits for flash ADCs have been demonstrated at 5 to 8 GHz (27).

It is traditionally thought that voltage-state logic operates only in the range of a few gigahertz. This limited speed operation is usually attributed to the "punch-through" problem. There are two types of punch-through, one caused by the plasma oscillations of the junction (36), and another called low-probability punch-through (37). The first type of punch-through is related to the plasma oscillation of the Josephson junction when it resets. Using the mechanical analog, when the undamped pendulum stops spinning, there are finite oscillations at the bottom of the arc, as the pendulum damps to zero; see Eq. (5). This is analogous to the latching Josephson junction.



**Figure 15.** (a) Photograph of 4 GHz input test data taken from a sampling scope. A pulse splitter and a short length of cable were used to make the phase delay of the top trace. (b) Output of the COSL OR gate circuit clocked at 15 GHz. The circuit switching is observed when either of the two lower speed signals is switched "on." (c) The AND gate switching occurs when the two inputs overlap.

As the junction resets, voltage oscillations called plasma oscillation are observed at the end of the logical pulse. If the clock is applied before complete damping has taken place, then the junction can misfire. This misfire, caused by nonresetting of the junction, is termed punch-through. However, circuit parameters can be chosen so that for a given clock speed and junction process the damping oscillations never misfire the circuit. Therefore, although the long resetting time of the junction limits the ultimate clocking speed, it will not generate random punch-through errors if the circuits are designed correctly.

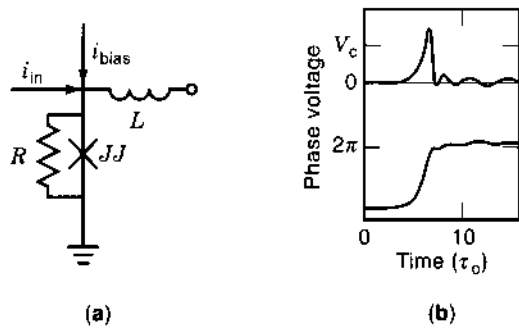
To understand the concept of low-probability punch-through, return to the pendulum model of the Josephson junction, Eq. (4). When the undamped spinning pendulum is reset by removing the external torque (clocked dc bias), there is a small but finite probability that it will stop at the top of the arc rather than at the bottom, that is, there is a small probability that the pendulum will end up balanced at the unstable equilibrium point. If this happens for a Josephson junction, then the junction is easily switched to the voltage state without any input when the next clock cycle is applied. The device "punches through" to give an error "one" output even though there may have been a "zero" input. In a process with high Josephson junction initial current, the probability of this type of punch-through is very small, much less than  $10^{-12}$  for COSL. Therefore for applications, such as analog-to-digital conversion, low-probability punch-through is not a significant source of errors. Bit error rate measurements on the COSL gates demonstrate that voltage-state logic can operate with very high clock speeds without significant errors from punch-through.

The design challenge for complex voltage-state logic circuits is the distribution of the multiphase clocks. For large circuits the clock lines have low impedance and will carry all of the power for the circuit (of the order of mW). Considerable cross talk can occur between transmission lines, so that the design and testing of complex voltage-state logic circuits that operate at speeds beyond 10 GHz is challenging. The ultimate application of voltage-state COSL circuits may be as interfaces between rapid single-flux quantum logic (RSFQ) circuits and room temperature electronics. RSFQ circuits can operate at speeds in excess of 100 GHz and are described in detail in the following section.

### Rapid Single-Flux-Quantum (RSFQ) Logic

RSFQ logic was proposed by Likharev, Mukhanov, and Semenov (23,28) in 1985, and was based upon the principles proposed by Silver and his collaborators (38) and Sawada and co-workers (39). The basic RSFQ circuit element is shown schematically in Fig. 16(a). The Josephson junctions are resistively damped by parallel resistors and are connected to an inductor. A dc bias is applied to the junction. The value of the critical current, resistor, and the inductor are chosen so that the dimensionless parameter  $\beta_c = 2\pi i_c R^2 C_j / \Phi_0 \sim 1$  and  $Li_c \sim \Phi_0$ .

For the simple case where the inductor in Fig. 16(a) is connected to ground, Eq. (4) is applicable to the circuit where the inductance adds just a linear term in  $\phi$ . Returning to the pendulum mechanical analog for the Josephson junction, the junction (pendulum) cannot "spin" freely because of the damping resistor. Recall that a dc bias applied to the junction is equivalent to a torque on the pendulum. For a bias current approximately 80% of the critical current value, the pendulum is rotated close to the horizontal unstable equilibrium point. If the pendulum is kicked over the unstable equilibrium point, it flips in a circle.



**Figure 16.** (a) The basic RSFQ circuit element. The junction is shunted by a resistor that is typically  $1 \Omega$  to  $2 \Omega$ . (b) A single flux quantum pulse generated by flipping a damped Josephson junction in parallel with an inductor. The width of the pulse is a few picoseconds. The original part (b) is from Likharev and Semenov (23), © 1991 IEEE.

One can imagine that with gravity the pendulum moves slowly over the top, rapidly picks up speed toward the bottom of the arc, and then moves back to the original position. From Eq. (3) the rate of change of the pendulum angle (junction phase) is equivalent to the voltage across the device. Hence, as the junction flips, a voltage pulse is produced; see Fig. 16(b).

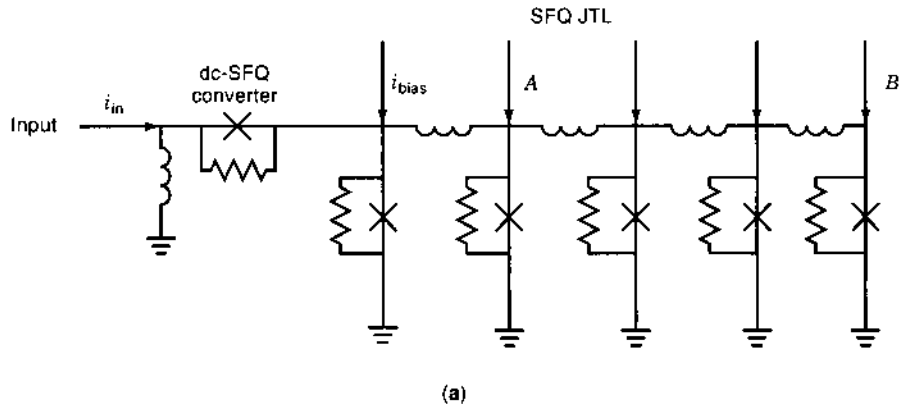
The junction switching is very fast, and the corresponding voltage pulse is typically a few picoseconds in width. These picosecond voltage pulses comprise the logical “ones” in an RSFQ digital circuit. Logical “zero” is the absence of a picosecond pulse within a clock period. As a junction flips, the phase (pendulum angle) changes by exactly  $2\pi$  radians. From Eq. (6) the junction

phase can be related to the magnetic field in the inductor by  $\phi = 2\pi \Phi / \Phi_0$ , so that a  $2\pi$  phase change corresponds exactly to the transfer of a single quantum of magnetic flux. Equivalently, the integral of the voltage pulse “area” is quantized,

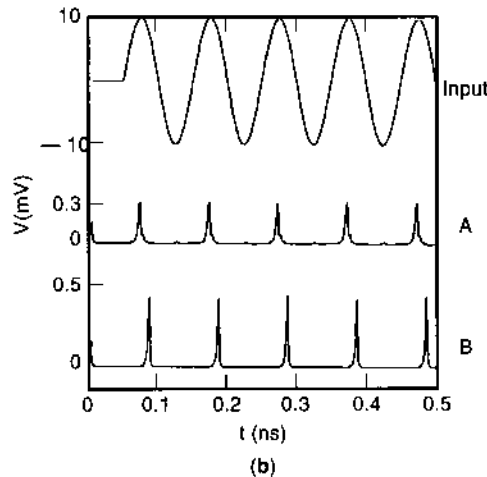
$$\int V(t) dt = \Phi_0 = 2.07 \text{ mV ps} \tag{12}$$

so that the total RSFQ voltage pulse integrated over time equals a quantum of magnetic flux given by Eq. (6). Hence the name single-flux-quantum logic. Also note from Eq. (12) that if the pulse is a few picoseconds wide the amplitude will be much less than 1 mV.

The simplest RSFQ circuit, called the Josephson transmission line (JTL) and shown schematically in Fig. 17(a), consists of the basic RSFQ component in Fig. 16(a) connected in series. The series junction and the inductor at the input convert a dc signal to single-flux quanta (dc-SFQ converter). Intuitively one can image the circuit as a series of pendula connected together by torsion springs (the inductors). The dc bias currents torque the pendulums close to the unstable equilibrium point. When a single-flux quantum pulse is fed in at the left end of the line, it flips the first junction (pendulum). Because this junction is connected to the adjacent junction by an inductor (a spring), when the first junction flips, it makes a current in the inductor which in turn flips the adjacent junction (the next pendulum). Hence in a JTL, a flux quantum input at the left “hops” down the line, flipping adjacent junctions until reaching the output. Figure 17(b) shows a WRspice (35) simulation of the input and output voltage of the JTL with a 10 GHz input.



**Figure 17.** (a) A Josephson transmission line with a dc-SFQ input circuit. The arrows represent the bias currents for the JTL. (b) WRspice simulation with the 10 GHz analog input and the resulting RSFQ voltage pulses measured at points A and B on the line. Single-flux quantum pulses propagate from A to B with a 10 ps delay. The pulse is re-shaped as it propagates from A to B, and the amplitude increases while the width decreases. The simulation assumes  $200 \mu\text{A}$  junctions,  $175 \mu\text{A}$  bias currents,  $3.6 \text{ pH}$  JTL inductors,  $1 \Omega$  resistors, and a  $5 \text{ pH}$  parallel inductor in the input dc-SFQ converter. The input is a 10 GHz 10 mV amplitude sine wave and a 10 mV dc bias both applied to  $50 \Omega$  resistors.



The clock signals for RSFQ circuits are single-flux-quantum pulses. The logical convention is that a “one” corresponds to an RSFQ pulse within a clock period, and a logical “zero” is the absence of a pulse in a clock period. JTLs are used for clock distribution. However one does not have to use JTLs for all data and clock distribution. For example, the digital signal processor in the ADC in Fig. 7 is broken into three components. The RSFQ and data pulses are transferred between modules by superconducting microstrip lines. Because the transmission lines are superconducting, they have low loss and low dispersion. With superconducting transmission lines one can transport RSFQ pulses for relatively large distances on a chip.

The switching speed of the damped junction is a function of the device capacitance. Reducing the junction area correspondingly decreases the junction capacitance and enables the junction (pendulum) to flip faster. Hence smaller junctions make faster RSFQ circuits, providing one scales the other parameters accordingly. For example, a 100  $\mu\text{A}$  critical current Josephson junction with a 2.5  $\mu\text{m}$  linear size corresponds to a 3 ps RSFQ pulse. A 1.25  $\mu\text{m}$  junction corresponds to a 2 ps pulse, and a 0.7  $\mu\text{m}$  junction corresponds to a 1 ps pulse. Lukens and co-workers at the State University of New York at Stony Brook have used e-beam lithography to fabricate simple RSFQ frequency divider circuits with 0.5  $\mu\text{m} \times 0.5 \mu\text{m}$  junctions and 50  $\text{kA}/\text{cm}^2$  critical current density.

Figure 18(a) shows a micrograph of an RSFQ T flip-flop circuit, and Fig. 18(b) shows experimental test data with a schematic of the circuit inset on the graph. The circuit consists of a Josephson transmission line connected to the RSFQ flip-flop, consisting of four Josephson junctions. A dc bias applied to the first junction in the JTL generates a high frequency SFQ pulse train which is input to the flip-flop circuit. The parameters of the SQUID loop are chosen so that a single-flux quantum can circulate in the loop. When the input pulses are injected into the loop, they have the effect of flipping the direction of the circulating flux quanta. Furthermore, if the current is circulating in the clockwise direction, a single flux quantum pulse is also produced at the output,  $V_{\text{out}}$ ; see Fig. 18(b). Therefore, the input flips the circulating current, and a half-frequency RSFQ pulse train is produced.

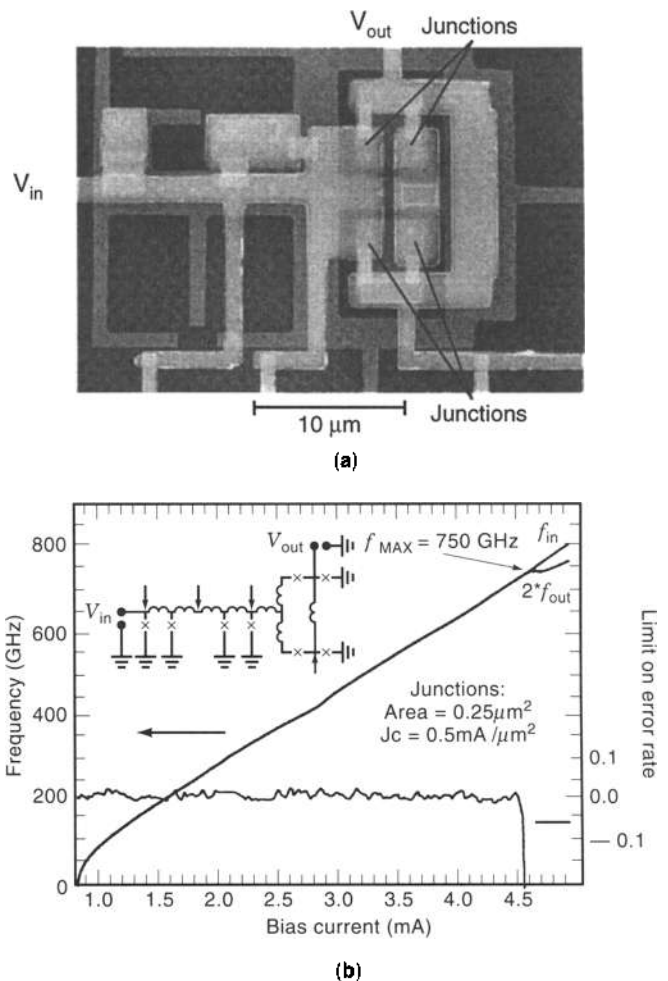
The diagonal plot on Fig. 18(b) is an overlay of the average voltage of the clock and twice the average voltage of the output. Increasing the input bias current increases the frequency of the input clock. Because the output is half the frequency of the input, we expect that, on average, the input voltage should equal twice the output voltage. The two measured quantities are indeed equal, and from the fundamental Josephson Eq. (3) one can calculate that the speed corresponds to 750 GHz.

The experimental data in Fig. 18(b) are dc average value measurements. The circuit operation is too fast to observe the digital outputs directly by room temperature electronics. Therefore, it is not clear how many errors the circuit makes at 750 GHz. The lower trace in Fig. 18(b) is the subtraction of the two curves and gives the error in the experiment. It corresponds very roughly to a limit on the maximum bit error rate in this experiment. The error is less than one part in 10,000 which is the limit of the test equipment, and the true bit error rate should be significantly less.

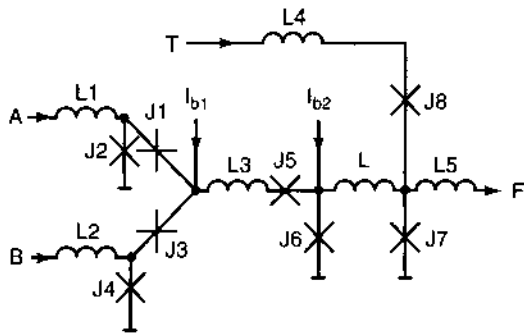
Therefore, RSFQ circuits have the potential of operating at hundreds of gigahertz clock rates. So far we have described the JTL, which can be used to transport RSFQ pulses, and

the T flip-flop which is the basic memory storage element. To make a complete logic family, Likharev, Mukhanov, and Semenov demonstrated OR, AND, and inversion functions for RSFQ (28). As one example, Fig. 19 shows the RSFQ 2-input OR gate.

The OR gate consists of two Josephson junctions J2 and J4 at the input connected to the inductor L3 and the junction J5. The junctions J1 and J3 are for isolation and stop feedback of RSFQ pulses to the input. Parameters are chosen so that an input RSFQ pulse at either A or B flips the junction J6. This junction is connected to a flip-flop circuit (J6, L, J7) which is a SQUID loop with an RSFQ readout. When junction J6 is switched by the input at either A or B, a single-flux quantum is held in the SQUID loop J6, L, and J7. The SQUID loop acts as a latch until the clock is applied. Quanta stored in the SQUID loop are read out by a clock pulse applied to the junction pairs J7 and J8. Parameters are chosen so that the clock pulse resets the SQUID loop and produces a flux quantum at F if there has been an input at A or B. This is the timed OR function. For



**Figure 18.** (a) Micrograph of an RSFQ T flip-flop fabricated using e-beam lithography. (b) Test data for a T flip-flop fabricated with 0.5  $\mu\text{m} \times 0.5 \mu\text{m}$  Josephson junctions. The circuit operates correctly up to a speed corresponding to 750 GHz. Photograph and data courtesy of W. Chen, A. V. Rylyakov, V. Patel, and J. E. Lukens at State University of New York at Stony Brook.

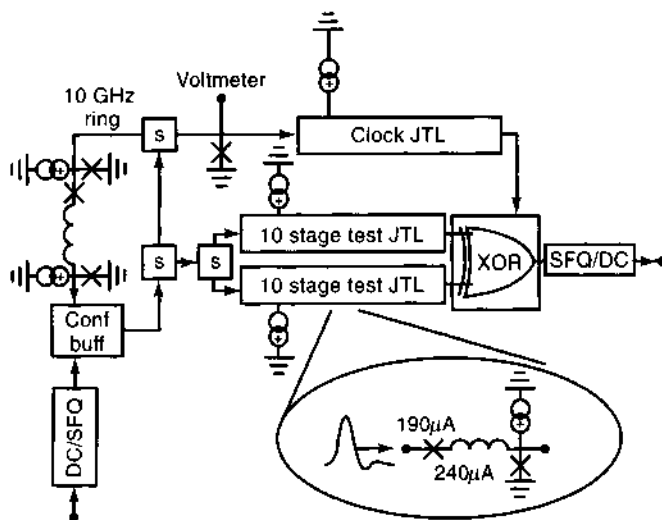


**Figure 19.** The basic RSFQ OR gate. The picosecond RSFQ pulses race through the circuit and are held in the SQUID loop J6, L, and J7 "latch." A clock RSFQ pulse is applied to the junction pair J8 and J7 to read out the latch. Original figure from Likharev and Semenov (23). © 1991 IEEE.

more detailed information on the complete RSFQ logic family, see Ref. 23.

Figure 20 is a schematic of a simple RSFQ circuit designed by Herr and Feldman at the University of Rochester to test bit error rates (40). The circuit consists of a 10 GHz RSFQ ring oscillator, which generates the clock, and two JTL stages connected to an XOR gate. Identical pulse trains from the clock propagate down the JTL stages, so that the XOR function should be logical "zero" for all correct outputs. Any error in the JTL stages produces an RSFQ pulse at the output of the XOR gate. The Rochester team measured the continuous operation of the circuit for nine days and measured a bit error rate of  $5 \times 10^{-17}$ . This error rate corresponds to 4 errors in 150 h ( $5\frac{1}{2}$  days). These results matched well with thermal noise analysis of errors for a single junction. Therefore, RSFQ circuits have the potential for ultrahigh-speed operation with very low bit error rates.

Ruck and coworkers in Jülich have done the same experiment with a circuit fabricated using high temperature superconductors, and have measured bit error rates less than  $10^{-11}$  at 39 K (41). However, note that in these experiments the circuit



**Figure 20.** Schematic diagram of the test circuit to measure RSFQ bit error rates. Figure reproduced courtesy of Q. Herr and M. Feldman, University of Rochester, and Applied Physics Letters (41).

under test is rather simple and the bit error rates may increase for complex RSFQ circuits.

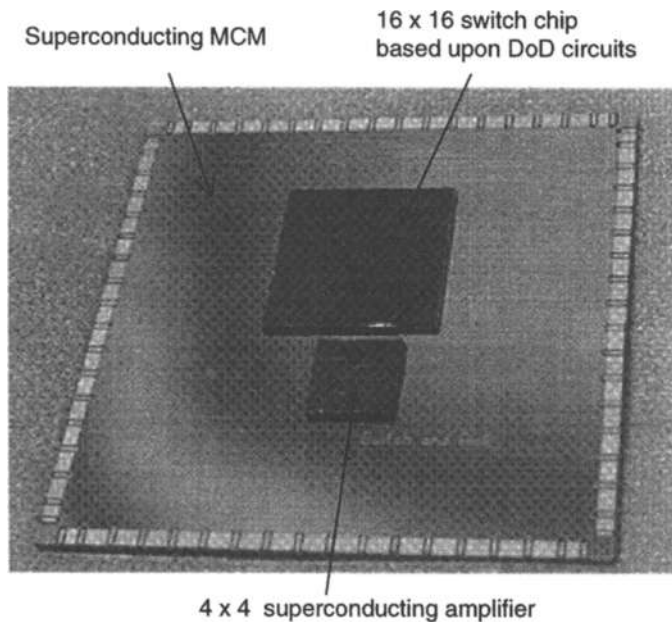
Picosecond RSFQ voltage pulses cannot be transferred between chips or to room temperature electronics with existing packaging. This is because they are easily attenuated or reflected by impedance mismatches and conventional room temperature electronics is simply neither fast enough nor sensitive enough to detect small picosecond signals. Therefore, the RSFQ pulses must be converted to lower frequency voltages at the output (23). Recently, researchers at HYPRES and Conductus have developed asynchronous amplifier circuits, so that one can interface RSFQ pulses directly to room temperature electronics at clock rates up to 8 GHz (42). High-speed testing has been performed on-chip at 20 GHz to 40 GHz by loading shift registers with data at low speed. Then these data are clocked through an RSFQ circuit at high speed, and the outputs are collected at high speed in shift registers. Finally, the output data are read out from the shift registers at low speed to verify the correct operation of the circuit.

One advantage of RSFQ over voltage-state logic is that it requires only dc power. The power dissipation of RSFQ circuits is approximately 10,000 times lower than GaAs room temperature electronics and approximately 50 times less than voltage-state logic. Therefore, thermal management is significantly easier for superconducting RSFQ digital circuits than for other room temperature technologies such as GaAs. The main challenge for RSFQ circuits and indeed any circuit operating above 10 GHz is clock distribution. For example, at 100 GHz there is only 10 ps between clock pulses. Assuming propagation at the speed of light, 10 ps corresponds to 3 mm. Because typical complex circuits are larger than 3 mm, global clock distribution is not possible at this speed. Researchers at UC Berkeley and elsewhere have been working to develop new types of asynchronous timing schemes to enable complex circuit operation at ultrahigh gigahertz clock rates (23,43).

## ADVANCED APPLICATIONS AND FUTURE DIRECTIONS FOR DIGITAL SUPERCONDUCTING ELECTRONICS

We are at an exciting point in the development of superconducting electronics. Fabrication techniques and design tools are enabling the development of the first real systems. For example Conductus, TRW, Stanford University, and UC Berkeley, funded through the US Department of Commerce Advanced Technology Program (ATP), collaborated to demonstrate a superconducting cross-bar packet switch. The system was packaged in a 4 K closed cycle refrigerator, all components were operational at 10 Gb/s, and the complete integrated system was demonstrated at 4 Gb/s (44). Although this system is not a final commercial product, it demonstrates the possibility of integrating and packaging superconducting circuits.

A significant accomplishment of the ATP project was the development of a superconducting multichip module (MCM) technology by TRW. Figure 21 shows a superconducting MCM with two chips used in the cross-bar demonstration flipped on top. Using reflow solder bumping and superconducting transmission lines between chips, researchers at TRW and Conductus demonstrated 10 Gb/s data transfer rates between chips and room temperature electronics. This MCM technology will facilitate the development of complex superconducting digital circuits from smaller components.



**Figure 21.** MCM module used in the ATP demonstration of a superconducting packet switch by Conductus, TRW, Stanford University and University of California, Berkeley. Both chips were fabricated using a  $2\ \mu\text{m}$ ,  $2\ \text{kA}/\text{cm}^2$  Nb Josephson fabrication technology. Photograph courtesy of G. Akerling, A. Smith, K. Yokoyama, and J. Spargo at TRW Space and Electronics.

Japanese industrial laboratories, such as Hitachi and NEC, continue to work on digital superconducting electronics. Recently completed projects at Hitachi include the development of a  $4 \times 4$  packet switch using voltage-state logic (45). NEC researchers have demonstrated a system consisting of three computers connected by a superconducting parallel-pipelined ring (26,46). Components of these systems, consisting of several hundred logic gates, have been demonstrated at multi-gigahertz data rates. The complete NEC system operated at 100 MHz, with an estimated ultimate throughput of 10 Gbit/s (46). NEC is working to develop micron size Josephson fabrication technologies in order to increase speed and integration density of their superconducting RAM. In 1998 the Japan Science and Technology Agency (STA) initiated a three year program for NEC, Hitachi, ETL, Fujitsu, Tokyo University, and Nagoya University. The goal of this program is to develop the core technology in Japan for ultrahigh-speed RSFQ logic circuits using both low-temperature and high-temperature superconductors.

Perhaps the most promising applications are in analog-to-digital conversion using RSFQ logic (see Fig. 7). Analog-to-digital converters (ADCs) are ideal candidates for superconducting electronics applications because all of the high-speed operations are internal to the circuit, and the outputs to room temperature electronics are relatively low speed. Furthermore, ADC circuits do not require frequent access to cache memory, which is at present difficult to implement in superconducting circuits. HYPRES Inc. in Elmsford, NY (18) has done considerable research to develop advanced ADC circuits.

Figure 22(a) is a 6-bit ( $2^6$  level) flash type ADC developed at HYPRES Inc. The circuit consists of six comparators. The analog input signal is applied to a resistor divider network

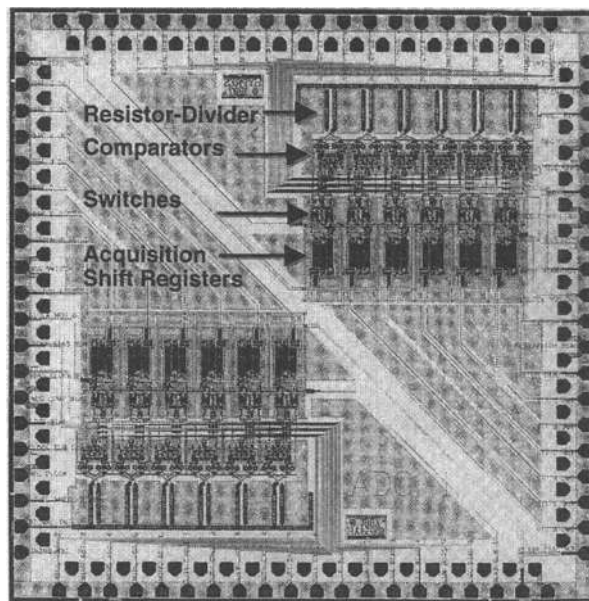
which feeds a one-junction SQUID at the input of each comparator. RSFQ circuits are used for ultrafast sampling of the one-junction SQUID current. With certain parameters (that is,  $\beta_L \leq 1$ ), the output current of the one-junction SQUID is periodic in the input current. Hence by using a resistor divider network, the one-junction SQUID currents make an interference pattern gray code for the input signal. Picosecond RSFQ sampling of this gray code gives a binary representation of the input analog signal. The circuit is compact, and requires only one comparator for each bit because the unique properties of the one-junction SQUID. To store the digitized data, a 32-word shift register memory is integrated within the ADC.

Figure 22(b) shows test data for the flash ADC for a 16 GHz sampling rate. The input signal (inverted by the amplifier) was acquired in real time, and the binary data were stored in a 32 bit shift register for low speed output to room temperature electronics. Research is in progress to directly interface this ADC to room temperature electronics, so that the performance specifications (spurfree dynamic range, effective number of bits, etc.) can be measured.

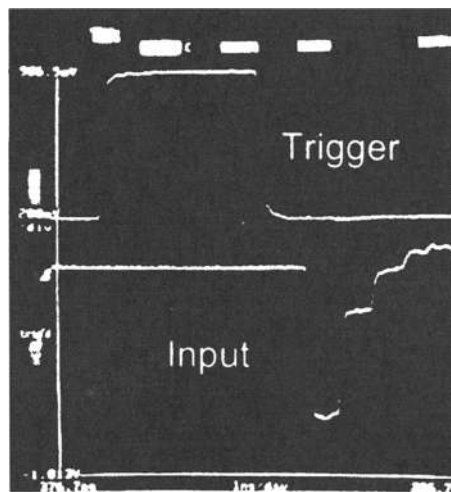
A fundamental advantage of superconducting electronics is low on-chip power dissipation compared to room temperature electronics, such as GaAs. Therefore, thermal management on-chip is significantly easier than with GaAs. Furthermore, compared to all other room temperature technologies, RSFQ has a raw speed advantage because digital operation is possible in excess of 100 GHz. These advantages make superconducting electronics a candidate for ultrahigh-speed supercomputer applications of the future. In fact, recent research has indicated the feasibility of a  $10^{12}$  floating-point operations per second (petaFLOP) RSFQ-based supercomputer. The petaFLOPs computer project, presently funded by the US Defence Advanced Projects Agency (DARPA), is studying the possibility of incorporating thousands of 100 GHz superconducting RSFQ vector processors (47).

The challenge for future digital superconducting technology is to unambiguously demonstrate circuits operating in excess of 100 GHz. These circuits will require small area ( $\sim 0.8\ \mu\text{m}$  linear size), higher current density ( $J_c \approx 10\ \text{kA}/\text{cm}^2$ ) junctions, and one must first demonstrate that a fabrication process can make hundreds of  $0.8\ \mu\text{m}$   $10\ \text{kA}/\text{cm}^2$  Josephson devices with small parameter spreads. In addition exact clocking of the ultrafast circuits needs to be achieved. At 100 GHz there is only a 10 ps window for a clock period, and at 200 GHz this window shrinks to 5 ps. Parameter variations and propagative time delays can easily introduce timing errors which will significantly limit the speed of complex circuits. Finally, the circuits must be tested unambiguously at ultrahigh speed. This is no easy task, since demultiplexer (DEMUX) circuits are needed to interface with room temperature electronics, which increases the complexity of even the simplest 100 GHz RSFQ circuit.

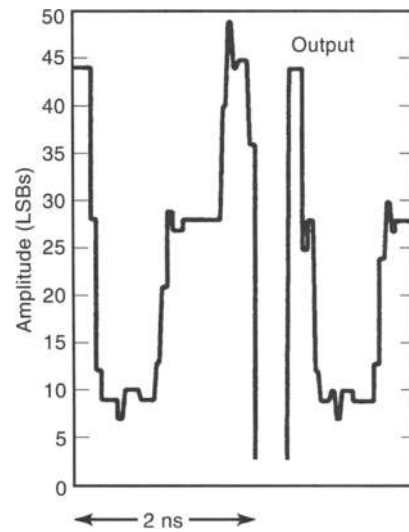
In the near future all systems based on digital superconducting circuits are expected to use conventional 4 K niobium superconductor technology that requires a relatively large refrigerator. For applications requiring low power and a small form factor, niobium nitride (NbN),  $T_c = 17\ \text{K}$ , will most probably become the processing technology of choice for digital superconducting circuits. NbN circuits operate at 10 K and use small pulse-tube or Stirling coolers described in the introduction. However, research and development is required for NbN fabrication technology to reach its full potential.



(a)



(b)



(c)

**Figure 22.** (a) HYPRES six-bit flash analog-to-digital converter fabricated on a  $1\text{ cm} \times 1\text{ cm}$  chip. Experimental test data (b) input to the circuit, and (c) collected in real time at 16 Gs/s. Photograph and data provided courtesy of S. Kaplan, S. Rylov, D. Gaidarenko, W. Li, and P. Bradley at HYPRES Inc. (18).

## ACKNOWLEDGMENT

The author would like to acknowledge gratefully the many researchers and laboratories that contributed to this article. He would also like to thank T. Van Duzer, John Rowell, K. Likharev, John Spargo, John Clarke, Oleg Mukhanov, and Andy Smith for some suggestions on the original manuscript.

## BIBLIOGRAPHY

1. For a review of superconducting electronics with many detailed references, see T. Van Duzer, Superconductor electronics 1986–1996, *IEEE Trans. Appl. Supercond.*, **7** (2): 98–111, 1997.
2. The classic text on superconducting electronics: T. Van Duzer and C. W. Turner, *Principles of Superconducting Devices and Circuits*, New York: Elsevier, 1981.
3. Another good text on superconductive electronics is K. K. Likharev, *Dynamics of Josephson Junctions and Circuits*, New York: Gordon and Breach, 1986.
4. J. G. Bednorz and K. A. Muller, Possible high  $T_c$  superconductivity in the Ba-La-Cu-O system, *Z. Phys. B. Condens. Matter*, **64** (2): 189–193, 1986.
5. For a good introduction to superconductivity at the undergraduate level, see R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman Lectures on Physics*, Menlo Park, CA: Addison-Wesley, 1966, Vol. 3, Chap. 21.
6. For a very readable introduction to the basic principles of superconductivity, see A. C. Rose-Innes and E. H. Rhoderick, *Introduction to Superconductivity*, Int. Ser. Solid State Phys., New York: Pergamon, 1978, Vol. 6.
7. For a more detailed introduction to superconductivity at the graduate level, see M. Tinkham, *Introduction to Superconductivity*, Malabar, FL: Krieger, 1980.



8. For more information on magnetic levitation trains, see Railway Technical Research Institute [Online], 1998. Available: <http://www.rtri.or.jp>
9. D. Zhang et al., Microstrip filters for wireless communications using high-temperature superconducting thin films, *Appl. Supercond.*, **3** (7–10): 483–496, 1995.
10. Conductus [Online]. Available: <http://www.Conductus.com>
11. For a very readable description of the BCS theory, see J. R. Schrieffer, *Theory of Superconductivity*, New York: Benjamin, 1964.
12. For the original paper on the Josephson effect, see B. D. Josephson, Possible new effects in superconducting tunneling, *Phys. Lett.*, **1**: 251, 1962; and for the first experimental verification, see P. W. Anderson and J. M. Rowell, Possible observation of the Josephson superconducting tunneling effect, *Phys. Rev. Lett.*, **10**: 230, 1963; for the original report of a dc SQUID see R. C. Jaklevic et al., Quantum interference effects in Josephson tunneling, *ibid.*, **12**: 159, 1964.
13. For a good introduction to SQUIDs and applications, see J. Clarke, SQUID fundamentals, in H. Weinstock (ed.), *SQUID Sensors: Fundamentals, Fabrication and Applications*, Proc. NATO ASI SQUID Sensors: Fundam. Fabric. Appl., Dordrecht, The Netherlands: Kluwer, 1996, p. 1.
14. For more details on the RF SQUID and references see A. H. Silver and J. E. Zimmerman, Josephson weak-link devices, in V. L. Newhouse (ed.), *Applied Superconductivity*, New York: Academic Press, 1975, vol. 1.
15. For a good review of scanning SQUID microscopes and applications see John Kirtley, Imaging magnetic fields, *IEEE Spectrum*, 40–48, Dec. 1996.
16. T. S. Lee, E. Dantsker, and J. Clarke, High transition temperature SQUID microscope, *Rev. Sci. Instrum.*, **67**: 4208, 1996.
17. For a review of the voltage standard see C. A. Hamilton, C. J. Burroughs, and S. P. Benz, Josephson voltage standard—review, *IEEE Trans. Appl. Supercond.*, **7** (2): 3756–3761, 1997.
18. For process design rules and product information, see <http://www.HYPRES.com>
19. See for example D. J. Durand et al., The distributed Josephson inductance phase shifter, *IEEE Trans. Appl. Supercond.*, **2**: 33–38, 1992, and J. Luine et al., Measurements of YBCO parametric amplifiers, *J. Appl. Supercond.*, in press.
20. See as an example and for references M. Jeffery, Subquantum limit Josephson magnetometry, *Appl. Phys. Lett.*, **66** (21): 2897, 1995.
21. A. Spooner et al., Superconducting direct digital synthesizer, *IEEE Trans. Appl. Supercond.*, **7** (2): 2270–2273, 1997.
22. For a review of voltage-state logic, see S. Hasuo and T. Imamura, Digital logic circuits, *Proc. IEEE*, **77**: 1177–1193, 1989.
23. For an excellent introduction to RSFQ, see K. K. Likharev and V. K. Semenov, RSFQ logic/memory family: A new Josephson-junction technology for sub-terahertz-clock-frequency digital systems, *IEEE Trans. Appl. Supercond.*, **1**: 3–28, 1991.
24. For a review of the IBM superconducting computer project, see the special issue of the IBM Journal of Research and Development: Josephson computer technology: An IBM research project, *IBM J. Res. Develop.*, **24** (Spec. Issue): March, 1980.
25. N. Fujimaki et al., Josephson modified variable threshold logic gates for use in ultrahigh-speed LSI, *IEEE Trans. Electron Devices*, **36**: 433–446, 1989.
26. S. Tahara et al., Josephson memory technology, *IEICE Trans. Electron.*, **E79-C**: 1193–1199, 1996.
27. M. Jeffery, W. Perold, and T. Van Duzer, Superconducting complementary output switching logic operating at 5–10 Gb/s, *Appl. Phys. Lett.*, **69**: 2746–2748, 1996.
28. K. K. Likharev, O. A. Mukhanov, and V. K. Semenov, Resistive single flux quantum logic for the Josephson junction technology, in *SQUID'85*, D. Hahlbolm and H. Lübbig (eds.), Berlin: de Gruyter, 1985, pp. 1103–1108.
29. V. K. Semenov, Yu. A. Polyakov, and D. Schneider, Implementation of oversampling analog-to-digital converter based on RSFQ logic, *Ext. Abst. ISEC'97*, 1997, Vol. 1 (Plenary Contrib.), pp. 41–43 (Rep. D6).
30. E. Goto, Y. Wada, and K. F. Loe (eds.), *Advances in Quantum Flux Parametron Computer Design*, Singapore: World Scientific, 1991.
31. G. L. Kerber et al., An improved NbN integrated circuit process featuring thick NbN ground plane and lower parasitic circuit inductances, *IEEE Trans. Appl. Supercond.*, **7** (2): 2638–2643, 1997.
32. For information on the Sumitomo and CTI cryocoolers see <http://www.janis.com> and for the Iwatani pulse tube see <http://www.kelvinic.com>
33. M. Jeffery et al., Magnetic imaging of moat-guarded superconducting electronic circuits, *Appl. Phys. Lett.*, **67**: 1769–1771, 1995.
34. For a systematic treatment of this and other examples, see A. Th. A. M. De Waele and R. De Bruyn Ouboter, Quantum-interference phenomena in point contacts between superconductors, *Physica*, **41**: 225, 1969.
35. For a good commercial Josephson circuit simulator see WRSpice at <http://www.srware.com>, for the public domain JSIM see <http://www.cryo.eecs.berkeley.edu>, and for HSPICE see <http://www.avantecorp.com>. The Josephson junction model is given in the Appendix to M. Jeffery et al., Monte Carlo optimization of superconducting complementary output switching logic circuits, *IEEE Trans. Appl. Supercond.*, **8** (3): 104–119, 1998.
36. T. A. Fulton, Punchthrough and the tunneling cryotron, *Appl. Phys. Lett.*, **19** (9): 310–313, 1971.
37. R. Jewett and T. Van Duzer, Low probability punchthrough in Josephson junctions, *IEEE Trans. Magn.*, **MAG-17** (1): 599–602, 1981.
38. J. P. Hurrell and A. H. Silver, SQUID digital electronics, *AIP Conf. Proc.*, **44**: 437–447, 1978.
39. G. Oya, M. Yamashita, and Y. Sawada, Single flux quantum 4JL-interferometer operating in the phase mode, *IEEE Trans. Magn.*, **MAG-21**: 880–883, 1985.
40. Q. P. Herr and M. J. Feldman, Error rate of a superconducting circuit with and without an applied magnetic field, *Appl. Phys. Lett.*, **69**: 694–695, 1995.
41. B. Ruck et al., Measurement of the dynamic error rate of a high temperature superconductor rapid single flux quantum comparator, *Appl. Phys. Lett.*, **72** (18): 2328–2331, 1998.
42. O. A. Mukhanov et al., Josephson output interfaces for RSFQ circuits, *IEEE Trans. Appl. Supercond.*, **7** (2): 2826–2831, 1997.
43. Z. J. Deng et al., Data-driven self-timed RSFQ digital integrated circuit and system, *IEEE Trans. Appl. Supercond.*, **7** (2): 3634–3637, 1997.
44. For an easy to read overview, see N. Dubash and R. Simon, Hybrid cryoelectronic communication systems—the future in ultrahigh performance communications devices, *Supercond. Ind.*, **10** (3): 18–22, 1997.
45. M. Hosoya et al., Contention solver for a superconducting packet switch, *IEEE Trans. Appl. Supercond.*, **7** (2): 2274–2275, 1997.
46. For a review of the NEC ring system and test results, see S. Yorozu et al., System demonstration of a superconducting communication system, *IEEE Trans. Appl. Supercond.*, in press.
47. For general information on the program, see <http://www.darpa.mil/ito/Summaries97/F187/0.html>, and for information on the superconducting RSFQ component, see <http://gamayun.physics.sunysb.edu/RSFQ/Research/PetaFLOPS>

MARK JEFFERY  
University of California at Berkeley

## SUPERCONDUCTIVITY: HYSTERESIS AND COUPLING LOSSES

A key issue for most applications of superconductivity involves ac losses. Designers need to understand the mechanisms of ac losses in order to lay out the conductors and windings correctly and to predict the performance range in operation. In devices operating at the grid frequency (transformers, current limiters, generators, motors, and power transmission cables) ac losses mostly affect the cryogenic load and hence the overall efficiency of the application. In large pulsed windings [fusion magnets, superconducting magnetic energy storage (SMEs)] the coupling current loss affects the stability, that is, the ability of the superconductor to withstand magnetic field transients. The dc magnetization of a superconductor, which leads to hysteresis loss, is a crucial issue in applications such as imaging and accelerator magnets, where the residual, low field magnetization of the superconductor affects the quality (linearity and homogeneity) of the generated magnetic field.

In normal conducting materials, both dc and ac losses are due to the finite electrical conductivity. At first glance, it appears contradictory that ac losses can occur in a superconducting material, with zero electrical resistivity.

The nonreversible magnetization behavior of bulk superconductors was known well before a practical superconducting wire was manufactured; afterward, it was called "ac loss", which sounds much more negative than "magnetization loop." For single-core wires and tapes, the only ac loss was the hysteresis loss, with the associated disruptive flux jumps. For multifilamentary composites, a new source of loss (filamentary coupling loss) was identified and stimulated the development of very sophisticated strand layouts, with high resistivity barriers and mixed matrices; on the other hand, flux jumps were no longer an issue. Later on, with large stranded conductors, the biggest concern was cable loss (interstrand coupling loss): the effort to limit ac losses was focused on the cable layout rather than the internal structure of the multifilamentary composite.

Because of ac losses, the competitiveness of superconductors has been limited in many fields of applied electrical engineering. Most of the superconducting coils in use today (commercial and research) are operated in dc mode, with limited ac loss occurring during the slow charging up. In those coils, the fastest field change is seen during a safety discharge, when concern about heating or even quenching the conductor is small.

Both hysteresis and coupling current losses occur in a time-varying magnetic field, but their relative weights depend on the specific application. In power transmission cables, the hysteresis loss is by far the largest source, as well as in small, slow rate, potted windings. In large cabled conductors for big SMES and fusion poloidal field coils, the coupling current loss dominates. The hysteresis loss per unit volume is a function of the filament size. The coupling current loss per unit volume increases with the conductor size: with few exceptions, the larger the conductor, the bigger the weight of the coupling current loss.

The physical mechanism of ac losses in superconductors is no longer the object of baseline research. Most R&D activities are devoted to developing low loss, stable conductors. For design purposes, the ability to accurately calculate (i.e. predict) the ac loss during operation is crucial for a reliable and cost-effective engineering approach.

## HYSTERESIS LOSS

In type I superconductors—for example, pure metals with defect-free lattices—the magnetic field does not penetrate the bulk of the material and the superconducting shielding currents flow only at the surface. Such materials have reversible magnetization and no hysteresis loss, with  $B = 0$  and  $M = -_0H$  (Meissner effect) inside the material.

In type II superconducting materials used for practical applications, both low and high  $T_c$ , the surface shielding currents have a marginal role for the magnetization and the flux penetrates the bulk of the conductor. The nonreversible magnetization of type II superconductors is the reason for the hysteresis loss (1–3). The energy loss per unit volume of superconducting material,  $Q$ , can be written in a general form for a closed cycle of applied magnetic field  $B$  (e.g., an oscillation at grid frequency, a charge–discharge cycle of a magnet, or the superposition of an ac field on the background field)

$$Q = \oint M(B) dB \quad (\text{J/M}^3) \quad (1)$$

where  $M$  is the average value of the magnetization inside the superconductor. The explicit expression for the local magnetization, needed to evaluate the integral, depends on the superconductor geometry and on the model selected to describe the magnetic flux penetration and the flux profiles inside the superconductor.

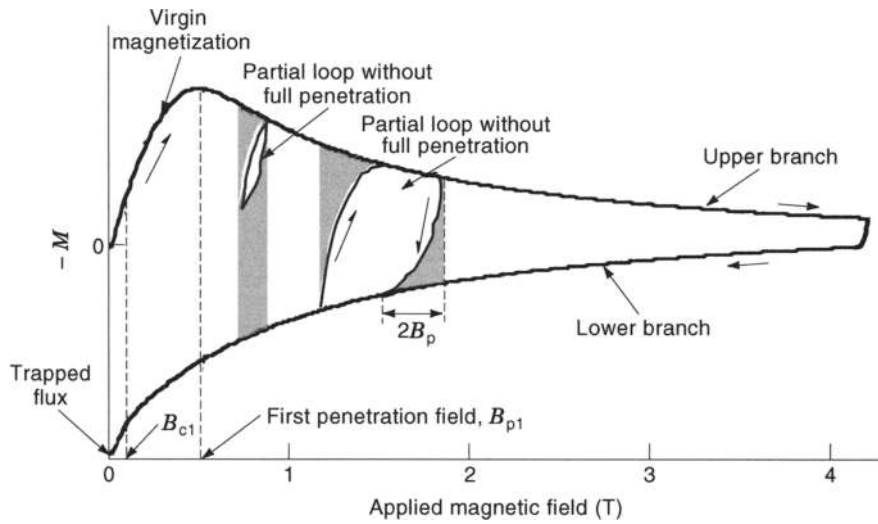
The formulas quoted below are for cylindrical superconducting filaments of diameter  $D$ : whenever the real cross section of the filaments is not round (e.g., oval, dendritic, or hollow filaments or clusters of bridged filaments), the parameter  $D$  in the formulas should be considered as an equivalent diameter. For tapes and flattened filaments with high aspect ratio (e.g., for some high  $T_c$  conductors), the formulas for an infinite slab can be used as a convenient approximation.

### The Magnetization Curve

A dc magnetization curve for a type II superconductor is shown in Fig. 1. After cooldown, at zero field,  $M = 0$ . As an external magnetic field is initially applied, the shielding currents at the filament surface prevent the flux penetration into the bulk superconductor. The diamagnetism is perfect (i.e.,  $M = -_0H$ ) as long as the applied field does not exceed the first critical field,  $B_{c1}$ . In a type II superconductor, total flux exclusion (the Meissner effect) occurs only at the beginning of the first, *virgin* magnetization: if the superconductor is cooled down in the presence of a magnetic field, no flux exclusion occurs.

Above  $B_{c1}$ , the average magnetization increases until the flux penetrates to the center of the filament, at  $B = B_{p1}$ . As the field increases, the diamagnetism decreases (upper branch of the curve) and the magnetization eventually becomes 0 at the upper critical field,  $B = B_{c2}$  (not shown in Fig. 1). When the field is decreased, the flux profiles reverse their gradient in the filament and the average magnetization is  $>0$  (lower branch of the curve). At  $B = 0$ , the flux trapped in the filaments is called *residual magnetization*.

The magnetic flux enters the filament as discrete flux quanta. The diffusion of the flux quanta in the bulk type II superconductor is restrained by the *pinning centers*, which establish field gradients (flux profiles) inside the filaments. The

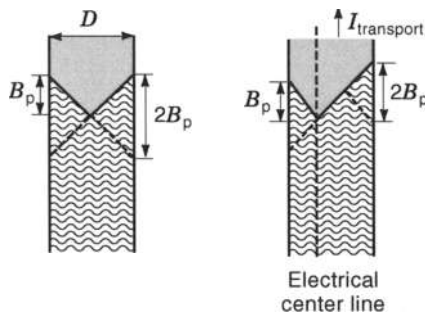


**Figure 1.** Typical magnetization loop of a type II superconductor and related terminology.

pinning centers are microscopic normal zones (e.g., metallurgical or lattice defects) that provide a potential hole to trap the flux quantum. According to the critical state model, a supercurrent encircles each pinning center: the strength of the pinning centers (i.e., their ability to hold the flux quanta) is a function of the fraction of the external field to the critical field,  $b = B/B_{c2}(T, \epsilon)$  where  $\epsilon$  is the mechanical strain. Strong pinning centers are able to store large density of magnetic energy. The magnetization current density  $J_c$  is directly proportional to the volumetric density of the pinning forces in the superconductor.

The penetration field  $B_p$  is the field difference between the filament surface and the electrical center line (see Fig. 2): in Bean's model (4),  $J_c$  is assumed constant inside the filaments for a given  $b$ , and  $B_p$  is a linear function of the critical current density. For a round superconducting filament of diameter  $D$ , the penetration field  $B_p(b)$ , in perpendicular and parallel orientation, is proportional to the density of the magnetization currents flowing respectively in the axial ( $J_{c\parallel}$ ) and the azimuthal ( $J_{c\phi}$ ) direction (5, 6)

$$B_{p\perp}(b) = \frac{\mu_0 D J_{c\parallel}(b)}{\pi} \quad (\text{T}), \quad B_{p\parallel}(b) = \frac{\mu_0 D J_{c\phi}(b)}{2} \quad (\text{T}). \quad (2)$$



**Figure 2.** Penetration field in a superconducting cylinder of diameter  $D$  in parallel applied field (or infinite slab of thickness  $D$ ), without (left) and with (right) transport current. The dashed profile models the lower branch of the magnetization curve.

For an infinite slab of thickness  $2a$ , with applied field parallel to the slab surface, the penetration field is

$$B_p(b) = \mu_0 a J_c(b) \quad (\text{T}) \quad (3)$$

A partial magnetization loop occurs when a small external field variation is superimposed on a background dc field. If the amplitude of the field change,  $\Delta B$ , is larger than  $2B_p$ , full penetration is achieved, that is, the magnetization moves from the upper to the lower branch of the curve.

It is questionable how far Bean's model for linear flux profiles is an acceptable approximation. When the penetration field is much smaller than the applied field,  $B_p \ll B$ , the  $J_c$  variation over  $B_p$  can reasonably be neglected. As  $B_p$  is proportional to the filament size, the range of field over which Bean's model reliably applies is larger for thin-filament superconductors. Whenever Bean's model is not considered adequate, an explicit formula for  $J_c(b)$ , such as the expression proposed by Kim (7)

$$J_c = \frac{\alpha}{B_0 + b} \quad (4)$$

must be substituted in Eqs. (2) and (3) ( $B_0$  and  $\alpha$  are fitting parameters). The magnetization and loss formulas become more complex using Eq. (4), but the improvement in the accuracy of the loss results is not dramatic. At very low field, when the linear profile approximation is rough, the parameters in Eq. (4) cannot be satisfactorily fitted by experimental results, as a direct measurement of  $I_c$  close to 0 field is impossible due to the self-field and instabilities. The formulas below apply only to the linear profile assumption. An example of hysteresis loss formulas using a nonlinear profile approach can be found in Ref. 8.

Integrating in cylindrical coordinates the linear flux profiles from Eqs. (2) and (3) over the filament volume, the upper and lower branches of the magnetization curve are obtained as explicit functions of the critical current density and filament diameter. In perpendicular and parallel applied field orientations the average magnetization per unit volume, according to Ref. 6, is

$$M_{\perp}(b) = \frac{2D J_{c\parallel}(b)}{3\pi} \quad (\text{T}), \quad M_{\parallel}(b) = \frac{D J_{c\phi}(b)}{6} \quad (\text{T}) \quad (5)$$

Table 1. Summary of Hysteresis Loss<sup>a</sup> Formulae

|                                 | Infinite Slab<br>Parallel to Field                                   | Cylinder<br>Parallel to Field  | Cylinder<br>Perpendicular to Field  |
|---------------------------------|--|--|---|
| $\Delta B \leq 2B_p$            | $\frac{\Delta B^3}{12\mu_0 B_p}$                                     | $\frac{\Delta B^3}{3\mu_0^2 D J_{c\phi}} \left(1 - \frac{\Delta B}{2\mu_0 D J_{c\phi}}\right)$<br>$\equiv \frac{\Delta B^3}{6\mu_0 B_{p\parallel}} \left(1 - \frac{\Delta B}{4B_{p\parallel}}\right)$                              | $\frac{\pi \Delta B^3}{3\mu_0^2 D J_{c\parallel}} \left(1 - \frac{\pi \Delta B}{4\mu_0 D J_{c\parallel}}\right)$<br>$\equiv \frac{\Delta B^3}{3\mu_0 B_{p\perp}} \left(1 - \frac{\Delta B}{4B_{p\perp}}\right)$   |
| $\Delta B \geq 2B_p$            | $\frac{\Delta B B_p}{\mu_0} \left(1 - \frac{4B_p}{3\Delta B}\right)$ | $\frac{\Delta B D \bar{J}_{c\phi}}{3} \left(1 - \frac{\mu_0 D (J_{c\phi}^A + J_{c\phi}^B)}{4\Delta B}\right)$<br>$\equiv \frac{2\Delta B \bar{B}_{p\perp}}{3\mu_0} \left(1 - \frac{B_{p\perp}^A + B_{p\perp}^B}{2\Delta B}\right)$ | $\frac{4\Delta B D \bar{J}_{c\parallel}}{3\pi} \left(1 - \frac{\mu_0 D (J_{c\parallel}^A + J_{c\parallel}^B)}{2\pi \Delta B}\right)$<br>$\equiv \frac{4\Delta B \bar{B}_{p\perp}}{3\mu_0} \left(1 - \frac{B_{p\perp}^A + B_{p\perp}^B}{2\Delta B}\right)$ |
| $\Delta B \gg 2B_p$             | $\frac{\Delta B B_p}{\mu_0}$   | $\frac{\Delta B}{3} D \bar{J}_{c\phi} \equiv \frac{2\Delta B}{3\mu_0} \bar{B}_{p\parallel}$  | $\frac{4\Delta B}{3\pi} D \bar{J}_{c\parallel} \equiv \frac{4\Delta B}{3\mu_0} \bar{B}_{p\perp}$  |
| $Q_{\max}$ at $\Delta B = 2B_p$ | $\frac{\Delta B^2}{6\mu_0}$  | $\frac{\Delta B^2}{6\mu_0}$  | $\frac{\Delta B^2}{3\mu_0}$   |

<sup>a</sup> As energy per unit volume of superconductor,  $Q$ , for a closed cycle of magnetic field  $\Delta B$ .

### Basic Formulas for Hysteresis Loss

Substituting Eq. (5) into Eq. (1), the hysteresis loss for a closed field cycle of amplitude  $\Delta B = B^a - B^b$  is obtained as a function of the average critical current density or the average penetration field, defined by

$$\bar{J}_{c\parallel} = \frac{\int_{B^b}^{B^a} J_{c\parallel}(B) dB}{B^a - B^b}, \quad \bar{B}_{p\perp} = \frac{\int_{B^b}^{B^a} B_{p\perp}(B) dB}{B^a - B^b} \quad (6a)$$

and

$$\bar{J}_{c\phi} = \frac{\int_{B^b}^{B^a} J_{c\phi}(B) dB}{B^a - B^b}, \quad \bar{B}_{p\parallel} = \frac{\int_{B^b}^{B^a} B_{p\parallel}(B) dB}{B^a - B^b} \quad (6b)$$

The results of the integration are summarized in Table 1 for the three cases of an infinite slab of thickness  $2a$  with field parallel to the slab surface and a cylinder with diameter  $D$  perpendicular and parallel to the applied field. The formulas are different for partial penetration ( $\Delta B \leq 2B_p$ ) and full penetration ( $\Delta B \geq 2B_p$ ). A further, easier formula is proposed for  $\Delta B \gg 2B_p$ ; this formula overestimates the loss. The shaded areas in (Fig. 1) give a measure of the excess, which is accounted for by using the formula reported in Table 1 for  $\Delta B \gg 2B_p$ .

For a given  $\Delta B$ , the loss maximum occurs when  $\Delta B = 2B_p$ . The loss maximum,  $Q_{\max}$ , reported in Table 1, is a fraction of the magnetic field energy density; it does not depend on the critical current, critical temperature, strain, or filament diameter. In some cases, it may be useful to use  $Q_{\max}$  to get a feeling for the worst-case loss without performing time-consuming calculations. For a given  $\Delta B$ , the filament parameters  $J_c$  and  $D$  determine the reduced field  $b$  for which  $\Delta B = 2B_p(b)$  is fulfilled, that is, the loss is maximum (9). The loss formulas in Table 1 may also be written as a fraction of  $Q_{\max}$  or of the magnetic field energy density  $\Delta B^2/2\mu_0$ , thus introducing a dimensionless loss factor, which is, for a given geometry, only a function of  $\Delta B/B_p(b)$  (10). In Table 1,  $Q$  is the hysteresis loss per unit filament volume and  $J_c$  is the filamentary critical current density. For some superconducting strands, such as  $\text{Nb}_3\text{Sn}$  and high  $T_c$  superconductors, the noncopper critical current is referred to instead of the filamentary critical current density. Whenever

the exact filament fraction is not known, it is possible to use the loss formulas for  $\Delta B \gg 2B_p$ , replacing  $J_c$  by the noncopper critical current  $I_c$ . The hysteresis loss is then expressed in joules per meter of conductor length.

### Anisotropy and Variable Angle Orientation

Due to the integration path of the flux profiles in the axial and radial direction of the cylindrical filament, the ratio of the magnetization in parallel and perpendicular field orientation is, according to Eq. (5),

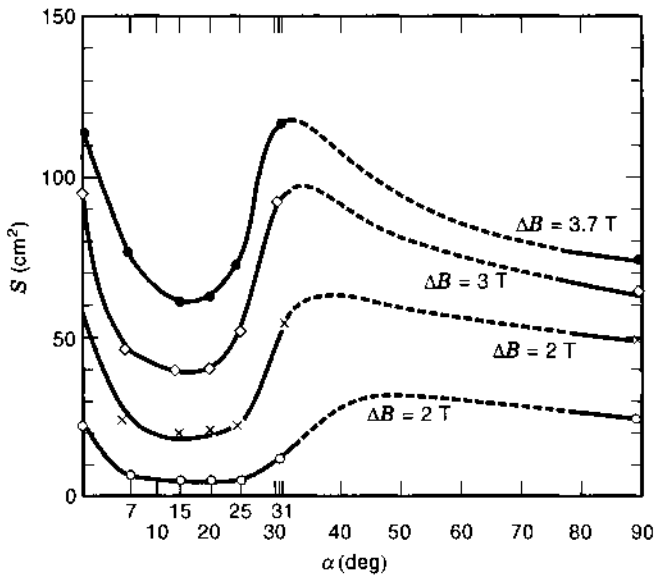
$$\frac{M_{\parallel}}{M_{\perp}} = \frac{4J_{c\phi}}{\pi J_{c\parallel}} \quad (7)$$

The difference observed in the amplitude of magnetization measurements at  $0^\circ$  and  $90^\circ$  orientation is larger than  $4/\pi$  and reveals of the anisotropy of the critical current density, that is,  $J_{c\phi} \neq J_{c\parallel}$ .

In NbTi filaments, the largest source of pinning centers is the precipitation of  $\alpha\text{-Ti}$  and cell dislocation (11). During the manufacturing process (drawing and annealing), the pinning centers are created and strongly oriented in the axial direction. The pinning forces are very anisotropic, resulting in a critical current density much larger in the azimuthal direction than in the axial direction. The ratio  $J_{c\phi}/J_{c\parallel}$  in NbTi conductors is a function of the field and also depends on the manufacturing history (filament size, alloy composition, cold work): from experimental magnetization measurements, the critical current anisotropy is  $J_{c\phi}/J_{c\parallel} \approx 3$  (12).

In  $\text{Nb}_3\text{Sn}$  conductors, the major source of pinning centers is the grain boundaries that form during the reaction heat treatment. The anisotropy of the critical current density is linked to the grain orientation, which is influenced by the heat treatment schedule. The Sn diffuses radially into the Nb filaments, and the  $\text{Nb}_3$  grains are elongated in the radial direction, giving a higher density of the boundary lines for the axial critical current (13). Typical values of the anisotropy in filamentary  $\text{Nb}_3\text{Sn}$  are  $J_{c\phi}/J_{c\parallel} \approx 0.5$  (14).

The anisotropy of the critical current density should not be confused with the variation of the transport critical current



**Figure 3.** Areas of the magnetization loops versus the orientation angle between filaments and applied field. Single core NbTi strand,  $D = 127 \mu\text{m}$ . From Ref. 6. Reprinted from *Cryogenics*, 18, A. P. Martinelli and B. Turck, Some effects of field orientation on the magnetization of superconducting wires, pp. 155–161, copyright 1978, with permission from Elsevier Science.

as a function of the orientation angle  $\alpha$  of the applied field,  $J_c(B_\alpha)$ . The azimuthal critical current density,  $J_{c\phi}$ , to be used for the hysteresis loss in parallel field orientation, is not the same as the transport critical current measured with parallel field orientation,  $J_{c\phi} \neq J_c(B_\parallel)$ , but  $J_{c\parallel} \equiv J_c(B_\perp)$ . For both NbTi and Nb<sub>3</sub>Sn conductors, a larger transport current has been observed in the parallel applied field,  $J_c(B_\parallel) > J_c(B_\perp)$  (6, 14–16).

The orientation of the superconducting filaments in cabled conductors with respect to the cable axis changes continuously over a broad range of angles. For large, multistage conductors, average strand angles of 16° to 25° are commonly observed: the range of the strand angles and its statistical distribution depend on the number of cable stages and the pitch sequence. The hysteresis loss at intermediate angles cannot be interpolated from the formulas in parallel (0°) and perpendicular (90°) fields. The behavior of the loss as a function of the angle has been observed to be not monotonic, with a peak around 30° and a minimum at small angle (<10°); see Fig. 3 from Ref. 6. The interference of the magnetization currents flowing in longitudinal and azimuthal directions distorts the flux profiles and does not allow a practical definition of the penetration field. At angles close to 0°, the flux profiles adjust themselves for consecutive field cycles and the loop area decreases until a reproducible magnetization is obtained after 10 to 20 cycles. An attempt to model the magnetization at intermediate angles can be found in Ref. 17.

### Filament Diameter

The filament diameter is a key parameter for the hysteresis loss formulas. It can be either estimated directly from metallographic examination of the strand cross section or deduced from the magnetization and critical current measurements.

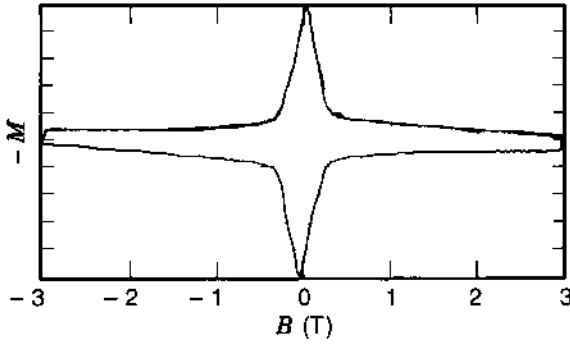
For most of the commercial NbTi strands, the magnetization currents are confined to the individual filaments: perma-

nent currents linking groups of filaments by proximity effects are observed only in very thin, highly packed filaments, with submicron interfilament spacing (11). The critical field for proximity effects,  $B_{c1p}$ , is a function of the temperature, transport current, ratio of spacing to filament size, matrix resistivity and impurities, twist pitch, and sample length (18–20).

For A15 superconductors, as well as for high-temperature superconductors, the estimation of the filament diameter from metallographic investigation is not accurate. In the Nb<sub>3</sub>Sn strands, because of different access to the Sn source (for both bronze method and internal Sn strands), the filaments do not all grow to the same size. A nonreacted Nb core may be left in some region of the filamentary zone, turning the Nb<sub>3</sub>Sn cylinders into hollow cylinders. A major problem affecting the assessment of the filament diameter in Nb<sub>3</sub>Sn composites is *bridging*: when the Nb filaments are tightly packed in the matrix, the Nb<sub>3</sub>Sn layers grow during the reaction heat treatment to build either continuous superconducting links between filaments (21) or mechanical contacts (22), which behave like the proximity effects. The density of bridging is a function of the spacing to filament ratio  $s/d$  (or local area ratio between Cu–CuSn matrix and Nb filaments) and of the heat treatment schedule (23). The superconducting properties of the bridges linking the filaments may be different from those of the bulk filament: at higher field, temperature, or strain, some of the links may become too weak for the magnetization currents. In these cases, the filament diameter is a function of  $b$ . The paths of the magnetization currents in a cluster of randomly bridged filaments cannot be analytically modeled: whenever bridging occurs and the loss formulas for cylinders are used, the equivalent filament diameter must be determined from the magnetization curve, preferably at different  $b$ .

Three main methods are used to derive the filament diameter from the magnetization measurement in a perpendicular field. All methods use Bean's model and assume that the filament critical current density is directly measured in the same field range. The same methods can also be applied to deduce the critical current density once the filament diameter is known (e.g., to estimate  $J_{c\phi}$  or the low field  $J_{c\parallel}$ , when a direct measurement of  $I_c$  is not possible).

- **Diameter from the Penetration Field.** The minimum field change to move from the upper to the lower branch of the magnetization curve (see Figs. 1 and 2) is  $\Delta B_p = 2B_p$ . The filament diameter can be estimated using Eq. (2) or (3) and  $\Delta B_p(B)$  from the magnetization curve. The advantage of this method is that no calibration of the magnetization is necessary to estimate  $\Delta B_p$ .
- **Diameter from the Amplitude of the Magnetization.** The filament diameter can be deduced using Eq. (5) from the amplitude of the magnetization, measured as half of the distance from the lower to the upper branch of the curve. The accuracy of this method is limited by the calibration of the magnetization.
- **Diameter from the Energy Loss of a Closed Field Cycle.** The hysteresis energy loss for a closed field cycle (magnetization loop) can be estimated either by the line integral of the magnetization curve, according to Eq. 1, or by the calorimetric method, after subtracting the coupling loss contribution, if any. If the calorimetric method is used, the magnetization does not need to be calibrated. According to the amplitude of the applied field  $\Delta B (B_{p\perp} > \Delta B \text{ or } B_{p\perp} < \Delta B)$ ,



**Figure 4.** Magnetization loop of a Nb<sub>3</sub>Sn multifilamentary strand with Nb + Ta diffusion barrier. The low field peak of the magnetization is due to the pure Nb shell with diameter  $\approx 0.5$  mm.

the measured energy is compared with the formulas in Table 1 to deduce the filament diameter.

### Crossing the Zero Field

The formulas for magnetization and hysteresis loss have limited validity at low applied fields, especially at the zero-field crossing. On one side, the linear flux profile approximation (Bean's model) is very rough at fields smaller than the first penetration field, and below  $B_{c1}$  the surface screening currents prevent any flux change inside the filament. On the other hand, non-current-carrying superconducting materials are sometimes included for manufacturing reasons in technical superconducting strands, resulting in low field perturbations of the magnetization curve. In soldered cables, the low field superconductivity of the solder may also play a similar role.

In thin filament NbTi strands, a Nb shell encircles each filament to prevent TiCu intermetallic formation during the intermediate heat treatment process, and some Nb<sub>3</sub>Sn suppliers use a Nb layer as diffusion barrier or include it to buffer the Ta or V barrier on the side facing the stabilizer. In both cases, a continuous Nb shell is left on the outer side of the barrier. When a continuous Nb<sub>3</sub>Sn ring grows from the Nb diffusion barrier, its magnetization is as much as one order of magnitude larger than that in the filament (24).

The pure Nb behaves like a *soft* type II superconductor, with  $B_c \approx 0.18$  T. The effect of the screening currents in the Nb layer on the outer side of the diffusion barrier, whose diameter is more than 100 times larger than that of the filament, can be clearly recognized as a low field peak in the magnetization curve; see Fig. 4.

Whenever an anomaly of the magnetization curve occurs at the zero crossing, large errors are likely in the hysteresis loss calculation. If the filament diameter is derived from microscopic examinations or from the higher field magnetization, the loss at low field will be substantially underestimated by the loss formulas. In contrast, if the energy loss of a bipolar field cycle is used to deduce the filament diameter, this, and hence the higher field loss, will be overestimated, because of the additional contribution of the Nb below 0.18 T. The range of the operating conditions should dictate the decision on the criterion to be used for the filament diameter. If necessary, a correction factor can be added in the calculation code to include the Nb magnetization contribution at the zero-crossing field.

### Hysteresis Loss with DC Transport Current

When a longitudinal current is superimposed on the transverse field magnetization currents of a filament, the electrical center line is moved to the periphery of the filament (or slab), the flux profiles are asymmetric, and the penetration field decreases by a factor  $1 - i$ , where  $i$  is the ratio of the longitudinal current (also called the *transport* current) to the critical current

$$i = \frac{I_{tr}}{I_{c1}(b)}. \quad B_{p\perp}^i = B_{p\perp}(1 - i) \quad (8)$$

Below penetration (i.e., for  $\Delta B < 2B_{p\perp}^i$ ), the low  $\Delta B$  amplitude formula in Table 1 can also be used in the presence of a dc transport current. Above penetration, the magnetization decreases as a function of the transport current, dropping to 0 for  $i = 1$ , that is,  $I_{tr} = I_c$ . For large  $\Delta B$ , the area of the magnetization loop (i.e., the energy supplied by the external field change) decreases when a transport current is superimposed on the magnetization currents [see Fig. 5(b) from Ref. 25]

$$\text{for } \Delta B \gg 2B_{p\perp}^i, \quad Q_{\perp}^i \approx Q_{\perp}^i \approx Q_{\perp}(1 - i^2) \text{ (J/m}^3\text{)} \quad (9)$$

As a function of the transport current, the magnetization loss increases until full penetration is achieved for  $\Delta B = 2B_{p\perp}(1 - i)$ . A further increase of the transport current decreases the magnetization loss; see Figs. 5(a) and 6 from Ref. 25.

A change of the filament magnetization beyond  $2B_{p\perp}^i$  is opposed by the power supply, which works to maintain the transport current, that is, the asymmetric flux profiles. A voltage appears along the filament, and an extra energy  $Q_d$  due to the dynamic resistance  $R_d$  must be added to the magnetization loss (26, 27). Above penetration, the dynamic resistance is proportional to the amplitude of the field change and inversely proportional to the duration  $t_0$  of the field change

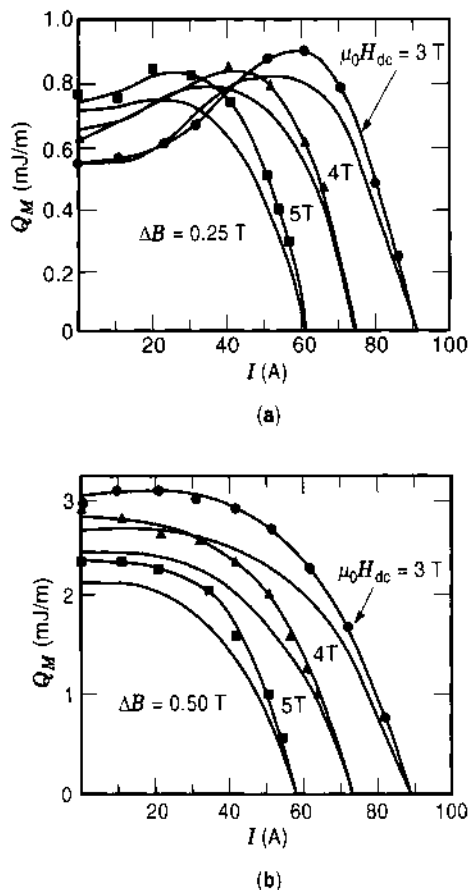
$$\text{at } \Delta B < 2B_{p\perp}^i, \quad R_d = 0 \quad (10)$$

$$\text{at } \Delta B > 2B_{p\perp}^i, \quad R_d \propto \frac{\Delta B - 2B_{p\perp}^i}{t_0} \quad (11)$$

For  $\Delta B \gg 2B_{p\perp}^i$ ,  $R_d \propto \Delta B$ , that is, the energy loss per cycle,  $Q_d$ , is proportional to  $i^2$  but is independent of the field rate. The total loss in filaments carrying a dc transport current is

$$\begin{aligned} \text{for } \Delta B \gg 2B_{p\perp}^i, \quad Q^{\text{total}} &= Q_d + Q_{\perp}^i \\ &= Q_{\perp}(1 + i^2) \text{ (J/m}^3\text{)} \end{aligned} \quad (12)$$

From Eq. (12), at  $\Delta B \gg 2B_{p\perp}^i$  the ratio of the total loss with transport current to the magnetization loss is  $\leq 2$  for any transport current. For  $B_{p\perp} \geq \Delta B/2 > B_{p\perp}^i$ , the loss enhancement factor can be much larger than a factor of two (10). This can be understood by remembering that the low  $\Delta B$  magnetization loss is proportional to  $\Delta B^3$ , but the dynamic resistance loss  $Q_d$  is proportional to  $\Delta B \cdot i^2$ .



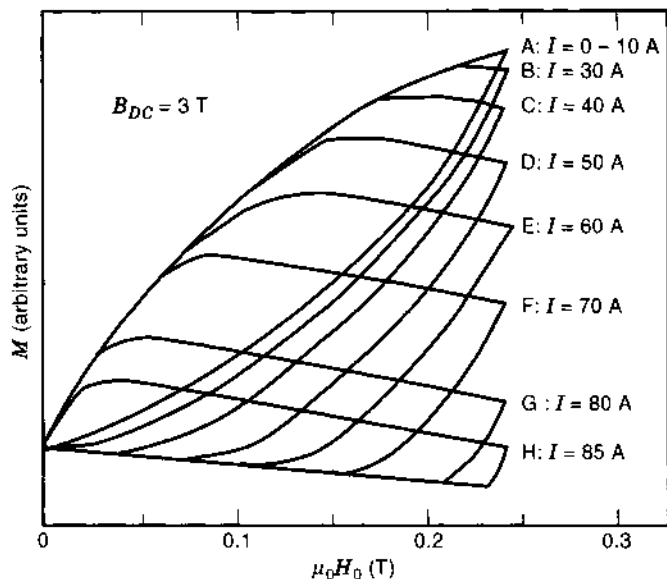
**Figure 5.** Magnetization loss as a function of the dc transport current for a single core NbTi conductor: (a)  $\Delta B = 0.25 \text{ T} < 2B_{p1}^{i=0}$ , (b)  $\Delta B = 0.50 \text{ T} > 2B_{p1}^{i=0}$ . From Ref. 25. Reprinted from *Cryogenics*, 25, T. Ogasawara, Y. Takahashi, K. Kanbara, Y. Kubota, K. Yasohama, and K. Yasukochi, Alternating field losses in superconducting wires carrying dc transport currents: Part 1. Single core conductors, pp. 736–740, copyright 1979, with permission from Elsevier Science.

The transport current also affects the azimuthal magnetization currents, modifying the local field orientation angle. Some kind of dynamic resistance is also expected because of the interference of  $I_{tr}$  with  $J_{c\phi}$ .

Equations (9) and (12) give a satisfactory and validated (25) model for a single-core strand. However, two assumptions are required to extend them to a multifilamentary strand or to a cable of stranded wires:

- Each filament of each strand carries the same fraction of critical current.
- The longitudinal current is constant during the external field change.

Both assumptions are highly unlikely: On one hand the current distribution is not homogeneous across either the cable or the filaments of an individual strand. On the other hand, the several coupling current loops induced by a field change cause local, time-dependent, very large variations of the current density compared with the average value.



**Figure 6.** Magnetization loops at increasing transport current for a single core NbTi conductor ( $D = 250 \mu\text{m}$ ), with  $\Delta B = 0.25 \text{ T}$ , from Ref. 25. The magnetization loss is maximum at F, when  $2B_{p1}^{i=0} = 0.25 \text{ T}$ ; see also Fig. 5(a). Reprinted from *Single core conductors*, pp. 736–740, copyright 1979, with permission from Elsevier Science.

**Self-Field Loss**

A special case of hysteresis loss occurs for a straight, solid conductor carrying a longitudinal current in the absence of any external field. The flux penetration is due only to the self-field, which is proportional to the longitudinal current. Full penetration and maximum loss occur for  $I = I_c$ . For ac operation, it is convenient to write the hysteresis loss per unit length per cycle as a function of  $i$ , the ratio of the peak current to the critical current. Norris (28) proposed a formula for self-field loss in an isolated thin slab and a round (or elliptical) filament

For round or elliptical cross section

$$q_r = \frac{l_c^2 \mu_0}{\pi} \left( (1-i) \ln(1-i) + \frac{(2-i)i}{2} \right) \quad (\text{J/m}) \quad (13)$$

For a thin slab

$$q_s = \frac{l_c^2 \mu_0}{\pi} [(1-i) \ln(1-i) + (1+i) \ln(1+i) - i^2] \quad (\text{J/m}) \quad (14)$$

For conductors with the same critical current, the loss ratio at saturation (i.e., at  $i = 1$ ) is  $Q_r/Q_s = 1.3$ . At small fractions of the critical current,  $Q_r/Q_s = 1/i$ , showing that the advantage of the thin slab geometry is significant only at very small current density.

In a round multifilamentary composite, the filaments are not transposed for self-field, and the filamentary zone of diameter  $D_{fz}$  can be treated as a single core, applying a filling factor  $\lambda$  for the critical current. Wilson (10) discusses the self-field loss for a round multifilamentary composite in terms of penetration field. The complete penetration field is  $B_{ps} = \lambda J_c D_{fz}/2$ , and the partial penetration field is  $B_{ms} = i B_{ps}$ . The loss per cycle per

unit filament volume is

$$Q = \frac{E_{ms}^2}{2\mu_0} \left( \frac{2}{i} - 1 + \frac{2(1-i)}{i^2} \ln(1-i) \right) \quad (\text{J/m}^3) \quad (15)$$

Equation (15) can be written in terms of critical current, as a loss per unit length, and becomes

$$q = \frac{\lambda I_c^2 \mu_0}{\pi} \left( (1-i) \ln(1-i) + \frac{(2-i)i}{2} \right) \quad (\text{J/m}) \quad (15a)$$

which is identical to Eq. (13) except for the filling factor  $\lambda$ . In the case of an oscillating, unidirectional current,  $i$  is defined as the ratio of transport to critical current, and the loss formula proposed by Wilson (10) as loss per unit length, becomes,

$$q = \frac{\lambda I_c^2 \mu_0}{\pi} \left[ 2(2-i) \ln \left( \frac{2-i}{2} \right) + \frac{(4-i)i}{2} \right] \quad (\text{J/m}) \quad (16)$$

The use of thin filaments does not help to reduce the self-field loss, as the nontransposed filamentary zone behaves like a single core with critical current reduced by the filling factor  $\lambda$ . When the self-field loss becomes a crucial issue, it is recommended to select a transposed cable or braid, where the filamentary zone of the individual strands is kept as small as possible.

The preceding formulas have become very popular in the high temperature superconductor community, although both Norris and Wilson warned about the limits of their applicability. The assumption of constant critical current may result in significant errors at low field. The twist in the multifilamentary composites introduces a spiral component of the self-field. For cabled conductors, the field from the neighboring strands may give rise to coupling current loss, not taken into account in the preceding formulas.

#### Accuracy of Hysteresis Loss Estimation

The overall accuracy of the hysteresis loss estimation is affected by the simplifications assumed in the model, the accuracy of the conductor parameters, the local field orientation, the distribution of the transport current (29, 30), and the nonfilamentary magnetization at low field. The weight of the individual error sources depends on the conductor layout and operating conditions. The use of sophisticated computer codes does not help much to improve the accuracy of the hysteresis loss prediction, which lies, in the best cases, around 20%.

**Model Accuracy.** The geometrical basis for hysteresis formulas is either a cylinder or an infinite slab. The actual filament geometry, especially in the case of bridging, is not a cylinder: even when an equivalent diameter is defined, it does not perfectly model the real filament over the entire range of operating conditions.

Bean's assumption of linear flux profile may be a source of inaccurate estimation of the magnetization and penetration field at low magnetic field, especially for thick filaments, strands, and tapes. In addition, the model does not account for surface screening currents below  $B_{c1}$ .

The loss formulas may include the effect of the strain and temperature on  $J_c$  and  $B_p$ , but the integration of the magnetization is done under isothermal conditions. A step-by-step integration, calculating the magnetization from the instantaneous value of  $J_c(T, \epsilon)$ , also would not be correct: an increase of  $T$  or  $|\epsilon|$  at constant field decreases the magnetization, as  $J_c$  decreases and more flux penetrates. However, a decrease in  $T$  or  $|\epsilon|$  at constant field leaves the flux profiles, and hence the magnetization, unchanged.

#### COUPLING CURRENT LOSS

Two filaments in a strand, as well as two noninsulated strands in a cabled conductor, constitute a loop for induced currents under a perpendicular time-varying field, that is, they are *coupled* in a current loop. A large portion of the loop is superconducting, that is, the linked area is large but the loop resistance is small. To reduce the linked area, the filament bundle is twisted and the strands are cabled with tight pitches, leading to transposition with respect to the perpendicular field (31).

The magnetic energy, initially stored in the coupling current loop, is released as Joule heating by the resistive decay of the induced currents, with a time constant,  $\tau$ , which is the ratio of the loop inductance to the loop resistance. In a round, ideal multifilamentary strand, with the filaments homogeneously distributed over the cross section, the time constant for the interfilament coupling currents is a function of the twist pitch  $l_t$  and the transverse resistivity  $\rho$

$$\tau = \frac{\mu_0 l_t^2}{8\pi^2 \rho} \quad (\text{s}) \quad (17)$$

The transverse resistivity  $\rho$  in a multifilamentary composite is a function of the bulk resistivity  $\rho_m$ , of the matrix and the superconducting fraction in the filamentary zone  $\lambda$ . If no resistance barrier is found at the interface between filaments and matrix, according to Carr (32) the transverse resistance is

$$\rho = \rho_m \frac{1-\lambda}{1+\lambda} \quad (\Omega \cdot \text{m}) \quad (18)$$

When a high resistivity barrier builds up around the filaments (e.g., in hot extruded NbTi composites), the transverse resistance is assumed to be

$$\rho = \rho_m \frac{1+\lambda}{1-\lambda} \quad (\Omega \cdot \text{m}) \quad (19)$$

In a cable of noninsulated strands, the interstrand coupling currents add to the interfilament loops. In a multistage cable, a large variety of coupling current loops exists, each with an individual time constant  $\tau_i$ . The size of the loops depends on the length and sequence of the cable pitches, but the exact path of the coupling currents, and hence the transverse resistance, is hard to predict (33). As a general trend, the transverse resistance is larger and the loss is smaller when the pitches of the different cable stages all have the same direction (34) and their ratio is close to one, that is, short pitches are used for the higher cable stages and long pitches for the lower cable stages.

In most cases, the loss is not homogeneously distributed over the strand or cable volume: a dimensionless geometry factor,



$n_i$ , is associated with each current loop with time constant  $\tau_i$ . The geometry factor (35–37) allows for the demagnetization effects (e.g., round versus flat conductor) and normalizes the loss to the overall strand volume (e.g., when the filament bundle is surrounded by a large normal metal shell) or to the cable volume (e.g., when an interstrand current loop is restricted to a fraction of the cable volume).

### Steady State Coupling Loss Formulas

When the time scale of a field change (e.g., the duration of a linear ramp or the period of a field oscillation) is much larger than any of the conductor time constants, steady state conditions are established for the coupling currents. For linear field change, with constant  $dB/dt$ , the power loss (10, 35, 38) per unit volume of strand material is

$$P = \frac{n\tau}{\mu_0} B^2 \quad (\text{W/m}^3) \quad (20)$$

For sinusoidal field variations  $B = (\Delta B/2) \sin \omega t$ , with frequency  $\pi$  and  $\omega = 2\pi$ , the average power loss (10, 35, 39) is

$$P = \frac{n\tau}{8\mu_0} \Delta B^2 \omega^2 \quad (\text{W/m}^3) \quad (21)$$

The energy loss  $Q_c$  for a field cycle of amplitude  $\Delta B$  is respectively

$$Q_c = \frac{2n\tau}{\mu_0} B \Delta B \quad (\text{J/cycle} \cdot \text{m}^3),$$

$$Q_c = \frac{\pi n\tau}{4\mu_0} \Delta B^2 \omega \quad (\text{J/cycle} \cdot \text{m}^3) \quad (22)$$

For cabled conductors with multiple current loops and associated time constants, the  $n\tau$  in the preceding formulas is the sum of the individual terms

$$n\tau = \sum_i n_i \tau_i \quad (\text{s}) \quad (23)$$

In steady-state conditions, where all the current loops are fully activated, it is not necessary to know the breakdown of  $n\tau$  into individual components. The average coupling loss can be calculated from the overall  $n\tau$ , obtained, for example, from measurements on a short conductor section.

The tool for experimental assessment of the coupling currents loss is the *loss curve*, where the energy per cycle per unit volume of strand (or cable) is plotted as a function of the field rate, for a linear ramp, or the frequency, for a sinusoidal field sweep. The hysteresis loss is the extrapolation of the loss curve to  $dB/dt = 0$ . From the initial slope of the loss curve,  $n\tau$  is derived using the steady-state formulas, Eq. (22).

### Transient Coupling Loss Formulas

The energy loss per unit volume in Eq. (22) is linear in the field rate or frequency. However, the energy loss has an obvious upper limit set by the magnetic field energy density,  $\Delta B^2/2\mu_0$ . At  $\tau > 0.3$  and at ramp time  $t_0 < 10\tau$ , Eqs. (20–22) give a loss overestimation larger than 10% and should be replaced by transient field loss formulas.

For multifilamentary strands in an oscillating field, the currents flowing in the outer filament layers screen the inner volume of the conductor. In a fast ramped field, the field penetrates the innermost layers with the time scale of the decaying screening currents  $\tau$ , even if the duration of the applied field change is smaller. In fully transposed cables, the mechanism of screening depends on the interstrand current loops and is hardly predictable.

For conductors characterized by a single time constant  $\tau$ , the transient formulas for sinusoidal oscillations (35), linear ramp (10), and exponential decay (40) are respectively for  $B = (\Delta B/2) \sin t$ ,

$$\bar{P} = \frac{n\tau \Delta B^2 \omega^2}{8\mu_0(1 + \omega^2 \tau^2)} \quad (\text{W/m}^3)$$

$$Q_c = \frac{\pi n\tau \Delta B^2 \omega}{4\mu_0(1 + \omega^2 \tau^2)} \quad (\text{J/cycle} \cdot \text{m}^3) \quad (24)$$

for  $B = \Delta B/t_0$

$$Q_c = \frac{\Delta B^2 n\tau}{\mu_0 t_0} \left[ 1 - \frac{\tau}{t_0} (1 - e^{-t_0/\tau}) \right] \quad (\text{J/m}^3) \quad (25)$$

for  $B = \Delta B(1 - e^{-t/t_0})$

$$Q_c = \frac{\Delta B^2 \cdot n\tau}{2\mu_0(t_0 + \tau)} \quad (\text{J/m}^3) \quad (26)$$

In case of multiple time constants, Eqs. (24–26) cannot be applied using the  $n\tau$  defined in Eq. (23). If each current loop behaved independently (i.e., the screening currents of the largest loops did not affect the applied field at the other loops), the total transient loss would be the sum of the individual contributions. From Eqs. (24) and (25)

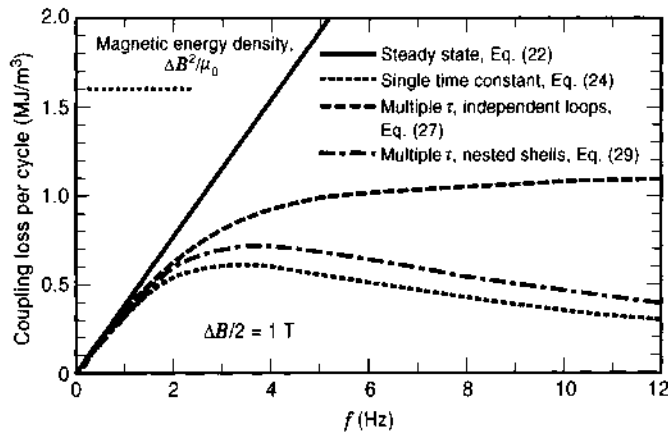
$$Q_c = \sum_i Q_c^i = \frac{\pi \omega \Delta B^2}{4\mu_0} \sum_i \frac{n_i \tau_i}{1 + \omega^2 \tau_i^2} \quad (\text{J/cycle} \cdot \text{m}^3) \quad (27)$$

$$Q_c = \sum_i Q_c^i = \frac{\Delta B^2}{\mu_0 t_0} \sum_i n_i \tau_i \left[ 1 - \frac{\tau_i}{t_0} (1 - e^{t_0/\tau_i}) \right] \quad (\text{J/cycle} \cdot \text{m}^3) \quad (28)$$

On the assumption that the larger current loops screen the smaller current loops, which is what happens in nontransposed conductors, a formula for the transient loss in a sinusoidal field has been proposed in Ref. 37 for a conductor with  $N$  loops

$$Q_c = \frac{\pi}{4\mu_0} \Delta B^2 \left( \frac{\omega n_N \tau_N}{1 + \omega^2 \tau_N^2} + \sum_{k=1}^{N-1} \frac{\omega n_k \tau_k}{1 + \omega^2 \tau_k^2} \prod_{i=k+1}^N \frac{1}{1 + n_i \omega^2 \tau_i^2} \right) \quad (\text{J/cycle} \cdot \text{m}^3) \quad (29)$$

It is hard to reliably predict the transient coupling loss in a large multistage cable. On one hand, the overall  $n\tau$  should be broken down into the individual  $n_i \tau_i$ ; the procedure to add the  $n\tau$ 's measured separately on the lower cable stages is not satisfactory, as the transverse resistance and the current loops change when the subcables are bundled together. On the other hand, because the current path for each loop is not known, it



**Figure 7.** Example of prediction of coupling loss in oscillating field, for a conductor with multiple time constant,  $n\tau = n_1\tau_1 + n_2\tau_2 + n_3\tau_3 = 4 + 16 + 80 = 100$  ms. The plot shows the result obtained neglecting the shielding effect, Eq. (22), and with three different approaches to taking account of shielding.

is hard to decide to what extent the higher loops do screen the smaller loops, that is, to decide between Eqs. (28) and (29).

In conductors with multiple time constants, the largest underestimation of the transient field loss occurs when the Eqs. (24–26) for a single time constant are applied. The steady state formulas, Eqs. (20–22), give the largest overestimation. Whenever the breakdown of the overall  $n\tau$  is known, Eqs. (27–28) give a better, but still conservative, estimate, because they assume no screening. On the contrary, Eq. (29) is rather optimistic, because it treats the current loops as nested shells. Equation (29) has been used to find the  $n\tau$  components from the experimental loss curves in Refs. 37 and 41.

An example of coupling loss prediction for a conductor with multiple time constants is shown in Fig. 7. The steady-state, overall time constant is assumed to be  $n\tau = 100$  ms (e.g., drawn from an experimental loss curve). The breakdown of the time constant is assumed to be  $\tau_1 = 2$  ms,  $\tau_2 = 8$  ms,  $\tau_3 = 40$  ms, and  $n_1 = n_2 = n_3 = 2$  (round cross section). The plot in Fig. 7 shows the loss according to Eqs. (22), (24), (27), and (29), for an applied field oscillation  $B = (\Delta B/2)/\sin t$ , with  $\Delta B = 2$  T. At low frequencies (i.e., for  $\tau < 0.2$ ), all the formulas give the same result, but at higher frequencies the assumptions about the current paths and the multiple screening lead to substantially different results.

### Saturation of Coupling Currents

According to the coupling loss formulas, the energy dissipation occurs in the resistive section of the induced current loops, that is, in the composite matrix for interfilament coupling loss and at the strand-to-strand contacts for the interstrand coupling loss. At a first approximation level, the coupling currents are assumed not to change the filament magnetization. Actually, the coupling currents flowing in the outer filament layers create in steady state a field difference  $\tau B$  across the multifilamentary zone. The associated magnetization loss is referred to as *penetration loss* and can be treated in analogy to the hysteresis loss of a solid filament of the size of the filamentary zone,  $D_{fz}$ , with a critical current  $\lambda J_c$ , where  $\lambda$  is the superconductor fraction in the filamentary zone (39). In steady state, the penetration loss per cycle  $Q_p$ , normalized to the volume of the filamentary zone,

can be easily found by replacing  $B_{p,l}$  by  $\tau B$  in the formulas of Table 1:

$$Q_\mu = \frac{4\tau\Delta B\dot{B}}{3\mu_0} \quad (\text{J/cycle} \cdot \text{m}^3) \quad (30)$$

The total loss in a round multifilamentary composite with  $n = 2$  is the sum of the coupling loss, according to Eq. (20), and the penetration loss

$$Q_{\text{tot}} = Q_c + Q_p = \frac{4\tau\Delta B\dot{B}}{\mu_0} \\ \frac{4\tau\Delta B\dot{B}}{3\mu_0} = \frac{16\tau\Delta B\dot{B}}{3\mu_0} \quad (\text{J/cycle} \cdot \text{m}^3) \quad (31)$$

Whenever the loss is experimentally assessed, the penetration loss does not need to be added to the coupling loss, because it is already buried in the  $n\tau$  inferred from the loss curve. The penetration loss for an oscillating field and for a transient field are discussed in Refs. 39 and 10.

At high field rates, the coupling currents may reach the critical current. The outer filament layer is saturated, and the difference between outer and inner fields is the penetration field for the filamentary zone,  $\tau\dot{B} = B_p^{\dot{B}} = \mu_0\lambda J_c D_{fz}/\pi$ . Saturation in a multifilamentary composite occurs whenever

$$\frac{\pi\tau_{\text{str}}\dot{B}}{\mu_0\lambda J_c D_{fz}} \geq 1 \quad (32)$$

The saturation loss is the upper limit of the penetration loss and is obtained by substituting  $B_p^{\dot{B}}$  for  $\tau\dot{B}$  in Eq. (30):

$$Q_{\text{sat}} = \frac{4\Delta B}{3\pi} D_{fz}\lambda J_c \quad (\text{J/m}^3) \quad (33)$$

When the condition of Eq. (32) is fulfilled (i.e., when the current loops are saturated), the coupling loss does not increase for higher field rates. The maximum total loss in a multifilamentary composite for a long duration (steady-state) field change is independent of  $\tau$  and can be written by substituting Eqs. (32) and (33) into Eq. (31):

$$Q_{\text{tot}}^{\text{max}} = Q_c + Q_{\text{sat}} = \frac{4\tau\Delta B}{\mu_0} \frac{\mu_0\lambda J_c D_{fz}}{\pi\tau} + \frac{4\Delta B}{3\pi} D_{fz}\lambda J_c \\ = \frac{16\Delta B\lambda J_c D_{fz}}{3\pi} \quad (\text{J/cycle} \cdot \text{m}^3) \quad (34)$$

### Coupling Loss with Transport Current

As long as coupling currents and transport current use a small fraction of the superconducting cross section, the influence of the transport current is limited to the hysteresis loss change. At a higher field rate or higher  $i$  (ratio of transport current to critical current), all the superconducting cross section is eventually engaged to carry either the transport or the coupling currents. The criterion for saturation with transport current in Eq. (32) becomes

$$\frac{\pi\tau_{\text{str}}\dot{B}}{\mu_0\lambda J_c D_{fz}} \geq 1 - i \quad (35)$$

The larger  $i$  is the smaller the loop current (and hence the field rate) is to achieve saturation. Above saturation (i.e., when  $2I_{\text{loop}} + I_{\text{transport}} > I_c$ ), the excess of transport current must be accommodated in the superconducting cross section carrying  $-I_{\text{loop}}$ . The paths of the coupling currents, with the current direction reversing with the periodicity of the transposition pitch, force the transport current (or a fraction of it) to switch continuously from one to the other filament (interfilament coupling) or strand (interstrand coupling) to match  $-I_{\text{loop}}$ . The energy dissipated is at the expense of the power supply, and it is called the *dynamic resistance loss* (29), because of the analogy between hysteresis and coupling loss due to transport current (see Ref. 42 for a discussion of the limits of this analogy). Whenever a transport current is imposed, the magnetic energy density of the applied field cannot be considered as an upper limit for the overall loss.

Above saturation, the coupling currents (and coupling loss) decrease and the dynamic resistance loss sharply increases. The behavior of the total loss as a function of  $i$  over the full range of  $B$  has been calculated analytically for a slab; see Fig. 8 (from Ref. 29). A cylinder requires a numerical calculation for the saturation range (43), leading to a results similar to Fig. 8. Experimental results on interfilament loss with transport current (29, 43) confirm the behavior of Fig. 8. At a very large field rate (i.e., when saturation occurs even at  $i = 0$ ), the effect of the transport current on the overall loss is an increase by a factor  $1 + i^2$

$$Q_{\text{sat}}(i) = (1 + i^2)Q_{\text{sat}}(0) \quad (\text{J/m}^3) \quad (36)$$

At an intermediate field rate (i.e., when saturation is achieved only above a certain value of transport current)

$$1 > \frac{\pi \tau_{\text{str}} B}{\mu_0 \lambda J_c D_{\text{tr}}} > 1 - i \quad (37)$$

the behavior of the total loss as a function of  $i$  is complex [see (29, 42, 43)]. Using  $Q_c(i) = (1 + i^2)Q_c(0)$  is not recommended and may lead to large errors in the actual loss values.

In large cable-in-conduit conductors, the occurrence of saturation cannot be exactly predicted from Eq. (35). If the current loops (i.e., the coupling currents) are not homogeneously distributed, a redistribution of the transport current at the start of the field change may avoid the occurrence of saturation and dynamic resistance loss in steady state. Using Eq. (35) with the average  $i$  and the overall cable time constant is very conservative and may result in an underestimate of the field rate causing saturation. In cable in conduit with multiple time constants, where only the overall  $n\tau$  is known, it may be difficult to select the correct  $\tau$  to be used in Eq. (35). An example of saturation of coupling currents in a large cable-in-conduit conductor with nonhomogeneous current distribution is reported in Ref. 44.

### Coupling Loss in Flat Cables

In flat cables and in rectangular composites with sides  $a$  and  $b$ , both  $n$  and  $\tau$  are much larger for field perpendicular to the broad side  $a$  of the cable. The loss ratio for the two orthogonal orientations has been calculated as a function of the aspect ratio  $\alpha = a/b$  for conductors with homogeneous transverse resistivity. The results obtained by Murphy et al. with concen-

tric ellipses (45), Turck et al. with concentric rectangles (46), and Campbell with rounded-edge concentric layers (35) are respectively

$$\frac{Q_c^{\perp a}}{Q_c^{\perp b}} = \alpha^4, \quad \frac{Q_c^{\perp a}}{Q_c^{\perp b}} = \alpha^2 \frac{35 + 20\alpha + 7\alpha^2}{35\alpha^2 + 20\alpha + 7}, \quad \frac{Q_c^{\perp a}}{Q_c^{\perp b}} = \frac{\alpha^4}{4} \quad (38)$$

In practical large flat cables, an insulating (or high-resistivity) strip is placed in the midplane to increase the transverse resistance of the coupling currents for field orientation perpendicular to  $a$ ; see for example Ref. 47. In such cases, the equations (38) are no longer valid and the loss anisotropy is reduced (48).

At intermediate angles, an analytical loss formula is proposed in Ref. 45, modeling the flat cable as an ellipse with homogenous transverse resistivity. In practical, nonhomogeneous flat cables, the coupling loss for field orientation at an angle  $\theta$  with respect to the broad side  $a$  can be roughly estimated by splitting the field into the orthogonal components and adding the loss contributions

$$Q_c(\theta) = Q_c^{\perp a} \sin^2 \theta + Q_c^{\perp b} \cos^2 \theta \quad (\text{J/m}^3) \quad (39)$$

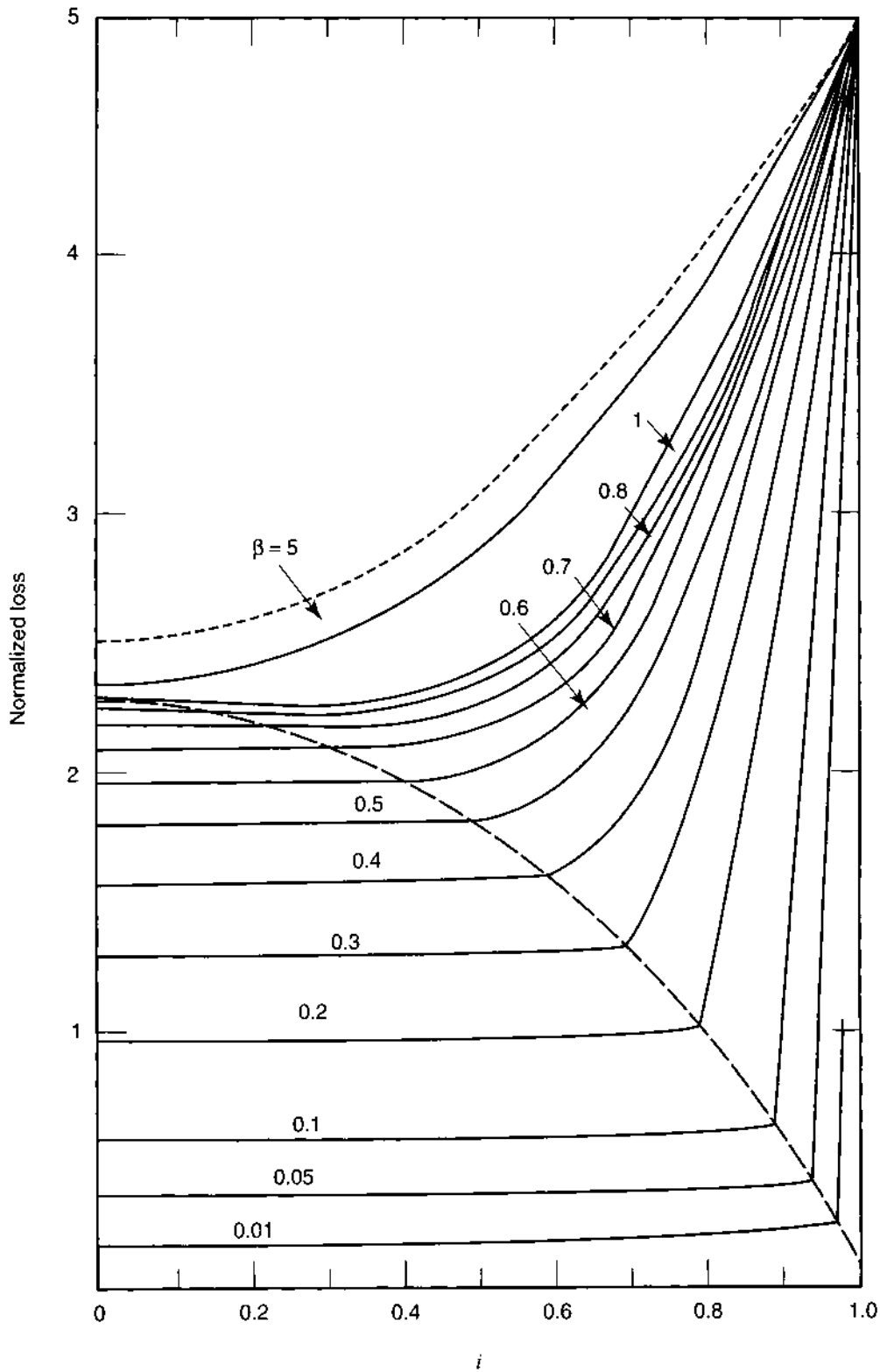
For flat cables with a large aspect ratio, the second term in Eq. (39) can be neglected over a broad range of angles.

### Coupling Loss in Spatially Changing Magnetic Field

When a superconducting cable is exposed to a time-varying magnetic field that is not homogeneous along the conductor, the periodicity of the boundary conditions for the coupling currents loops is affected. If the flux linked by two geometrically identical current loops next to each other is not balanced, the coupling currents extend beyond the boundary of the pitch length. In one-stage cables (e.g., one-layer flat cables and Rutherford cables), adjacent strand pairs may build current loops with different flux balance in the presence of a spatial gradient of magnetic field. Because of the different boundary conditions, the individual strand pairs carry coupling currents of different amplitude. The inductance associated with these *extended* current loops is larger and the resistance is smaller, resulting in time constants that may be orders of magnitude larger than with strictly periodic boundary conditions. The result is a strong, quasi-steady-state current imbalance and larger loss.

Ries and Takács (49, 50) first did an analysis of the coupling loss in a spatially changing magnetic field for flat cables. The subject, later named *boundary-induced coupling currents*, (BICCs), assumed a much larger relevance in the context of Rutherford cables for accelerator magnets. Here, the spatial field gradients along the conductor at the saddles of the dipole magnets are large and occur over a length smaller than the cable pitch. The long-lasting current imbalance across the cable leads to field distortions (51, 52) and ramp rate limitations in the accelerator dipoles (53, 54).

The variation of the strand crossover resistance along Rutherford cables has also been shown to be a potential reason for BICCs (53). In multistage cable-in-conduit conductors, the current loops do not have a regular pattern, as a result of the nonhomogeneous distribution of the interstrand resistance. Flux imbalance for current loops next to each other is expected to be frequent in large cable-in-conduit conductors, even in a spatially homogeneous magnetic field.



**Figure 8.** Normalized loss,  $Q/(B_p^{t_0})$ , in a slab geometry as a function of the transport current fraction  $i$  for different  $\beta = \tau B/B_p^{t_0}$  (29). The locus of the knees of the curves shows the saturation threshold. Above saturation, the loss is dominated by the dynamic resistance. Reprinted from *Cryogenics* 20, T. Osagawara, Y. Takahashi, K. Kanbara, Y. Kubota, K. Yasohama, and K. Yasukochi, Transient field losses in multifilamentary composite conductors carrying dc transport currents, pp. 216–222, copyright 1980, with permission from Elsevier Science.

### Interstrand Resistance in Cable-in-Conduit Conductors

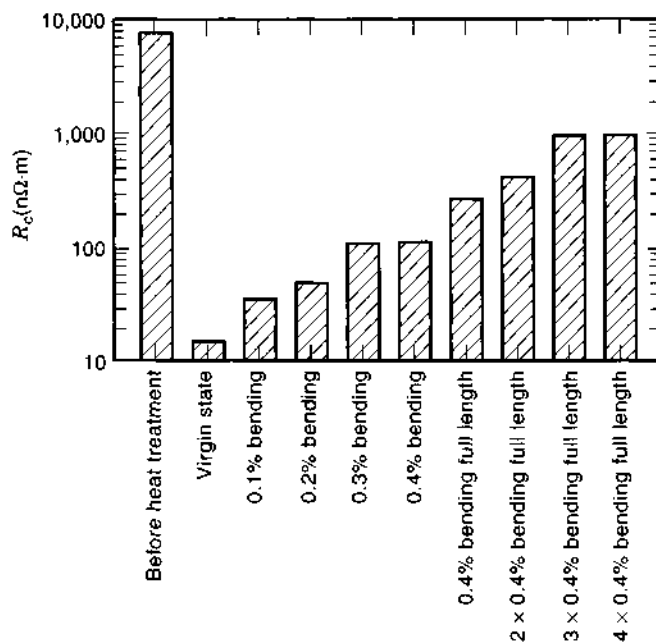
Cable-in-conduit conductors (CICCs) are a special case of multi-stage cables. What makes the CICCs different is the tribological nature of the transverse resistance, which is only marginally determined by the bulk properties of the metallic components. A database for coupling currents loss in CICCs and its implication for stability is discussed in Ref. 55.

In CICCs with void fraction in the range of 30% to 40%, the coupling currents may follow complex, zigzag paths through a number of good electrical contacts at the strand crossovers. Rather than the 2-D smeared transversal resistivity, the interstrand resistance is the critical parameter, together with the pitch length, for assessing and controlling the coupling loss in CICCs. The interstrand resistance in CICCs has units of ohm-meters and is measured as the dc resistance between a strand pair embedded in a cable, multiplied by the length of the cable section, which should be longer than a pitch length. For the same conductor, the interstrand resistance in a CICC may vary as much as 20% from piece to piece and as a function of the sample length and homologous strand pair (56).

**NbTi Strands.** The interstrand resistance depends on the operating transverse load and on the conductor history, including contact surface oxidation, heat treatment, and cycling. In CICCs made of bare NbTi strands, the interstrand resistance, as well as the coupling loss, may have a broad range of results (57): the thin layer of copper oxide that develops at room temperature at the strand surface provides a precarious resistive barrier, which may either partly dissolve or break under heat treatment and/or applied load (58–60). The electromagnetic load in operation can also produce a dramatic increase in the coupling loss in CICCs made of bare NbTi strands (61, 62). Whenever long term, reliable control of the coupling loss is desirable, it is recommended to use a surface coating for the NbTi strands.

Several coatings have been investigated on NbTi strands for contact resistance, mostly in the frame of the Rutherford cable development for accelerator magnets, including Zn, SnAg, Ni, and Cr. The SnAg (Stabrite) soft coatings produce very low contact resistance with applied load (59, 63) and are not recommended for low-loss cables. A resistive barrier can be obtained by a diffusion heat treatment at 200°C of the SnAg-coated strand before cabling, building a Sn-rich bronze shell at the surface; however, the diffusion heat treatment also affects the residual resistance ratio (RRR) of a fraction of the stabilizer. The Ni coating, as well as the Zn, has higher contact resistance and is not sensitive to curing heat treatment and applied load (63). The Cr coating has proved to be effective in cutting the interstrand loss in medium-sized CICCs (64, 65). A direct comparison of contact resistance for bare and Cr- and Ni-plated strands indicates almost one order of magnitude higher resistance for Cr than for Ni and another order of magnitude difference between Ni and clean Cu (57).

**Nb<sub>3</sub>Sn Strands.** In CICCs of bare Nb<sub>3</sub>Sn strands, diffusion bonding (sintering) occurs at a number of strand crossovers during heat treatment, resulting in low interstrand resistance and high coupling loss (66). The Cr coating has been identified as a reliable, thin coating to avoid sintering during the heat treatment, with moderate effect on the strand RRR.



**Figure 9.** Evolution of the interstrand resistance in a Cr-plated Nb<sub>3</sub>Sn CICC (81 strands, 30%), before and after the heat treatment and after controlled bending strain.

The effect of the Cr plating on the coupling loss of Nb<sub>3</sub>Sn-based CICCs has been the object of systematic investigations. A comparison of CICC samples identical except for Cr plating by vendors (67) has found a variation as high as a factor of four for the interstrand resistance (and the loss), depending solely on the proprietary electroplating processes. The influence of the Cr plating process parameters on the contact resistance is discussed in Ref. 68.

The void fraction is an important parameter affecting the interstrand resistance and coupling loss in CICCs, with a low loss range above 40% voids and a large loss range below 30% (69). The looseness of the cable in the jacket may play an even more important role than the average void fraction: the electromagnetic forces associated with the coupling currents tend to pull apart the strand bundle and relax the contacts at the strand crossovers.

The most striking effect observed in the interstrand resistance of Cr-plated Nb<sub>3</sub>Sn CICCs is the decrease in resistance after heat treatment and the increase after mechanical (70) and/or electromagnetic loading (44, 71). The diagram in Fig. 9 shows the evolution of the interstrand resistance at different steps: the large drop after the heat treatment may be due to a partial depletion of O from the Cr oxide at the sealed surface of the strand crossovers. The subsequent increase of the interstrand resistance after bending loads (as happens in the react and transfer coil manufacturing process) or transverse loads (Lorentz forces in operation) can be understood in terms of microscopic movements of the crossovers, which partly lose their initial engagement. An important lesson learned from these results is that most loss test results obtained for short samples in the virgin state (i.e., as heat-treated, without any load) overestimate the loss by about one order of magnitude compared with the actual coil operating conditions.

**Target Interstrand Resistance for CICCs.** The goal of the designer is to obtain in operation a value of interstrand resistance

large enough to reduce the coupling loss to an acceptable level but small enough to allow fast and effective current redistribution among the strands and provide stability under transient local disturbances.

The experimental results on the interstrand resistance  $R_c$  and the coupling current constant  $n\tau$ , measured on the same  $Nb_3Sn$  CICC specimens, have been correlated in Ref. 55 to obtain

$$R_c \approx \mu_0 \frac{l_t^2}{n\tau} \quad (\Omega \cdot m) \quad (40)$$

Once the acceptable coupling loss is assessed and the cable pitch  $l_t$  is known, Eq. (40) indicates the *minimum* design value for the interstrand resistance, that is, for the kind of strand coating to be selected. The *maximum* allowable value for the interstrand resistance is assessed by the requirement on the interstrand current sharing length,  $l_0$ . According to Ries (72)

$$l_0 = \sqrt{R_c/R_{||}} \quad (m) \quad (41)$$

where  $R$  ( $\Omega/m$ ) is the normal longitudinal resistance of the strand per unit length. Combining Eqs. (40) and (41), a criterion for interstrand resistance in CICC summarizing the coupling loss and stability requirements is obtained according to Ref. 55

$$\frac{\mu_0 l_t^2}{n\tau} < R_c < l_0^2 R_{||} \quad (\Omega \cdot m) \quad (42)$$

In very large CICC, Eq. (42) may need a correction if the interstrand resistance for the first cable stage, to be used for the current-sharing length, is significantly smaller than in the cable stage where the largest loss occurs (73).

#### Accuracy of Coupling Current Loss Estimation

The calculation of the coupling current loss in steady-state conditions is based on the experimental value of  $n\tau$ , drawn from the loss curve. In other words, to estimate the coupling loss for a coil in operation, it is necessary to measure first the loss on a conductor specimen under controlled conditions. The assessment of  $n\tau$  from the pitch and the expected transverse resistivity may lead to very rough loss estimates.

The accuracy of  $n\tau$  depends on the loss calibration method, the resolution and reproducibility of the loss measurements, and the number of test points in the linear range of the loss curve ( $n\tau$  is inferred by linear regression from the initial slope of the loss curve). In multistage, cabled conductors, the loss curve may show slight deviations from linearity at a very low field rate, possibly due to very large time constants associated with the very small factor  $n$  (41). In such cases, the actual operating field rate should drive the choice of the range of the loss curve from which  $n\tau$  must be inferred. Typical experimental values of  $n\tau$  have an error bar larger than 10% (74).

When the coupling currents are not in steady state (transient field change or high frequency oscillations), the product  $n\tau$  from the loss curve must be broken down into  $n$  and  $\tau$  to use the transient formulas from Eqs. (24–29). This is very hard to do, especially in the case of multiple time constants. A measurement of  $\tau$  from the decay time of the induced currents is discussed in Ref. 75 for conductors with a single (or dominant) time constant.

Besides the issue of breaking down the overall  $n\tau$ , the prediction of non-steady-state loss in conductors with multiple time constants is made difficult by the choice between a model with independent current loops or with nested loops (see discussion of Fig. 7). For  $\tau \gg 1$  or ramp time  $t_0 \leq \tau$ , the error in loss estimation may be large and the transient formulas should be used only to provide the order of magnitude of the expected loss.

In large cable-in-conduit conductors, the accuracy of the coupling loss estimation is much worse than in multifilamentary composites. On one hand, it is hard to reproduce in a short sample measurement the actual load history for interstrand resistance and hence to obtain a reliable result for  $n\tau$  in operation. On the other hand, the nonhomogeneous distribution of the interstrand resistance may give rise, in a coil, to complex patterns of BICCs, not measurable on a short sample (61, 62). In the presence of a transport current, more incertitude is added because of the occurrence of dynamic resistance loss and the variation of current density across the cable (44, 73).

#### CONCLUDING REMARKS

Although the basic mechanisms of ac losses in superconductors are well known, the optimization of low-loss conductor design and the prediction of ac losses in operation are still the subject of several studies and much R&D activity. The complexity of the conductor layout and the variety of the boundary conditions conclusively show that the existing formulas and models need systematic improvement.

The challenging task for the designer consists in distinguishing, for an individual application, the relevant from the negligible sources of ac losses and implementing measures to reduce their effect on the overall performance of the device. Reduction of ac losses is always a welcome result, but it must be carefully balanced with possible negative effects (e.g., reduction of the overall current density due to passive components or poor stability due to impaired current sharing).

The task of calculating ac losses may be challenging, although no complex finite-element models are required, as they are for example in thermohydraulic and mechanical analyses. For the hysteresis loss a rough estimate is easy to make, because the results for a single filament can be reliably scaled to large conductors. However, very accurate predictions are not easy.

Estimating the coupling current loss for large conductors may be difficult. The electrodynamic behavior can be predicted only on the basis of several assumptions, which can hardly be verified. Small-scale experiments, even on full-sized conductors, are mostly unable to fully reproduce the actual operating conditions.

#### BIBLIOGRAPHY

1. D. Saint-James, E. J. Thomas, and G. Sarma, *Type II Superconductivity*, New York: Pergamon, 1969.
2. A. M. Campbell and J. E. Evetts, Flux vortices and transport currents in type II superconductors, *Adv. Phys.* **21** (90): 333–357, 1972.
3. H. Ullmaier, *Irreversible Properties of Type II Superconductors*, New York: Springer-Verlag, 1975.
4. C. P. Bean, Magnetization of hard superconductors, *Phys. Rev. Lett.*, **8**(6): 250–253, 1962.

5. C. Y. Pang, P. G. McLaren, and A. M. Campbell, Losses in superconducting cylinders in transverse field, *Proc. ICEC*, **8**: 739–743, 1980.
6. A. P. Martinelli and B. Turck, Some effects of field orientation on the magnetization of superconducting wires, *Cryogenics*, **18**: 155–161, 1978.
7. Y. B. Kim, C. F. Hempstead, and A. R. Strnad, Critical persistent currents in hard superconductors, *Phys. Rev. Lett.*, **9**: 306–309, 1962.
8. H. Brechna and C. Y. Hwang, Hysteretic effects in pulsed superconducting magnets, *Cryogenics*, **19**: 217–223, 1979.
9. K. Kwasnitza and P. Bruzzone, Hysteresis losses of multifilament superconductors in superimposed dc and ac fields, *Cryogenics*, **21**: 593–597, 1981.
10. M. N. Wilson, *Superconducting Magnets*, London: Oxford Univ. Press, 1983.
11. E. W. Collings, *Applied Superconductivity, Metallurgy and Physics of Titanium Alloys*, New York: Plenum, 1986.
12. M. P. Mathur et al., Anisotropy of the critical current density in the NbTi filaments of round composite superconductors, *J. Appl. Phys.*, **43** (9): 3831–3833, 1973.
13. H. H. Farrell, G. H. Gilmer, and M. Suenaga, Grain boundary diffusion and growth of intermetallic layers: Nb<sub>3</sub>Sn, *J. Appl. Phys.*, **45**: 4025–4035, 1974.
14. T. Schild, J. L. Duchateau, and D. Ciazynski, Influence of the field orientation on the critical current density of Nb<sub>3</sub> strands, *IEEE Trans. Appl. Supercond.*, **7**: 1512–1515, 1997.
15. S. Takacs, M. Polak, and L. Krempasky, Critical currents of NbTi tapes with differently oriented anisotropic defects, *Cryogenics*, **23**: 153–159, 1983.
16. M. Jergel and S. Takacs, The dependance of critical current of Nb<sub>3</sub>Sn tapes on the direction of a transverse magnetic field, *J. Low Temp. Phys.*, **10** (3/4): 469, 1973.
17. J. F. Artaud, Distribution de courant dans des fils supraconducteurs soumis à un champ magnétique d'orientation quelconque, PhD Thesis, Univ. de Aix-Marseille, France, 1994.
18. T. S. Kreilick, E. Gregory, and J. Wong, Influence of filament spacing and matrix material on the attainment of high quality, uncoupled NbTi fine filaments, *IEEE Trans. Magn.*, **24**: 1033–1036, 1988.
19. M. Polak et al., Anomalous magnetization behavior in fine filamentary NbTi superconducting wires, *IEEE Trans. Appl. Supercond.*, **3**: 150–153, 1993.
20. M. D. Sumption and E. W. Collings, Influence of twist pitch and sample length on proximity effect coupling in multifilamentary composites described in terms of a field independent, two current region model, *Cryogenics*, **34**: 491–505, 1994.
21. A. K. Gosh, K. E. Robins, and W. B. Sampson, Magnetization measurements of multifilamentary Nb<sub>3</sub>Sn and NbTi conductors, *IEEE Trans. Magn.*, **21**: 328–331, 1985.
22. R. B. Goldfarb and K. Itoh, Reduction of interfilament contact loss in Nb<sub>3</sub> superconductor wires, *J. Appl. Phys.*, **75**: 2115–2118, 1994.
23. R. B. Goldfarb and J. W. Ekin, Hysteresis losses in fine filament internal-tin superconductors, *Cryogenics*, **26**: 478–481, 1986.
24. S. S. Shen, Effect of the diffusion barrier on the magnetic properties of practical Nb<sub>3</sub>Sn composites, *Adv. Cryog. Eng. Mat.*, **28**: 633–638, 1982.
25. T. Ogasawara et al., Alternating field losses in superconducting wires carrying dc transport currents: Part 1. Single core conductors, *Cryogenics*, **19**: 736–740, 1979.
26. T. Ogasawara et al., Effective resistance of current carrying superconducting wire in oscillating magnetic field 1: Single core composite conductor, *Cryogenics*, **16**: 33–38, 1976.
27. S. S. Shen and R. E. Schwall, Interaction of transport current and transient external field in composite conductors, *IEEE Trans. Magn.*, **15**: 232–235, 1979.
28. W. T. Norris, Calculation of hysteresis losses in hard superconductors carrying ac: Isolated conductors and edges of thin sheets, *J. Phys. D*, **3**: 489–507, 1970.
29. T. Ogasawara et al., Transient field losses in multifilamentary composite conductors carrying dc transport currents, *Cryogenics*, **20**: 216–222, 1980.
30. D. Ciazynski et al., Ac losses and current distribution in 40 kA NbTi and Nb<sub>3</sub>Sn conductors for NET/ITER, *IEEE Trans. Appl. Supercond.*, **3**: 594–599, 1993.
31. M. N. Wilson et al., Experimental and theoretical studies of filamentary superconducting composites, *J. Phys. D*, **3**: 1517–1546, 1970.
32. W. J. Carr, Jr., Conductivity, permeability and dielectric constant in a multifilament superconductor, *J. Appl. Phys.*, **46** (9): 4043–4047, 1975.
33. K. Kwasnitza and P. Bruzzone, Large ac losses in superconducting Nb<sub>3</sub>Sn cable due to low transverse resistance, *Proc. ICEC11*, 741–745, Berlin, 1986.
34. K. Kwasnitza and I. Horvath, Experimental evidence for an interaction effect in the coupling losses of cabled superconductors, *Cryogenics*, **23**: 9–14, 1983.
35. A. M. Campbell, A general treatment of losses in multifilamentary superconductors, *Cryogenics*, **22**: 3–16, 1982.
36. B. Turck, Effect of the respective position of filament bundles and stabilizing copper on coupling losses in superconducting composites, *Cryogenics*, **22**: 466–468, 1982.
37. A. Nijhuis et al., Coupling loss time constant in full size Nb<sub>3</sub>Sn CIC model conductors for fusion magnets, *Adv. Cryog. Eng. Mat.*, **42B**: 1281–1288, 1996.
38. H. Brechna and G. Ries, Ac losses in superconducting synchrotron magnets, *IEEE Trans. Nucl. Sci.*, **18** (3): 639–642, 1971.
39. G. Ries, Ac losses in multifilamentary superconductors at technical frequencies, *IEEE Trans. Magn.*, **13** (1): 524–527, 1977.
40. J. P. Soubeyrand and B. Turck, Losses in superconducting composites under high rate, pulsed transverse field, *IEEE Trans. Magn.*, **15**: 248–251, 1979.
41. A. Nijhuis et al., Electromagnetic and mechanical characterization of ITER CS-MC conductors affected by transverse cyclic loading, part 1: Coupling currents loss, *IEEE Trans. Appl. Supercond.*, **9**: 1063–1072, 1999.
42. A. M. Campbell, The effect of transport current and saturation on the losses of multifilamentary superconducting wires, *Cryogenics*, **21**: 107–112, 1981.
43. D. Ciazynski, Effect of the transport current on the losses of a superconducting composite under fast changing magnetic field, *IEEE Trans. Magn.*, **21**: 169–172, 1985.
44. P. Bruzzone et al., Test results for the high field conductor of the iter central solenoid model coil, *Adv. Cryog. Eng.*, **45**, to be published.
45. J. H. Murphy et al., Field orientation dependence of ac losses in rectangular multifilamentary superconductors, *Adv. Cryog. Eng.*, **22**: 420–427, 1975.
46. B. Turck et al., Coupling losses in a rectangular multifilamentary composite, *Cryogenics*, **22**: 441–450, 1982.
47. E. Seibt, Investigations of a steel reinforced NbTi superconducting flat cable for toroidal field magnets, *IEEE Trans. Magn.*, **15**: 804–807, 1979.
48. P. Bruzzone, Ac losses in high current superconductors for nuclear fusion magnets, Ph. D. Thesis, ETH 8224, Zurich, 1987.
49. G. Ries and S. Takács, Coupling losses in finite length of superconducting cables and in long cables partially in magnetic field, *IEEE Trans. Magn.*, **17**: 2281–2284, 1981.

50. S. Takács, Coupling losses in cables and in spatially changing ac fields, *Cryogenics*, **22**: 661–665, 1982.
51. A. A. Akhmetov, A. Devred, and T. Ogitsu, Periodicity of crossover currents in a Rutherford-type cables subjected to a time-dependent magnetic field, *J. Appl. Phys.*, **75**: 3176–3183, 1994.
52. L. Krempasky and C. Schmidt, Theory of “supercurrents” and their influence on field quality and stability of superconducting magnets, *J. Appl. Phys.*, **78**: 5800–5810, 1995.
53. A. P. Verweij, Electrodynamics of superconducting cables in accelerator magnets, Ph.D. Thesis, Univ. of Twente, Enschede, The Netherlands 1995.
54. L. Krempasky and C. Schmidt, Ramp rate limitation in large superconducting magnets due to “supercurrents,” *Cryogenics*, **36**: 471–483, 1996.
55. P. Bruzzone, Ac losses and stability on large cable-in-conduit superconductors, *Phys. C*, **310**: 240–246, 1998.
56. A. Nijhuis, P. Bruzzone, and H. H. J. ten Kate, Influence of Cr plating on the coupling loss in cable-in-conduit conductors, *Appl. Supercond. 1997, Inst. Phys. Conf.*, **158**: 921–924, 1997.
57. M. D. Sumption et al., Contact resistance and cable loss measurements of coated strands and cables wound from them, *IEEE Trans. Appl. Supercond.*, **5**: 692–695, 1995.
58. M. D. Sumption et al., Calorimetric measurements of the effect of nickel and Stabrite coatings and resistive cores on ac loss in accelerator cables under fixed pressure, *Adv. Cryog. Eng.*, **42**: 1303–1311, 1996.
59. E. W. Collings et al., Magnetic studies of ac loss in pressurized Rutherford cables with coated strands and resistive cores, *Adv. Cryog. Eng.*, **42**: 1225–1232, 1996.
60. D. Richter et al., Dc measurement of electrical contacts between strands in superconducting cables for the LHC main magnets, *IEEE Trans. Appl. Supercond.*, **7**: 786–789, 1997.
61. T. Hamajima et al., Ac loss performance of the 100 kWh SMES model coil, *Proc. of Magnet Technology Conf.*, **16**, Sept. 99, Ponte Vedra, FL.
62. A. Kawagoe et al., Increase of interstrand coupling losses of superconducting cable-in-conduit for actual condition of sweep rate, *IEEE Trans. Appl. Supercond.* **9**: 727–730, 1999.
63. J. M. Depond et al., Examination of contacts between strands by electrical measurement and topographical analysis, *IEEE Trans. Appl. Supercond.*, **7**: 793–796, 1997.
64. T. M. Mower and Y. Iwasa, Experimental investigation of ac losses in cabled superconductors, *Cryogenics*, **26**: 281–292, 1986.
65. K. Kwasnitza, A. Sultan, and S. Al-Wakeel, Ac losses of a 10 kA NbTi cable-in-conduit superconductor for SMES application, *Cryogenics*, **36**: 27–34, 1996.
66. M. Nishi et al., Test results of the DPC-TJ, a 24 kA–40 A/mm<sup>2</sup> superconducting test coil for fusion machines, *Fusion technology 1992, Proc. 17th SOFT*, 1993, p. 912–916.
67. P. Bruzzone, A. Nijhuis, and H. H. J. ten Kate, Effect of Cr plating on the coupling current loss in cable-in-conduit conductors, *Proc. ICMC 96, Kitakyushu, Japan, May 1996, Amsterdam, The Netherlands Elsevier*, p. 1243–1248.
68. Y. Ipatov, P. Dolgosheev, and V. Sytnikov, Prospective barrier coatings for superconducting cables, *Supercond. Sci. Technol.*, **10**: 507–511, 1997.
69. A. Nijhuis et al., Parametric study on coupling loss in subsize ITER Nb<sub>3</sub>Sn cabled specimen, *IEEE Trans. Magn.*, **32**: 2743–2746, 1996.
70. P. Bruzzone, A. Nijhuis, and H. H. J. ten Kate, Contact resistance and coupling loss in cable-in-conduit of Cr plated Nb<sub>3</sub>Sn strands, *Proc. MT-15, Beijing, October 1997, Science Press*, 1998, p. 1295–1298.
71. A. Nijhuis, H. H. J. ten Kate, and P. Bruzzone, The influence of Lorentz forces on the ac loss in sub-size cable-in-conduit conductors for ITER, *IEEE Trans. Appl. Supercond.*, **7**: 262–265, 1997.
72. G. Ries, Stability in superconducting multistrand cables, *Cryogenics*, **20**: 513–519, 1980.
73. P. Bruzzone, Stability under transverse field pulse of the Nb<sub>3</sub>Sn ITER cable-in-conduit conductor, *Proc. of Magnet Technology Conf.*, **16**, Sept. 1999, Ponte Vedra, FL.
74. P. Bruzzone et al., Coupling currents losses bench mark test of ITER subsize conductor, *IEEE Trans. Magn.*, **32**: 2826–2829, 1996.
75. L. Krempasky and C. Schmidt, Time constant measurements in technical superconductors, *Cryogenics*, **39**: 23–33, 1999.

PIERLUIGI BRUZZONE  
Centre de Recherches en Physique  
des Plasmas

## SUPERCONDUCTIVITY: MICROWAVE TECHNOLOGY

Superconductivity is a remarkable state of matter in which electric current can flow inside materials without any detectable resistance. This phenomenon was discovered by the Dutch physicist H. K. Onnes in 1911. While studying the temperature dependence of the electric resistance of mercury in his newly invented helium liquefier, Onnes found that below a temperature of about 4 K, the resistance abruptly fell to below measurable limits.

In a superconductor, perfect electric conductivity or zero resistivity occurs only below a “critical temperature”  $T_c$ . This happens because the electric current, instead of being carried by single electrons, is carried by pairs of electrons called Cooper pairs which have the ability to conserve energy because of quantum mechanical reasons. A direct current induced in a superconducting ring has been shown to persist for over two years without any measurable decay. From this type of experiment the upper limit on resistivity  $\rho$  is  $\sim 10^{-25} \Omega \cdot \text{m}$ . Since  $\rho$  for copper is  $\sim 10^{-8} \Omega \cdot \text{m}$ , a factor of  $10^{17}$  larger, it is believed that the electric resistance of a superconductor is truly zero.

Another important characteristic of superconductors was the discovery by W. Meissner and R. Ochsenfeld in 1933 that a superconductor expels all magnetic flux from its interior; that is, a superconductor is also a perfect diamagnet. This phenomenon is known as the Meissner effect.

A key area of technological applications of superconductors is in high-frequency devices, particularly at frequencies in the radio-wave, microwave, and millimeter-wave spectral ranges. Broadly speaking, two principal areas of high-frequency applications can be identified:

- *Passive devices*, typically using resonant and transmission line structures, exploiting the low loss properties of superconductors.
- *Active elements*, such as oscillators, mixers, logic elements, and other similar devices, which exploit a macroscopic quantum coherence property of the charge carriers, called the Josephson effect.

The price to pay to achieve the superior performance is the need to cool the superconducting device to temperatures well



**Table 1. Some Commercially Used Superconducting Materials and Their Physical Properties**

| Material           | $T_C$ (K) | $\lambda_L$ (nm) | $\xi$ (nm)       | $H_{C1}$ (Oe) | $H_{C2}$ (Oe)        | Crystal Structure   |
|--------------------|-----------|------------------|------------------|---------------|----------------------|---------------------|
| Pb                 | 7.2       | 37               | 83               | 803           | 803                  | Face-centered cubic |
| Nb                 | 9.2       | 32               | 39               | p 1600        | p 3200               | Body-centered cubic |
| NbN                | 16        | 50               | 4                | 300           |                      | B1                  |
| Nb <sub>3</sub> Ge | 23.2      |                  |                  |               | $3.6 \times 10^{6a}$ | A15                 |
| YBCO <sup>b</sup>  | 93        | 140              | p 2              | 200           | p $10^6$             | Orthorhombic        |
| TBCCO <sup>b</sup> | 127       | 220 <sup>c</sup> | 2.6 <sup>c</sup> |               |                      | Tetragonal          |
| BSCCO <sup>b</sup> | p 90      | 500              | 0.16             |               |                      | Tetragonal          |

<sup>a</sup> At 4.2 K.

<sup>b</sup> YBCO, TBCCO, BSCCO, and most other high- $T_C$  superconductors are highly anisotropic. The values given are for microwave current flowing in the  $ab$ -plane, which is often the plane of epitaxial growth and the desired geometry in most technological applications.

<sup>c</sup> Average of  $ab$ -plane and  $c$ -axis value.

below room temperature. This means the use of cryogenic fluids like liquid nitrogen or liquid helium, or a mechanical cryocooler. Despite the need for a cryogenic environment, there are many applications where the superior performance of superconductors prevails over conventional devices.

## SUPERCONDUCTING MATERIALS

Superconductivity has been observed in diverse types of materials including pure metals, alloys, semimetals, organic materials, semiconductors, polymers, and even elemental insulators. Presently, at least 26 of the naturally occurring elements are known to be superconducting at sufficiently low temperatures at ambient pressure, and the number of alloys and intermetallic compounds are well above 1000.

A significant advancement in understanding superconductivity was the theory of Bardeen, Cooper, and Schrieffer in 1957. This theory successfully described the microscopic mechanism of superconductivity and explained the properties of most superconductors for nearly 30 years.

Then came the era of high-temperature superconductivity, with the groundbreaking discovery by Bednorz and Müller in 1986 of superconductivity in a Ba-La-Cu-O compound with  $T_C$  around 30 K, which subsequently earned them a Nobel Prize. This discovery rekindled the interest in superconductivity and spurred the hopes of finding new technological applications. The efforts led to the discovery of other copper-oxide-based superconductors that have critical temperatures well in excess of the boiling point of nitrogen (77 K), an easily available and cheap cryogenic fluid usable for commercial applications. Presently the highest critical temperature is about 150 K at high pressures in HgBa<sub>2</sub>Ca<sub>2</sub>Cu<sub>3</sub>O<sub>8</sub>.

Superconductors fall in two broad categories, namely, "high critical temperature" (or "high- $T_C$ ") and "low critical temperature" (or "low- $T_C$ ") materials. (There are many other ways of classifying them, such as Type I and Type II superconductors, heavy fermion and conventional superconductors, and so on.) Some examples of low- $T_C$  superconductors, particularly relevant to commercial applications, are niobium and its alloys, such as NbN and Nb<sub>3</sub>Ge. These have critical temperatures in the region of 10 K to 20 K. The high- $T_C$  superconductors that are most commonly used include YBa<sub>2</sub>Cu<sub>3</sub>O<sub>7</sub> (commonly referred to as YBCO or Y-123,  $T_C \sim 93$  K), Bi<sub>2</sub>Sr<sub>2</sub>CaCu<sub>2</sub>O<sub>8</sub> (commonly referred to as BSCCO or Bi-2212,  $T_C \sim 90$  K), and Tl<sub>2</sub>Ba<sub>2</sub>Ca<sub>2</sub>Cu<sub>3</sub>O<sub>10</sub> (TBCCO or Tl-2223,  $T_C \sim 125$  K).

Besides the critical temperature, some other relevant physical parameters that characterize a superconductor are as follows:

- *The London penetration depth ( $\lambda_L$ ):* The depth to which an applied direct current (dc) magnetic field is confined because of the Meissner effect.
- *The coherence length ( $\xi$ ):* The length scale over which the two electrons forming the Cooper pairs are separated.
- *The (thermodynamic) critical field ( $H_C$ ):* The magnetic field whose associated free energy is equal to the free energy change in the superconducting transition. A "Type I" superconductor will completely expel the applied field below  $H_C$  and will become normal abruptly as the applied field exceeds this magnitude, assuming no demagnetization effects. For "Type II" superconductors (which includes practically all of the superconductors of technological interest), the field starts to penetrate in the form of vortices at a "lower critical field"  $H_{C1}$  ( $< H_C$ ), but superconductivity is not quenched until the applied field exceeds the "upper critical field"  $H_{C2}$  ( $> H_C$ ).

Some superconductors of relevance to microwave applications are described in Table 1 along with their relevant properties ( $\lambda_L$ ,  $\xi$ ,  $H_{C1}$ , and  $H_{C2}$  quoted are extrapolated to  $T = 0$  K).

### Substrates for Thin-Film Superconductors

For most electronic applications, including microwave devices, superconductors are necessarily used in the form of thin films. This requires the use of a foreign material for a substrate. Some of the popular substrates include sapphire ( $\alpha$ -Al<sub>2</sub>O<sub>3</sub>), lanthanum aluminate (LaAlO<sub>3</sub>), and magnesium oxide (MgO). Ferroelectric substrates like strontium titanate (SrTiO<sub>3</sub>) and KTa<sub>1-x</sub>Nb<sub>x</sub>O<sub>3</sub> are also used to achieve tunability at the expense of additional dielectric loss.

The electrical parameters that characterize the properties of a substrate material are its dielectric constant  $\epsilon$  and its loss tangent  $\tan \delta$ . (The loss tangent of a substrate is defined as  $\tan \delta = \sigma / \omega \epsilon$ , where  $\sigma$  is the conductivity.) Other factors such as environmental stability, mechanical strength, chemical inertness, and absence of magnetic moment are also important. Close lattice match between the deposited film and the substrate is essential to achieve good epitaxial growth. These parameters have been measured and investigated by various researchers using different techniques for most substrate materials.

In general, the dielectric constant would be more or less frequency independent unless the frequency is close to a resonance frequency of the material. The dielectric constant is also expected to have a weak temperature dependence for nonferroelectric substances. The loss tangent is expected to increase

**Table 2. Some Commonly Used Substrate Materials for Superconducting Thin-Film Devices**

| Material           | Dielectric Constant | Loss Tangent         | Crystal Structure | Growth Surface | Remarks                              |
|--------------------|---------------------|----------------------|-------------------|----------------|--------------------------------------|
| LaAlO <sub>3</sub> | 25                  | $5 \times 10^{-6}$   | Rhombohedral      | (110)          | Usually twinned                      |
| YAlO <sub>3</sub>  | 16                  | $10^{-3}$            | Orthorhombic      | (110)          |                                      |
| MgO                | 9.65                | $6.2 \times 10^{-6}$ | Cubic             | (100)          | Good lattice match                   |
| Sapphire           | 8.6                 | $3.8 \times 10^{-8}$ | Hexagonal         | (11 02)        | Very low loss                        |
| NdGaO <sub>3</sub> | 23                  | $3.2 \times 10^{-4}$ | Orthorhombic      | (110)          | Good for multilayer circuits         |
| SrTiO <sub>3</sub> | 300                 | $3 \times 10^{-4}$   | Cubic             | (100)          | Good for tunable device applications |
| YSZ                | 27                  | $7.4 \times 10^{-4}$ | Cubic             | (100)          |                                      |

with temperature. Table 2 provides the most commonly accepted values of these material parameters for some of the substrates, measured at 10 GHz and 77 K unless otherwise stated.

Most of these materials have high  $\epsilon$ ; this imposes certain restrictions, especially for high-frequency applications. A high  $\epsilon$  reduces the dimension of the device and thus imposes stringent dimensional tolerance, especially in the millimeter-wave region. But consideration of dielectric constant has to be sacrificed for some even more important parameters, namely, a lattice match between the substrate and the film, and not widely different thermal expansivity from the superconductor. A lattice mismatch would create atomic level strain and leads to poor superconducting properties. Another criterion is that losses in the substrate must be negligible compared to the losses in the superconductor. Heavy twinning in materials like LaAlO<sub>3</sub> creates problems in multilayer films. When the substrate is reheated for the deposition of a new layer, the thermally induced movement of the twin boundaries strain the previously deposited layers.

There are numerous methods of depositing superconducting films on substrates, the most common being pulsed laser deposition (PLD), coevaporation, and off-axis sputtering or physical vapor deposition (PVD). Each of these methods has its pros and cons. For example, PLD produces high-quality films at a fast rate (a few angstroms per second) but only over relatively small areas, typically less than 2 in. diameter. The PVD method is useful for producing larger area films, but the deposition rate is only a fraction of an angstrom per second. Another method of film deposition, namely, metal-organic chemical vapor deposition (MOCVD), popular in the semiconductor industry, has also been tried for high- $T_C$  superconductors, although somewhat less successfully owing to the chemical and structural complexity of the high- $T_C$  materials. We refer the reader interested in deposition technique to some useful resources (see, for example, Ref. 1).

The "quality" of a deposited superconducting film is characterized by several parameters, such as the critical temperature  $T_C$ , the transition width  $\Delta T_C$ , and the critical current density  $J_C$ , which is the maximum lossless current per unit cross-sectional area that the film can carry. The higher the  $T_C$  and  $J_C$  and the lower the  $\Delta T_C$ , the better the film.

## THE SUPERCONDUCTOR-MICROWAVE INTERACTION

A superconductor has zero electrical resistance only at zero frequency (dc). At any finite frequency, it exhibits losses that increase with frequency. A superconductor can be regarded as being composed of two types of charge carriers, one con-

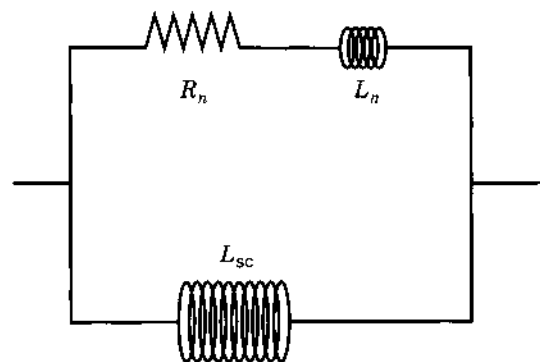
sisting of the lossless Cooper pairs and the other being lossy "quasiparticles." Such a description is known as the "two-fluid model" and is applied to all known superconductors.

The finite frequency [alternating current (ac)] response of a superconductor is characterized in terms of a generalization of Ohm's law:  $\mathbf{J} = \sigma_s \mathbf{E}$ , where  $\mathbf{E}$  is the applied electric field and  $\sigma_s = \sigma_1 - i\sigma_2$  is the complex conductivity. The normal part  $\sigma_1$  comes from the motion of the quasiparticles, and the imaginary part  $\sigma_2$  comes from that of the Cooper pairs. For most superconductors well below the transition temperature,  $\sigma_2 \gg \sigma_1$ . The frequency  $\omega$ , wave vector  $\mathbf{k}$ , temperature  $T$ , and current density  $\mathbf{J}$  dependence of  $\sigma_s$  contains all the information about the electrical properties of the superconductor. For a normal metal,  $\sigma_s = \sigma_n$  is purely real.

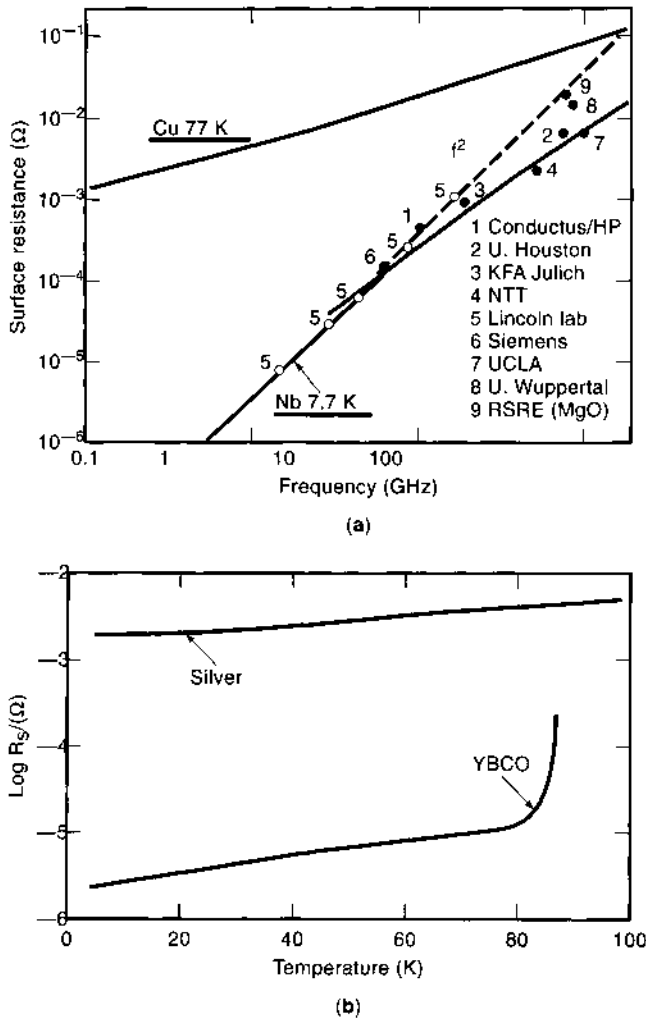
Figure 1 shows a lumped circuit equivalent of a superconductor. The large inductive element  $L_{sc}$  represents the lossless carriers, and the parallel branch with a series combination of a resistor  $R_n$  and an inductance  $L_n$  represents the quasiparticles. At dc, the large inductor shorts out the resistive branch, leading to zero resistance, while at finite frequencies there is always dissipation.

As the excitation frequency increases, the impressed electromagnetic field is confined closer to the surface. In a normal metal, this is known as the skin effect and the characteristic length of confinement (the "skin depth") is given by  $\delta_n = (2/\mu\omega\sigma_n)^{1/2}$  where  $\sigma_n$  is the (normal) conductivity.  $\delta_n$  for Copper at 10 GHz and 77 K is 0.55  $\mu\text{m}$ . In a superconductor at temperature  $T \ll T_C$ , the skin depth is replaced by the "London penetration depth"  $\lambda_L = 1/(\mu\omega\sigma_2)^{1/2}$ . The penetration depth is typically independent of frequency from dc to microwave frequencies and is comparable to the London penetration depth in Table 1.

The interaction of a time-varying electromagnetic field and a metal (normal or superconducting) at microwave frequencies



**Figure 1.** Two-fluid model equivalent circuit of a superconductor.



**Figure 2.** (a) Comparison of the surface resistance of YBCO at 77 K with other superconducting materials. The horizontal bands represent YBCO. (From Ref. 3, with permission.) (b) Temperature dependence of the surface resistance of YBCO grown by a number of techniques compared with niobium and niobium-tin superconductors at the same reduced temperature. The scale on the upper axis refers to the actual temperature of the Nb<sub>3</sub>Sn and Nb, while the lower scale refers to the actual temperature of the YBCO films. The  $T_C$  of the three materials coincides.  $f = 8$  GHz.

is best described in terms of the surface impedance defined as  $Z_S = R_S + iX_S = (i\mu\omega/2\sigma)^{1/2}$ .  $R_S$  and  $X_S$  are the surface resistance and reactance, respectively. As can be seen from the above discussion,  $R_S$  and  $X_S \propto \omega^{1/2}$  in a normal metal, where  $\sigma = \sigma_n$ . On the other hand,  $R_S \propto \omega^2$  and  $X_S \propto \omega$  in an "ideal" superconductor. This can be easily verified by assuming  $\sigma_1$  and  $\lambda_L$  to be frequency-independent and  $\sigma_2 \gg \sigma_1$ .

The property that makes superconductors attractive for passive microwave circuits is that their surface resistance is orders of magnitude lower than that of normal metals. For example, thin-film YBCO has a surface resistance of  $<0.1$  mΩ at 77 K and 10 GHz compared to 8.7 mΩ of copper under identical conditions, and this difference is even higher at lower temperatures. Figure 2 shows the measured frequency and temperature dependence of  $R_S$  of some of the most common superconducting and nonsuperconducting materials (3).

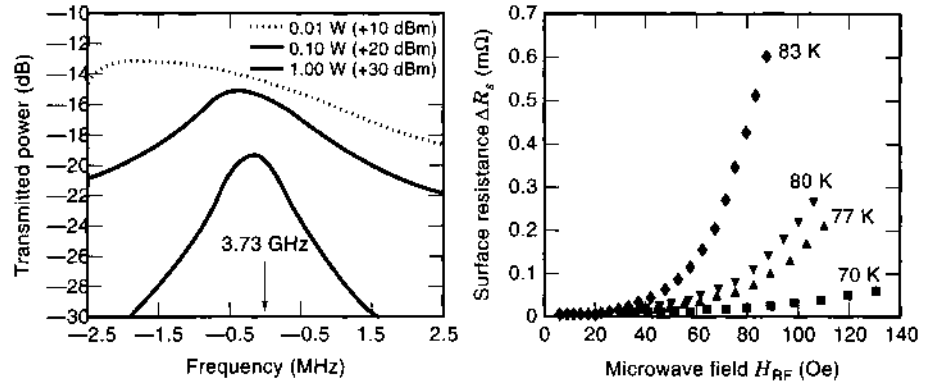
### Nonlinearity in High- $T_C$ Superconductors

For most superconductors, particularly the high- $T_C$  superconductors, the surface impedance increases with applied microwave power, even for moderate power levels ( $<0.1$  W). In other words (unlike normal metals), they act as nonlinear circuit elements. This nonlinear response often poses serious constraints on the utility of superconductors in practical devices. Some specific examples of such limitations are as follows:

- Degradation of insertion loss with increasing power.
- Generation of harmonic frequencies: For an applied signal at frequency  $f$ , most of the harmonic power is generated at a frequency of  $3f$ . Given today's tight use of the electromagnetic spectrum, practically all frequency ranges have been preassigned by the appropriate authorities (e.g., the Federal Communication Commission in the United States) with just the required amount of bandwidths. Thus, radiation by a device at frequencies far outside the designated frequency of operation is unacceptable.
- Frequency mixing, leading to intermodulation products (intermods): For two nearby frequencies  $f_1$  and  $f_2$ , the strongest intermods are produced at  $2f_1 - f_2$  and  $2f_2 - f_1$ . The result is that spurious frequencies are produced in devices.

Figure 3 shows the effect of nonlinearity on a suspended line resonator with a resonance frequency of 3.73 GHz (4). At low power levels the response in the frequency domain is close to a Lorentzian. As the input power is increased, the shape begins to distort, and the  $Q$  and the resonance frequency go down. This also highlights the point that at high power levels the  $Q$  of a superconducting resonator cannot be defined in terms of a "3 dB bandwidth." A quantitative measure of the nonlinearity of a device is specified in terms of its third-order intercept (TOI), defined as the input power at which power output at the fundamental and the third harmonic equal each other. Single-tone TOI for most commercially used HTSC materials is at least 70 dBm (10 kW).

All these effects together limit the maximum power that a passive superconducting circuit can handle, and raising the power handling capacity has been one of the prime concerns of material scientists and engineers alike. Much of the nonlinearity of the high- $T_C$  materials is ascribable to material properties (e.g., granularity and "flux pinning") and is expected to improve with synthesis and deposition techniques. From a design point of view, nonlinearities can be suppressed by increasing the surface area of the device, leading to lower current densities and avoiding current crowding at the edges. Thus, a planar filter can handle significantly higher power than a stripline, and a cavity resonator has  $Q$  value orders of magnitude higher than one built out of a transmission line, even factoring out the losses in the substrate. However, higher-dimensional circuits are not always practical for many applications because they take up more space and often have the problem of mode degeneracy (i.e., more than one mode of oscillation at the same frequency). The latter problem can often be solved by making the circuit slightly asymmetrical, but then modeling them with a computer becomes more cumbersome.



**Figure 3.** (Left) Frequency-domain response of a suspended resonator. Notice the deviation from a Gaussian shape as the input power increases. (Right) The  $R_s$  of a similar resonator against applied power at different temperatures. (From Ref. 4, with permission.)

## PASSIVE MICROWAVE CIRCUITS

Passive microwave circuits are some of the most promising applications of superconductors to date. In the following we shall focus on the advantages of high-temperature superconductors (HTS) materials and the special considerations that applies to them. General ideas about passive microwave devices are available in the literature. For a good review of the current state of the art, see Ref. 5. For a more elementary introduction, see Ref. 6.

Superconductors are an attractive alternative to “conventional” materials for three principal reasons:

1. *Very low surface resistance* compared to normal metals: For example, insertion loss of a superconducting filter ( $\sim 0.1$  dB) is usually at least 3 dB lower than one made out of high-quality metal (e.g., oxygen-free high-purity copper) under equivalent conditions. Thus in applications where the signal strength is low, superconductors are the material of choice.
2. *A frequency-independent penetration depth.* In normal metals, the skin depth is proportional to the inverse square root of the frequency of the signal, giving rise to strong dispersion of a wide-band signal, unacceptable in many applications. In superconductors, the penetration of electromagnetic field is dominated by the “London penetration depth,” which is independent of frequency, giving rise to phase-dispersion free propagation in the TEM mode.
3. *The ability to be fabricated in extremely compact geometries,* arising from the low loss mentioned above. Thus, a several-meter-long coaxial delay line made out of a normal conductor can be replaced with a superconducting one that’s only a few square centimeters in area, and bulky dielectric resonator filters can be replaced by compact planar superconducting filters.

### Transmission Line

A microwave transmission line can be constructed in several geometries, each having its pros and cons. The most common geometries are microstrip, stripline, and coplanar. Many other novel geometries such as the suspended line and coplanar strips are also used for special-purpose applications. The basic concepts behind transmission line and distributed circuits is the same whether or not they are made out of superconductors, and

the interested reader can find them elsewhere (see, for example, Ref. 7). Here we will briefly discuss only those issues that are specific to superconducting transmission lines.

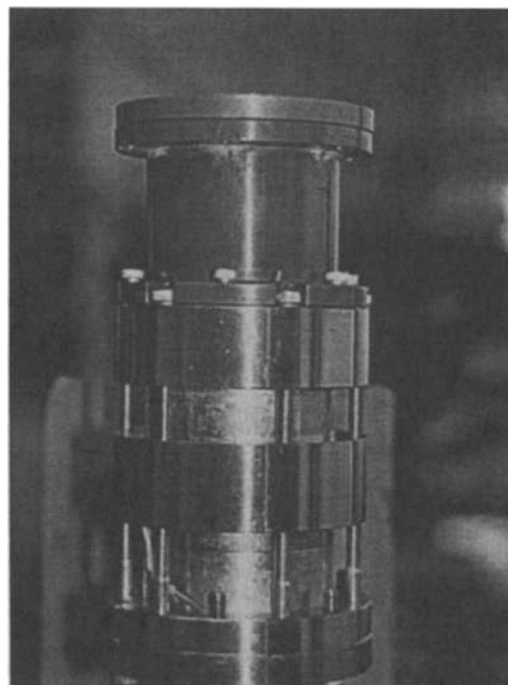
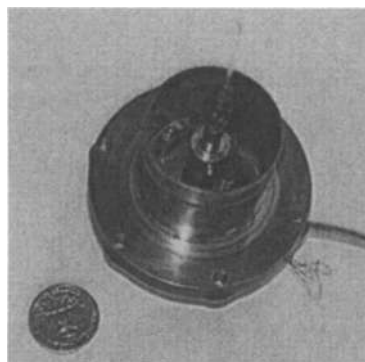
The two major distinctions between superconducting and nonsuperconducting transmission lines stem from the temperature and frequency dependence of the characteristic impedance of the line, defined as  $Z_0 = (L/C)^{1/2}$  where  $L$  and  $C$  are the equivalent lumped inductance and capacitance, respectively. In addition to the usual geometric inductance, superconductors have temperature-dependent kinetic inductance because of the presence of the Cooper pairs. Also, for sufficiently low temperatures ( $T \ll T_c$ , or more pertinently,  $\lambda_L \ll \delta$ ) the penetration depth is independent of the frequency, giving rise to a phase dispersion free propagation for the TEM mode. Nonlinear effects are another major complication in dealing with superconducting transmission lines as has been discussed earlier. The nonlinear effects are particularly worse in the edges of the line where the current density is maximum. Therefore, a good design for a passive superconducting circuit tends to avoid sharp edges as much as possible. Owing to these, full-wave analysis of superconducting transmission lines is quite a daunting task which cannot be done satisfactorily using most commercially available CAD programs and often must be carried out for individual needs.

One of the most promising applications of superconducting transmission lines is in the form of hybrid interconnects between components, both semiconducting and superconducting and both analog and digital.

### Resonators

The “quality factor” ( $Q$ ) of a resonator is defined to be  $Q = Uf_0/P$ , where  $U$  is the total energy stored in the oscillator,  $f_0$  is the resonance frequency, and  $P$  is the power dissipated per cycle. Since  $P = \int R_s H^2 dA$ , where the integral is over the surface of the material comprising the resonator walls, the  $Q$  can then be written as  $Q = \Gamma/R_s$ , where  $\Gamma$  is a geometric factor, for a resonator comprised entirely of a single material.

The advantage of a superconducting resonator over a normal metal one is the very high  $Q$  owing to the much smaller surface resistance of the former. At 10 GHz,  $Q$  as high as  $10^{11}$  can be achieved (using niobium cavities in  $TE_{011}$  mode), compared with a maximum  $Q$  of around  $10^4$  using copper. Superconducting cavities are typically operated in the  $TE_{011}$  mode since this has no electric fields on the cavity walls. Cavities have been made out of both low-temperature and high-temperature superconductors (8), although the latter is a lot harder to make owing to the ceramic nature of the material. Superconducting cavities such



**Figure 4.** A superconducting niobium cavity used for surface resistance measurement. The specimen under test is thermally insulated from the cavity using a sapphire rod, allowing it to be probed up to nearly room temperature while the cavity is still superconducting. (Right) The disassembled cavity showing the sapphire sample mount. (Left) The packaged structure.  $Q$  values of  $\sim 10^7$  to  $10^8$  at 4.2 K are routinely obtained in this setup. (Courtesy of Z. Zhai and H. Srikanth.)

as these play an important role as a research tool for precision measurements of surface impedance of superconductors and other materials. Figure 4 shows a picture of a niobium cavity that was fabricated at Northeastern University and is used to measure the surface impedance of other crystalline superconductors using the cavity perturbation technique (9).

A major application of superconducting cavities is in providing large high-frequency electric fields for accelerating subatomic and atomic particles. The fields in a cavity resonator are given by  $E_{\max} = \gamma(QP_{\text{abs}})^{1/2}$ , so that significantly larger fields can be achieved with high- $Q$  cavities for much lower absorbed powers  $P_{\text{abs}}$ . A notable example is the continuous electron beam accelerator facility (CEBAF) for heavy ions, which uses superconducting niobium cavities operating at megahertz frequencies.

Superconducting microwave cavity resonators are among the most stable frequency standards available today owing to their very high  $Q$  values. One of the applications of such high stability frequency standards is in satellite and deep space communications, where it is important to maintain precise clocks on board the satellite to improve synchronization between the ground-based clock and the satellite or space vehicle's on-board clock.

Another potential application is in the master oscillator of a Doppler radar. A Doppler radar identifies the target velocity by measuring the frequency shift of the reflected beam. Clearly, for the measurement to be reliable, the frequency of the source has to be highly stable. This is especially important for detecting targets flying close to the ground (e.g., a cruise missile), since reflections from the ground (the "ground noise") can completely mask the signal.

### Filters

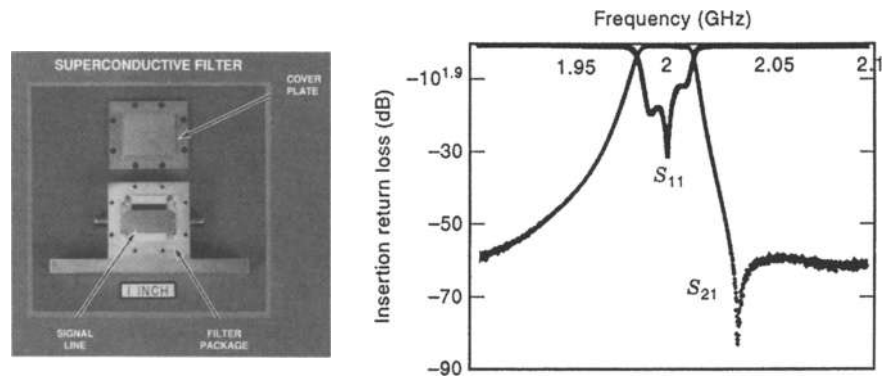
A filter is realized by a set of coupled resonators with closely spaced resonance frequencies. The response of each resonator is represented as a pole in the frequency domain. Thus a filter

with  $n$  resonating elements is called an  $n$ -pole filter. The resonators can be lumped (e.g., tank  $LC$  oscillators) or distributed (e.g., half-wave transmission lines.) Clearly, the greater the number of poles in a filter, the greater the bandwidth that can be achieved for a preset filter skirt and band ripple, or the steeper the filter skirt and smaller the band ripple for a given bandwidth. Both are desirable. However, increasing the number of elements increases the insertion loss proportionally. Thus, many applications which require *both* a small insertion loss and steep filter skirt with low band ripple cannot be realized using conventional materials such as high-purity copper.

Alternative technologies such as surface acoustic wave (SAW) devices do produce better results, but are still too lossy (typically 2 dB to 4 dB). Dielectric resonator based filters have low loss, but are too bulky for many lucrative applications. This is precisely why superconductors are the material of choice for making high-performance filters. The extremely low surface resistance allows a designer to use a lot more elements for a given amount of insertion loss.

Figure 5 shows a four-pole microstrip Chebyshev filter fabricated at the MIT Lincoln Laboratory with a 4.8 GHz center frequency and 100 MHz bandwidth patterned on an  $\text{LaAlO}_3$  substrate with YBCO film. The filter uses a "hair pin geometry"; that is, the resonators are bent with a U-turn to save wafer space. The right-hand side of Fig. 4 shows the measured insertion loss ( $S_{21}$ ) of the filter at 77 K temperature. Also shown for reference is the performance of equivalent filters made out of silver (Ag) at 77 K and gold (Au) at room temperature (300 K). The advantages of high- $T_c$  superconducting filters are quite obvious.

One of the problems in working with HTS material is the lack of availability of large wafers of acceptable quality. The maximum size of high-quality HTS films are presently limited to about 5 cm, compared to 30 cm wafers of semiconductors. This poses a problem in two cases: (1) where the center frequency is low, requiring the length of the transmission lines of a distributed filter to be long, and (2) where the bandwidth is

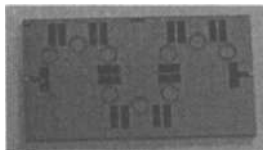


**Figure 5.** A four-pole superconducting filter and its measured performance. (Courtesy of Dr. Daniel Oates, MIT Lincoln Laboratory.)

tight (a fraction of 1%), requiring very weak coupling, and hence large separation between elements. The first problem calls for miniaturization of the device, and there are several means of accomplishing this, including the use of lumped resonators (10) and high-dielectric-constant substrates. The second problem is usually solved by a technique called staggering. Figure 6 shows a photograph of a high-performance lumped-element nine-pole Chebyshev filter.

**Frequency Agile Devices.** A filter is much more useful if it can be tuned. Several innovative approaches to tune a passive filter have been tried. One method consists of changing the resonance frequency of the individual resonators (due to change in the kinetic inductance) with temperature. The desired variation in temperature is achieved by a control line in the form of a heating element placed close to the resonators. A better method is to build the filter on a ferroelectric substrate with a low Curie temperature such as  $\text{KTa}_{1-x}\text{Nb}_x\text{O}_3$  and  $\text{Sr}_{1-x}\text{Pb}_x\text{TiO}_3$  (11). The permittivity of the substrate can be changed by applying a bias voltage, thus providing the necessary tuning. The optimal temperature of operation is slightly above the Curie temperature of the substrate to avoid hysteresis and to produce maximum tunability for a given bias voltage. Impedance matching of the input and output is done with a set of ferroelectric transformers. Since these techniques essentially manipulate the electrical length of the filter elements, the same methods can also be used to tune the parameters of other passive structures.

Another important type of filter is a "chirp" filter. A "chirp" is essentially a frequency-modulated signal whose frequency increases ("up-chirped") or decreases ("down-chirped") with time. One of the applications of chirp signals is in Doppler radars to optimize power output and bandwidth, both of which are desirable. Using a simple sinusoidal signal, however, both cannot be achieved simultaneously. A higher bandwidth means a shorter pulse width which limits the maximum power transmitted per pulse and vice versa. A chirped waveform gets around this lim-



**Figure 6.** A nine-pole lumped Chebyshev filter from Superconductor Technologies Inc. (Courtesy of Dr. Balam Willemsen, Superconductor Technologies, Inc.)

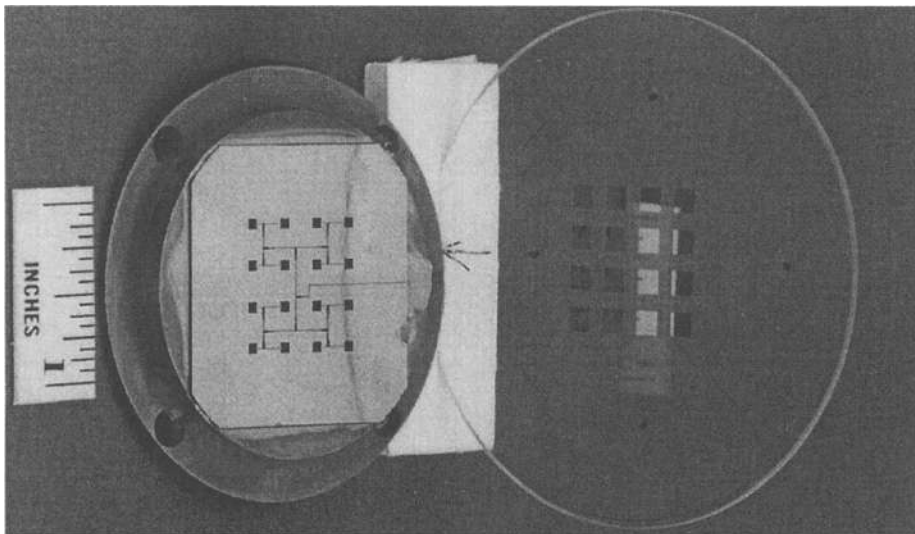
itation. Another application is to compensate the distortion of a wave packet after passing through a dispersive transmission line. The principal advantage of constructing a chirp filter out of superconducting material is low dispersiveness.

A superconducting chirp filter in microstrip geometry consists of a series of quarter-wave forward-coupled transmission lines which are deliberately decoupled for a specified length between the couplings. The transmission line is wound in the form of a spiral to optimally use the film surface. As a pulse is applied to the input, the individual frequency components couple to the output line that corresponds to their resonance frequencies. Depending on whether the resonators are increasing or decreasing in length down the line, a down-chirped or up-chirped signal is obtained at the output.

### Antenna

Using superconducting antennas can improve efficiency enormously because of their lower loss (for a general discussion of antennas, see Ref. 12). In particular, there are three situations when a superconducting antenna can outperform one made out of a normal metal sufficiently to merit its use:

1. *Electrically Short Antennas.* A dipole antenna has maximum radiation efficiency when its characteristic length is an integral multiple of the wavelength being radiated. For low-frequency applications (such as underwater communication, which involves frequencies of the order of 15 kHz), this length can be unrealistically large, and hence a shorter antenna has to be deployed. The radiated power  $P$  is proportional to  $(l/\lambda)^2$  for a linear (electric dipole) antenna and  $P \propto (D/\lambda)^4$  for a circular loop (magnetic dipole) antenna, where  $l$  is the length of the linear antenna,  $D$  is the diameter of the loop antenna, and  $\lambda$  is the radiated wavelength. For  $l, D \ll \lambda$ , the impedance of the antenna is mostly reactive and most of the power is dissipated as ohmic losses in the antenna and the feed network. Therefore, introduction of superconducting radiating elements and feed networks can dramatically improve the radiation efficiency at low-frequency regions.
2. *Superdirective Antennas.* A superdirective antenna (13) has directional gain much larger than a conventional one. This can be of advantage in radio beacons and radar transmitters. A superdirective antenna is realized in practice by an array of closely spaced (separation  $\ll$  wavelength) dipole elements that are excited approximately  $180^\circ$  out of phase with respect to their neighbor. Such a structure has a low efficiency when built out of normal



**Figure 7.** A 16-element superconducting phase array antenna and the feed network. An  $\text{LaAlO}_3$  substrate is holding the HTS film. To the right of it is a 1 mm quartz plate holding the corresponding copper patches that provides an electromagnetic coupling to room temperature environment. A 0.5 mm vacuum between the two and low thermal conductivity spacers provide the required thermal isolation. The quartz plate has the necessary mechanical strength to withstand the atmospheric pressure and also serves as a radome. The antenna operates at 20 GHz. (Courtesy of Dr. Jeffrey Herd, US Air Force Research Laboratory.)

metal because ohmic losses in each individual element add up. In addition, these antennas have inherently low radiation efficiency because of cancellation of the radiation field, and use of HTS material can enhance efficiency.

3. *Millimeter-Wave Antennas and Feed Networks.* Superconducting antennas also improve the performance of antenna arrays and other distributed feed systems, where the power gain because of the distributed structure has to compete with the loss in an increasing number of elements. In a copper microstrip antenna the overall gain begins to decrease beyond about 40 elements, whereas the same antenna employing superconducting elements shows an increase in gain up to about 400 elements. This increase in gain can be crucial in mission critical applications such as military target tracking systems.

Figure 7 is a photograph of a 16-element phased-array antenna that was developed at the US Air Force Rome Laboratory.

#### Delay Lines

Requirements for a good delay line are low loss, low dispersion, and large delay for unit size/weight of the material. Superconducting delay lines are far superior to normal metal ones in satisfying these requirements, so much so that superconducting delay lines were in application even before the advent of high- $T_C$  superconductivity (14). Delays of the order of 20 ns can be routinely achieved with HTS materials on a substrate with an area of a few square inches, with good possibilities of achieving more than 100 ns in the near future. A copper coaxial line has to be several meters long for producing the same amount of time delay and would also have a high insertion loss.

A novel application of a delay line is the measurement of instantaneous frequency of a nonperiodic signal (15). This is achieved by splitting the input signal power equally between a series of  $n$  phase discriminating units, where  $n$  is the number of bit of frequency resolution desired. Each phase discriminating unit consists of a mixer, one port of which is fed by the signal directly and the other one by the same signal delayed.

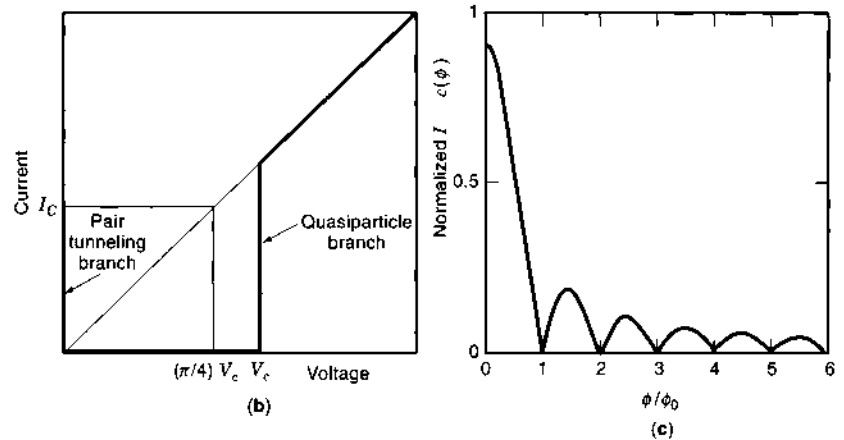
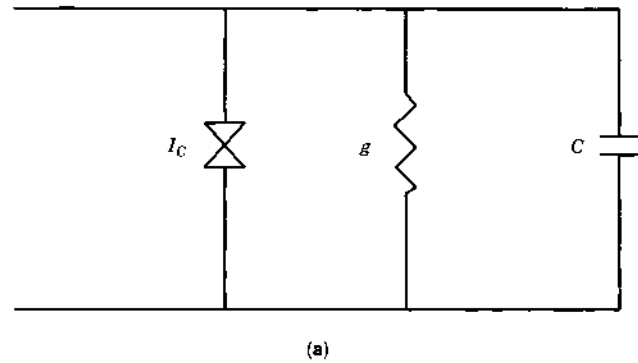
#### ACTIVE MICROWAVE CIRCUITS

In 1962 B. D. Josephson predicted that when two superconductors are separated by a thin layer of metal or an insulator, a

lossless current can flow up to a certain critical value  $I_C$  before a voltage appears across the junction. This is the manifestation of the phenomena that the Cooper pairs that carry lossless electric current can tunnel through the potential barrier represented by the junction material. Josephson's remarkable prediction was verified experimentally the following year and is known as the Josephson effect. The Josephson effect opened the way for a number of new applications of superconductors which exploit the fact that superconductivity is actually an amazing manifestation of quantum mechanics on a macroscopic scale. The most noteworthy of these is the superconducting quantum interference device (SQUID) magnetometer. With sensitivities approaching  $10^{-15}$  T/Hz $^{1/2}$  values, SQUIDs can measure magnetic field with precision that is unimaginable with any conventional techniques. The interested reader is referred to some useful reference for further details (see, for example, Ref. 16). In the following, we shall describe some of the microwave applications of the Josephson effect that have the potential to radically alter the future of microelectronics.

The Josephson current  $I$  is related to the phase difference  $\theta$  between the Cooper pair on the two sides of the junction as  $I = I_C \sin \theta$ . When current across the junction exceeds the critical value, a voltage  $V$  appears across it that is related to time rate of change of  $\theta$  as  $\partial \theta / \partial t = 4\pi e V / h$  (i.e., the phases of Cooper pairs on the two sides of the junction begin to "slip" relative to each other.) Figure 8(a) shows a lumped circuit equivalent of a Josephson junction (JJ), called a resistively shunted junction (RSJ model). The conductance and the capacitance represent the resistive and displacement current flow across the junction. The static characteristics of the circuit is given in terms of Stewart-McCumber parameter (17,18)  $\beta = 4\pi e I_C C / h g^2$ .

Elementary mathematical analysis of Fig. 8(b) shows that the time-domain response of the circuit can be written as  $I/I_C = \beta d^2 \theta / d\phi^2 + d\theta / d\phi + \sin \theta$ , where  $\phi = 4\pi e I_C t / h g$ ,  $t$  being time. This is equivalent to the equation describing the motion of a simple pendulum in a gravitational field. Thus the Josephson current behaves like a damped harmonic oscillator for small currents which is underdamped for  $\beta < 1/4$  and overdamped for  $\beta > 1/4$ . This also implies a hysteretic  $I - V$  behavior in the a region  $I_{\min} < I < I_C$ , and the value of  $I_{\min}$  must be obtained by solving the above equation. The value of  $I_{\min}$  decreases monotonically with increasing  $\beta$ . Another useful property of a JJ is the flux dependence of the critical current. If the JJ has a flux



**Figure 8.** (a) Lumped equivalent of a superconductor-insulator-superconductor (SIS) Josephson junction. (b)  $I$ - $V$  characteristics of an SIS Josephson junction. (c) Dependence of junction critical current on the threaded flux.

$\Phi$  threaded to it, then the critical current is given by  $I_{C\Phi} = I_C \sin(\pi\Phi/\Phi_0)/(\pi\Phi/\Phi_0)$ , where  $\Phi_0 = h/2e = 2 \times 10^{-15}$  Wb is the “flux quantum.” (The flux threaded through any superconducting ring must be an integral multiple of  $\Phi_0$ .) This is shown in Fig. 8(c).

To date, the most promising applications of superconducting active devices have been demonstrated only in low-temperature superconductors, most notably lead, niobium, and niobium nitride with aluminum oxide ( $\text{AlO}_x$ ), magnesium oxide ( $\text{MgO}$ ), and lead oxide ( $\text{PbO}$ ) as the insulating barriers. For a Cooper pair to tunnel coherently across a junction, its thickness must be comparable to the coherence length of the material, which is extremely small for high- $T_C$  superconductors (see Table 1). The critical current in these materials is also rather high (of the order of milliamperes), which is required for the Josephson coupling energy  $\hbar I_C/4\pi e$  to overcome the thermal energy  $k_B T$ , but it produces excessive Joule heating in the resistive shunts and high voltages across inductances. There are additional factors arising out of complex crystal structure and presence of “weak links” (i.e., small-angle grain boundaries) that make fabrication of JJs difficult in these materials. There are several possible analog and digital applications of the Josephson effect in microwave frequency regime, some of which are outlined in the following.

### Superconducting Digital Logic Circuits

From the early days of its discovery, the Josephson junction (JJ) has been eyed as a potential replacement of semiconducting logic gates (a detailed review is provided by Ref. 19). Josephson junctions are inherently bistable (with zero and finite

resistivity). However, the real driving force behind a “superconducting supercomputer” comes from three different sources:

- Josephson junctions can be switched *much* faster than a CMOS logic gate, where the parasitic and junction capacitance limits the minimum switching time. For a JJ, however, the theoretical limit on switching time is  $\hbar/2\pi\Delta$ , where  $\hbar$  is Planck’s constant and  $\Delta$  is the superconducting energy gap. For niobium, this corresponds to 0.22 ps, and practical circuits with switching times of 1.5 ps have been fabricated. The large difference between the theoretical upper limit in switching speed and those practically achieved stems from parasitic capacitance in the junction. However, there are no space-charge effects in a JJ, and hence these capacitances are much smaller compared to semiconductor circuits.
- Average power dissipation per gate in a JJ is at least two orders of magnitude lower than equivalent semiconductor gates. This means that the gates can be packed closer together, thereby reducing the propagation delay of the signal, another constraint in high-speed digital circuits. As an example, a four-bit microcontroller fabricated out of JJ and clocked at 770 MHz dissipated 5 mW, in contrast to a replica made out of gallium arsenide (GaAs) and clocked at 72 MHz that dissipated 2.2 W.
- Digital circuits operating over  $\sim 100$  GHz must use superconducting interconnects, because the inherent dispersion and loss in metal interconnects will degrade the signal sufficiently to make the circuit inoperable.



Owing to the hysteretic  $I$ - $V$  characteristics of the JJ, switching between zero resistance and finite resistance states (logic 0 and logic 1 in our convention) cannot be achieved as fast as it would be in a nonhysteretic device. If the current is reduced slightly below the value at which a logic 0 to logic 1 switching takes place, the circuit will remain in logic 1 state. Owing to this property, such circuits are called "latching circuits." The maximum clocking speed of these circuits cannot exceed a few gigahertz. To solve this problem, JJs are shunted with resistors that make them nonhysteretic. These circuits, called rapid single flux quantum (RSFQ) devices, are the basic building block of superconducting logic circuits. They differ from "conventional" logic in a fundamental way: The logic state is not decided by the voltage level of the gate but by the presence or absence of voltage pulse generated by the motion of single fluxons. Practical superconducting digital circuits with significantly higher performance have been demonstrated. Using niobium technology, a 4 bit microprocessor has been fabricated by Fujitsu and a 1 kbit random access memory (RAM) chip has been made by NEC (for a good review of recent progress in Josephson IC fabrication, see Ref. 20). Many other leading manufacturers are also pursuing this technology (21).

There are a few rather unusual problems in the practical realization of a superconducting computing device. It is impractical to realize a high-speed computer in a Von Neumann-type architecture that involves massive data transfer between a central processor and the memory, which are physically separated over a relatively long distance. An order-of-magnitude estimate of the maximum allowable data path (assuming microstrip interconnects on a substrate with dielectric constant  $\sim 20$ ) of a computer operating at 300 GHz is  $\sim 0.3$  mm. Clearly, this is very difficult to realize on a circuit board. However, applications such as dedicated digital signal processor with on-board cache memory and multichip modules can have phenomenal speed and performance increase if built using RSFQ logic.

### Detectors and Mixers

The nonlinearity of a Josephson junction can be exploited to make a mixing device. In addition to the low noise and high efficiency that these devices offer, they can easily be integrated to an all-superconductor radio-frequency (RF) receiver front end.

The general principles of mixing due to photon-assisted tunneling was derived and applied to superconductor-insulator-superconductor (SIS) junctions by Tucker (22,23). There can be two types of mixers using Josephson junctions: those using Cooper pair tunneling operating near zero bias voltage and those using quasiparticle tunneling operating near gap voltage. Quasiparticle mixers are usually preferable over Cooper pair mixers for several reasons. For one, the junction capacitance has to be small so that most of the current comes from Cooper pair tunneling and not displacement electric field. This means the use of point contact junctions which are difficult to fabricate reproducibly. Secondly, they are noisy because of harmonic mixing of all frequencies up to the gap frequency. According to the quantum theory of mixing, strong nonlinearity in  $I$ - $V$  characteristics (more precisely, when  $I_{dc}(V_{dc} + hv/e) - I(V_{dc}) \gg I_{dc}(V_{dc}) - I_{dc}(V_{dc} - hv/e)$ ) can result in conversion gain exceeding unity and conversion efficiency approaching the quantum limit even in a purely resistive mixer. Thirdly, the shot noise in a quasiparticle mixer is lower than that in Cooper pair

tunneling mixers. SIS quasiparticle mixers have been used in radio astronomy for quite a while.

### Analog-to-Digital Converters

Analog-to-digital converters (ADCs) are another type of circuit where phenomenal performance improvement has been demonstrated using superconducting circuitry. While superconducting ADCs can be and have been fabricated in most architectures and have the usual advantage of a large bandwidth because of their fast switching capability, there is the additional advantage that when exploiting the multithreshold characteristics of a JJ [see Fig. 8(c)], an  $n$ -bit flash ADC requires only  $n$  comparators as opposed to  $2^n - 1$  that would be normally required (24). This allows high-bandwidth and high-resolution ADCs to be fabricated reliably and compactly.

### Precision Voltage and Frequency Sources

The Josephson relation  $\partial\theta/\partial t = 4\pi eV/h$  can be used both as a source of high-frequency radiation and a dc voltage standard. The power radiated by the oscillating Josephson current in response to a small dc bias (483.5 GHz/mV) is usually too very small (a few nanowatts) to be of much practical use, but coherent Josephson arrays have been fabricated that can output as much as a few microwatts, and power levels of up to 1 W has been predicted. On the other hand, when irradiated with microwaves, a JJ develops a dc voltage across it. Since frequency of such radiation can be accurately controlled and measured, such devices are one of the accepted precision voltage standards today.

**The Future of Superconducting Microwave Electronics.** How are these myriad of possible superconducting circuits realized in practice? Just like any other microelectronic circuit, they are lithographically patterned on a substrate through a similar sequence of steps as in a semiconductor (25). Thus, the vast assortment of techniques developed during the last several decades for and by the semiconductor industry can be ported into commercialization of superconducting electronics. For a complete self-contained system, many of the subsystems have to be built out of semiconductor devices [e.g., the (IF) amplifiers for an RF transceiver]. An efficient way of manufacturing these systems is to mix HTS and high electron mobility transistor (HEMT) semiconductor circuits on the same substrate as a multichip module (MCM) (26). Typically the HTS film is grown on a GaAs substrate with a thin buffer layer for better lattice match. Semiconductor devices perform better at lower temperatures because of enhanced carrier mobility, so the overall performance increases. Decrease of thermal noise is another desirable byproduct. Also, with the increase in the packing density of the semiconductor circuitry, a decrease in operating temperature, and the advent of high-temperature superconductors, the disparity between the operating voltage levels associated with the two types of circuits is going down, minimizing the possibility of ground loops between these two types of elements. Any large-scale commercial application of HTS microwave circuit will rely on the ability to integrate them, either as a hybrid component or as a monolithic component, with active semiconductor components.

There are many hurdles in the road to success of superconducting devices. The absence of a room-temperature superconductor, essential for consumer market, is one of them.

Even if there were a room-temperature superconductor, the vast amount of techniques developed by the semiconductor industry for precise and predictable control of material properties are unavailable in the superconductor industry. The reason for this is obvious: The whole phenomenon of high-temperature superconductivity is only understood phenomenologically at best. There is no equivalent of "bandgap engineering" for superconductors. If we knew the material to dope to change the superconducting bandgap, we could then produce a room-temperature superconductor!

However, consumer electronics, though a large share of the market, still is not all of the market. There are niche markets, mostly in the defense sector and commercial/military wireless communication, who would pay the extra dollar to have the advantages of superconducting electronics (see, for example, Ref. 27). For example, the low insertion loss and steep skirt of a superconducting filter makes it cost effective in a cellular base station where the receiving/transmitting antenna must operate within tight bandwidth tolerance and divide up the available bandwidths among as many customers as possible. Such filters are available from many vendors and several are being field-tested by cellular service providers. One example of a large-scale attempt to use HTSC microwave circuits in communication application is the United States Naval Research Laboratory's "High Temperature Superconductivity Space Experiments" (HTSSE) (28). The US Air Force is another major player in this field and expects to utilize low-loss superconducting antennas in the next-generation radar systems. System integration is a crucial aspect toward commercialization, and significant progress has been made to this end. Integration of individual superconducting microwave components to produce a complete self-contained RF receiver/transmitter front end has been successfully demonstrated by many researchers. Integration of superconducting and semiconducting electronics on the same device has also been carried out.

At the present rate of progress, we can safely say that the growth of superconducting electronics will steadily increase with time. Whether it would radically alter the present state of the art, similar to the effect the advent of solid-state circuits had over the vacuum tube technology, will be seen in the years to come.

#### ACKNOWLEDGMENTS

This work was supported by NSF-ECS-9711910. I am grateful to the following people for providing me with figures and associated data: Dr. Daniel Oates of MIT Lincoln Laboratory for Fig. 5, Dr. Balam Willemsen of Superconductor Technologies Inc. for Fig. 6, and Dr. Jeffrey Herd of US Air Force Research Laboratory for Fig. 7.

#### SUGGESTIONS FOR FURTHER READING

The literature of microwave applications of superconductivity is quite extensive, and citations relevant to a particular subtopic have already been given wherever appropriate. A nonexhaustive list of periodicals and monographs of general interest in this area is provided below:

- *IEEE Transactions on Microwave Theory and Techniques* (Periodical).

- *IEEE Transactions on Applied Superconductivity* (Periodical).
- M. J. Lancaster, *Passive Microwave Device Applications of High-Temperature Superconductors*, Cambridge: Cambridge University Press, 1997.
- Zhi-Yuan Shen, *High-Temperature Superconducting Microwave Circuits*, Norwood, MA: Artech House, 1994.
- S. T. Ruggiero and D. A. Rudman (eds.), *Superconducting Devices*, New York: Academic Press, 1990.
- R. D. Parks (eds.), *Superconductivity* (in two volumes), New York: Marcel Dekker, 1969.
- T. Van Duzer and C. W. Turner, *Principles of Superconductive Devices and Circuits*, Amsterdam: Elsevier, 1981.
- Harold Weinstock and Martin Nisenoff (eds.), *Superconducting Electronics*, NATO Advanced Study Institute Series, New York: Springer-Verlag, 1989.

#### BIBLIOGRAPHY

1. McConnell, Wolf, and Noufi (eds.), *Science and Technology of Thin Film Superconductors*, Vols. 1–2, New York: Plenum, 1988, 1990.
2. J. D. Jackson, *Classical Electrodynamics*, 3rd ed. New York: Wiley, 1999.
3. M. J. Lancaster, *Passive Microwave Applications of High-Temperature Superconductors*, Cambridge: Cambridge Univ. Press, 1997.
4. B. A. Willemsen, J. S. Derov, and S. Sridhar, Non-linear response of suspended high temperature superconducting microwave resonators, *IEEE Trans. Appl. Supercond.*, **5**: 1753–1755, 1995.
5. *IEEE Trans. Microw. Theory Tech.*, **44** (7): 1996.
6. R. Chatterjee, *Elements of Microwave Engineering*, Ellis Horwood Series on Electrical and Electronic Engineering, Chichester, UK: Ellis Horwood, 1986.
7. R. A. Chipman, *Transmission Lines*, Schaum Outline Series, New York: McGraw-Hill, 1968.
8. C. Zahopoulos, W. L. Kennedy, and S. Sridhar, Performance of a fully superconducting microwave cavity made of the high  $T_C$  superconductor  $Y_1Ba_2Cu_3O_y$ . *Appl. Phys. Lett.*, **52**: 2168, 1988.
9. S. Sridhar and W. L. Kennedy, Novel technique to measure the microwave response of high  $T_C$  superconductors between 4.2 and 200 K. *Rev. Sci. Instrum.*, **59**: 531, 1988.
10. G. L. Hey-Shipton et al., *High temperature superconductor lumped element band-reject filters*, U.S. Patent No. 5,616,539, 1997.
11. S. Das, U.S. Patent No. 5,496,795, 1996.
12. R. J. Dinger, D. R. Bowling, and A. M. Martin, A survey of possible passive antenna applications of high-temperature superconductors, *IEEE Trans. Microw. Theory Tech.*, **39**: 1498–1507, 1991.
13. R. W. Conrad, U.S. Patent No. H000653, 1989.
14. J. T. Lynch et al., U.S. Patent No. 4,499,441, 1985.
15. Guo-chun Liang et al., Superconductive digital instantaneous frequency measurement subsystem, *IEEE Trans. Microw. Theory Tech.*, **41**: 2368, 1993.
16. J. C. Gallop, *SQUIDS, the Josephson Effects and Superconducting Electronics*, Adam Hilger Series on Measurement Science and Technology, Philadelphia, PA: Adam Hilger, 1991.
17. W. C. Stewart, Current-voltage characteristics of Josephson junctions, *Appl. Phys. Lett.*, **12**: 277–280, 1968.
18. D. E. McCumber, Effect of ac impedance on dc voltage-current characteristics of superconductor weak link junctions, *J. Appl. Phys.*, **39**: 3113–3118, 1968.

19. K. K. Likharev and V. K. Semenov, RSFQ logic/memory family: A new Josephson-junction technology for sub-terahertz-clock-frequency digital systems, *IEEE Trans. Appl. Supercond.*, **1**: 3, 1991.
20. K. Hara (ed.), *Superconductivity Electronics*, Ohmsha, Japan: Prentice-Hall, 1988.
21. Oleg A. Mukhanov, *Three-part Josephson memory cell for superconducting digital computer*, U.S. Patent No. 5,365,476, 1994.
22. J. R. Tucker, Quantum limited detection in tunnel junction mixers, *IEEE J. Quantum Electron.*, **QE-15**: 1234–1258, 1979.
23. Predicted conversion gain in superconductor–insulator–superconductor quasiparticle mixer, *Appl. Phys. Lett.*, **36**: 477–479, 1980.
24. P. D. Bradley, *Flash analog-to-digital converter employing Josephson junctions*, U.S. Patent No. 5,400,026, 1995.
25. Q. Ma and W. N. Hardy, *Superconductor logic and switching circuits*, U.S. Patent No. 5,345,114, 1994.
26. A. D. Smith and A. H. Silver, *Integrated superconductive heterodyne receiver*, U.S. Patent No. 5,493,719, 1996.
27. F. W. Patten and S. A. Wolf, The ARPA high temperature superconductor program, *IEEE Trans. Appl. Supercond.*, **5**: 3203, 1995.
28. M. Nisenoff et al., The high-temperature superconductivity space experiments: HTSSE I components and HTSSE II subsystems and devices, *IEEE Trans. Appl. Supercond.*, **3**: 2885–2890, 1993.

DURGA P. CHOUDHURY  
 Centre de Recherches en Physique  
 des Plasmas  
 S. SRIDHAR  
 Northeastern University

## SUPERCONDUCTIVITY: TUNNELING AND JOSEPHSON JUNCTIONS

Tunneling devices incorporating superconducting materials can exhibit what are known as Josephson effects. A typical example is two superconducting materials separated by a thin dielectric barrier. Remarkably, such a tunneling device can exhibit both a zero-voltage trace for currents below some critical value and a second trace that has tunneling characteristics akin to a forward-biased diode. Traversing the entire current–voltage characteristic of the device will give a hysteretic current–voltage trace. The scale of these phenomena is roughly 1 mV for low-temperature superconductors (such as Pb, Pb-alloys, and Nb) and 10 mV for high-temperature superconductors (such as YBaCuO, BiSrCaCuO, and TlBaCaCuO).

Low-temperature superconductors used for Josephson applications have transition temperatures,  $T_c$ , from the normal to the superconducting state of approximately 1 K to 20 K, while high-temperature superconductors have  $T_c \lesssim 90$  K.

The zero-voltage leg of the current–voltage characteristics of a Josephson junction is highly sensitive to magnetic fields. This property is exploited by superconducting quantum interference devices (SQUIDs) to measure magnetic fields down to the 10 fT range, making them the most sensitive detectors of field/flux in existence. Superconducting quantum interference devices have been employed in prototype diagnostic systems for noninvasive mapping of heart and brain function, but must typically be operated within the confines of a magnetically shielded room. Prototype devices using SQUIDs for nondestructive evaluation of metal components are currently being tested in the

field. SQUIDs are also employed in the detection of undersea anomalies (submarine and ordinance detection) and in a variety of scientific investigations such as earthquake, gravity-wave, free-quark, monopole, and dark-matter detection. Some practical SQUID devices have also used high-temperature superconductors.

The nonzero voltage leg of a Josephson junction current–voltage characteristic is also useful because it can be highly nonlinear. As such, these devices are employed for mixing in the 100 GHz region, where low noise-temperatures are a must (such as in radio astronomy). This work is restricted to low temperatures—on the order of 1 K—and the use of low-temperature superconductors. High-temperature–superconductor tunneling characteristics are not currently of sufficiently high quality for this application.

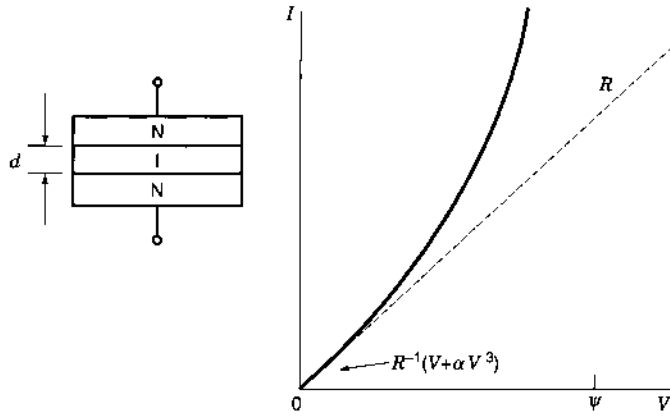
Another important potential application of Josephson junctions is their use in digital applications. The reason is that switching with Josephson-junction-based logic involves small voltages and currents and can occur over very short time scales. Intrinsic power and switching speed are in the 0.1  $\mu$ W and 1 ps range. Josephson junctions employed for this purpose, based on Nb, are a mature technology. For a number of reasons, current technology has focused on ultrahigh-speed analog/digital conversion rather than computation; one issue is the current lack of a suitable memory format compatible with Josephson technology. Nevertheless, high-speed Josephson processing circuitry continues to have promise for such targeted applications as video image processing. A number of books and edited works are available on the subject of the Josephson effect and its applications (1–8).

Ironically, for digital and SQUID applications the intrinsic hysteresis of the junctions must be suppressed. For junctions that employ a thin, insulating oxide barrier, this implies the use of thin-film resistors to shunt the junctions to reduce the hysteresis. A better solution would be to tailor the tunnel barrier material so that it was somewhere between a metal and insulator. More exactly, this suggests producing a material close to the metal–insulator transition. This subject is discussed at further length in connection with research in producing high-temperature superconductor junctions.

## METAL–INSULATOR–METAL TUNNELING

A basic tunneling system consists of a metal–insulator–metal (MIM) structure as depicted in Fig. 1. Generally the system comprises thin metal films  $\sim 10$  nm to 100 nm in thickness, separated by a thin dielectric layer  $\lesssim 1$  nm in thickness. This dielectric can either be a native oxide created by the oxidation of the base metal electrode (a good example of which is aluminum) or a deposited artificial barrier, a good example of which is a semiconductor such as silicon. Tunneling in native and artificial barriers has been reviewed as a separate topic (9).

If the metals are not in the superconducting state, we have normal–insulator–normal (NIN) tunneling (10). The current–voltage characteristics of NIN junctions have been extensively studied in the context of quantum-mechanical electron tunneling through a potential barrier. The height of the barrier,  $\psi$ , governs the rate of change of current with applied voltage. Barrier heights typically range from 0.1 eV to 3 eV. Simmons (11)



**Figure 1.** Shown is a sketch of a normal-insulator-normal, NIN, tunneling system. N represents a normal, or nonsuperconducting metal film. I (nominally an insulator) represents a thin, dielectric layer through which quantum-mechanical tunneling can occur. The normal-metal films are typically approximately 10 nm to 100 nm thick and the dielectric layer is approximately 1 nm thick. Barrier heights,  $\psi$ , are typically in the range of 0.1 eV to 3 eV. At low voltages  $I \sim R^{-1}(V + \alpha V^3)$  with  $\alpha \sim 1 V^{-2}$ .

has shown that for applied voltages  $V \ll \psi$

$$I = R^{-1}(V + \alpha V^3)$$

where  $\alpha \sim 1 V^{-2}$ .

At a given voltage, the current varies with barrier thickness,  $d$  and barrier height,  $\psi$  as

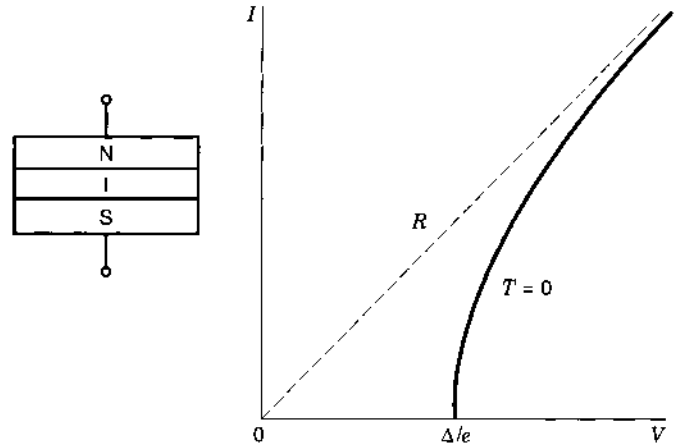
$$I \propto \frac{\psi^{1/2}}{d} e^{-10.25 d \psi^{1/2}}$$

where  $d$  is in nanometers and  $\psi$  is in volts (12). Tunneling in more complex although thematically similar semiconductor systems is discussed by Sze (13).

If we now introduce a superconductor (Fig. 2) as one of the metal elements of the tunnel junction, to produce a superconductor-insulator-normal (SIN) system, the current-voltage characteristic becomes highly modified. Most notable is the introduction of a region in which, at zero temperature, no current flows until a voltage  $\Delta/e$  is reached, where  $\Delta$  is the so called energy gap of the superconducting film. Energy gaps associated with conventional superconductors such as Pb, Pb alloys, Nb, NbTi, and NbSn are in the few meV (1 meV =  $10^{-3}$  eV) range, while energy gaps associated with the high-temperature superconductor materials, such as YBaCuO, are in the  $\gtrsim 30$  meV range.

If we introduce a second superconductor to make an superconductor-insulator-superconductor (SIS) system, something unique occurs (Fig. 3). In this case, two branches develop in the system. If we start at zero current and increase the current slightly, no voltage will develop across the junction until the critical current,  $I_c$ , is reached. The critical current is related to the energy gaps of the superconductors by  $I_c = \pi/4 \cdot (\Delta_1 + \Delta_2)/eR$ , at zero temperature. For example, for Pb and Nb,  $\Delta = 1.37$  and 1.53 meV, respectively.

If the critical current is exceeded, a junction driven by a current source will jump (at constant current) over to the quasi-particle curve. Further increasing the current moves one to



**Figure 2.** Shown is a sketch of a superconductor-insulator-normal, SIN, tunneling system. N represents a normal, or nonsuperconducting metal film and S represents a superconducting film. I (nominally an insulator) represents a thin, dielectric layer through which quantum-mechanical tunneling can occur. The normal-metal films are typically approximately 10 nm to 100 nm thick and the dielectric layer is approximately 1 nm thick. The current-voltage trace shown is for zero temperature ( $T = 0$ ).

higher voltages on the quasi-particle curve and reducing the current explores the rest of this curve. The zero-voltage state is not recovered until the current is reduced to zero. If a finite impedance source drives the junction, this switching will occur along the load-line of the system. This remarkable situation is due the fact that coherent tunneling of Cooper pairs of electrons (paired electrons in the superconducting state) can occur, producing a zero-voltage superconducting tunneling state in the system. When the critical current is exceeded, these pairs are broken into quasi particles with properties akin to electron-hole pairs created across a semiconductor energy gap (13).

The current can be described by the Josephson equations (14-16)

$$I = I_c \sin \phi \frac{d\phi}{dt} = \frac{2e}{\hbar} V$$

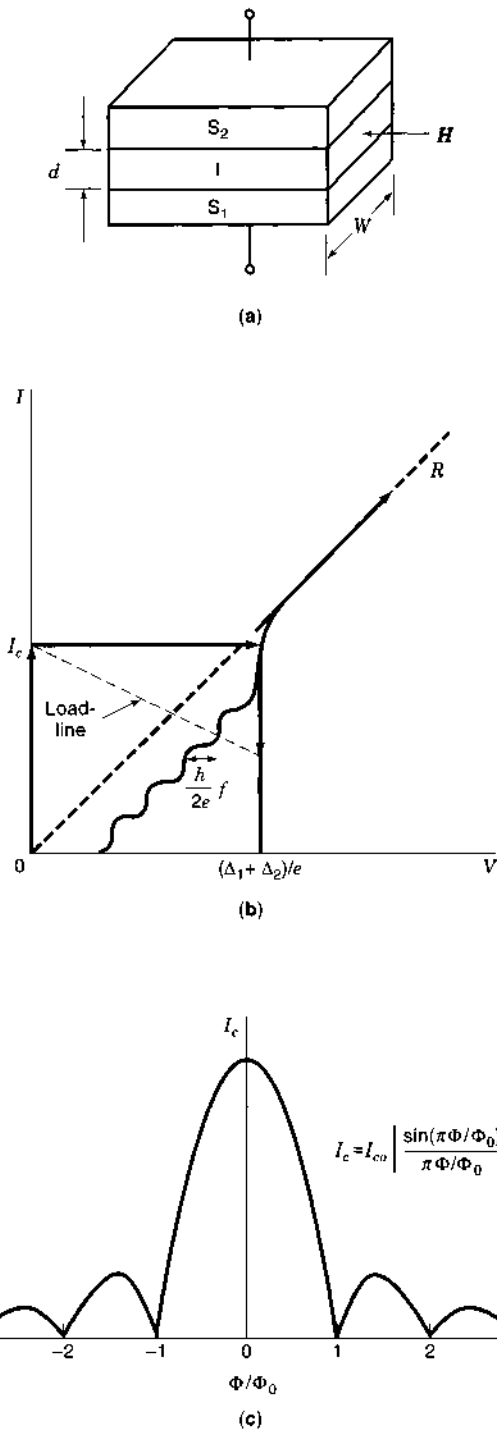
where  $\phi$  is the phase difference between the superconductors. Therefore, increasing the applied current from zero is equivalent to introducing a quantum-mechanical phase difference between the superconducting elements of the junction.

Beside SIS systems, properly configured SNS systems can also exhibit Josephson effects, as long as the normal-metal channel between the superconducting materials is of the appropriate geometry. This means that the normal metal is either a microscopic weak-link connection, point contact, or a thin-film metal constriction as opposed to a dielectric material (4, 17).

### ALTERNATING CURRENT JOSEPHSON EFFECT

A nonzero dc voltage,  $V = \text{constant}$ , across a Josephson junction implies that

$$\phi = \phi_0 + \frac{2e}{\hbar} Vt$$



**Figure 3.** Shown is a sketch of a superconductor–insulator–superconductor, SIS, tunneling system. S represents a superconducting film. I (nominally an insulator) represents a thin, dielectric layer through which quantum-mechanical tunneling can occur. The normal-metal films are typically approximately 10 nm to 100 nm thick and the dielectric layer is approximately 1 nm thick. Increasing the current applied across the device produces no voltage until the critical current,  $I_c$  is reached, at which point increasing or decreasing current drives the system along the nonzero-voltage, quasiparticle current–voltage characteristic. The load-line of a system driven by a nonzero impedance source is also indicated. Applying a magnetic field to a sufficiently small junction (see text) will produce a periodic depression of the critical current.  $\Phi$  is the magnetic flux entering the junction.  $\Phi = HW(d + \lambda_1 + \lambda_2)$ , where  $\lambda$  is the penetration depth and  $\Phi_0 = h/2e = 2.068 \times 10^{-15}$  Wb.

which means that

$$I = I_c \sin(\phi_0 + \frac{2e}{h} Vt)$$

Therefore, a finite voltage across a Josephson junction gives rise to a radio frequency (RF) current of

$$f = \frac{2e}{h} V$$

where  $2e/h = 483.598$  THz  $V^{-1}$ . The presence of an alternating current associated with the appearance of a dc voltage across the junction is the ac Josephson effect. This means that at finite voltage, the current comprises a dc component with a superimposed ac modulation at the Josephson frequency. A current–voltage characteristic thus represents the time-averaged voltage across the junction as a function of applied current (when the system is driven by a finite-impedance source). The junction thus acts like an RF frequency-to-voltage converter.

Josephson junctions are in fact actually used as high-frequency sources. For a given junction, roughly 10 nW of power can be produced, although most this is dissipated in the junction resistance. The linewidth of the radiation is given by

$$\Delta f = \frac{4k_B T R}{\Phi_0^2}$$

where  $\Phi_0 = h/2e = 2.068 \times 10^{-15}$  Wb is the magnetic flux quantum. For  $T \sim 1$  K and  $R \sim 1 \Omega$ , this means that  $\Delta f \sim 1$  MHz. The linewidth can be reduced by coupling to a high- $Q$  cavity, although with the concomitant loss of overall operational bandwidth.

To achieve useful output power levels, arrays of junctions are produced to create tunable millimeter wave sources—typically as low-noise local oscillators for radio astronomy mixer applications. Junction arrays can produce  $\sim 2 \mu\text{W}$  to  $6 \mu\text{W}$  of RF power in the 340 GHz to 440 GHz band. To achieve this, all junctions in the array must be phase locked (18).

The ac Josephson effect can be manifested as Shapiro steps (Fig. 3) by coupling microwave radiation into the junction. This produces a series of steps of equal width  $\Delta V = (h/2e)$  where  $f$  is the frequency of the applied microwaves. An important application that exploits this phenomenon, and also uses series arrays of (here Nb-based) junctions, is the standard volt. Here, a junction array is driven by a microwave source to produce in excess of 1 V dc on the output terminals of the device (19, 20). Only fundamental constants and the frequency of an external oscillator, which can be established with high accuracy, determine the output voltage. The National Institute of Standards and Technology (NIST) has developed Josephson arrays for this purpose and they now serve as the primary US standard volt.

Recent work in this particular area has also included Nb/AuPd/Nb (SNS-type) junctions. These systems use a normal-metal alloy instead of a dielectric barrier, through which Josephson coupling can also occur to produce a junction that is inherently resistively shunted (21). Another refinement is to use Ti (22) instead of AuPd. Titanium is potentially more desirable, because its resistivity (at 4.2 K) is higher than AuPd and it is compatible with whole-wafer processing techniques.

Experimental work based on a stacks of Josephson junctions (23) has also been explored for potential submillimeter oscillator applications with, for example, NbCN/MgO/NbCN (24), NbCN/NgO/NbCN (25), and Nb/Al/AlO<sub>x</sub>/Nb (26, 27) systems. Stacks of Josephson junctions may also be useful as inductive elements in resistive or rapid single-flux-quantum (RSFQ) digital circuitry. One technical challenge here is producing junctions with uniform Josephson critical currents,  $I_c$  (28).

Microwave irradiation has also been explored on a more purely experimental basis with high-temperature superconductors. These materials can be viewed as a stack of superconducting CuO planes, with an interplanar quantum-mechanical coupling that can vary from one material to another. For example, the layer-to-layer coupling in BaSrCaCuO is far less than in YBaCuO. In fact, in the former case the coupling is similar to the Josephson coupling occurring in an SIS tunnel junction. Thus BaSrCaCuO is thought capable of mimicking the behavior of a stack of individual Josephson junctions. With this picture in mind, researchers have applied microwave radiation to BaSrCaCuO mesas. Microwave steps were in fact observed, but with a voltage spacing greater than that expected for Shapiro steps (29). Mesa-type TlBaCaCuO devices have also been produced (30) with microwave applications in mind. Again, however, these systems remain experimental in nature.

#### MAGNETIC-FIELD EFFECTS

The current distribution in a Josephson junction will be uniform if the width,  $W$ , of the junction (see Fig. 3) is small compared with the Josephson penetration depth,  $\lambda_j$ , with

$$\lambda_j = \left\{ \frac{\hbar}{2e\mu t J_c} \right\}$$

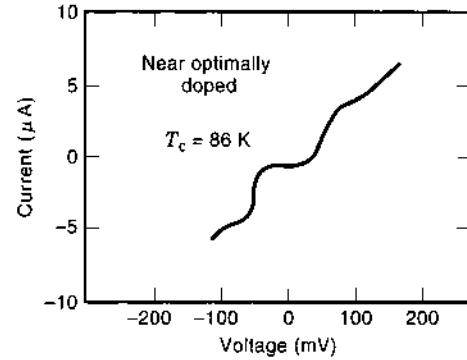
where  $J_c = I_c/A$  with  $A$  the junction area, and  $t = \lambda_1 + \lambda_2 + d$ . The quantities  $\lambda$  are the penetration depths of the superconductors composing the junction. This is the scale over which an applied external magnetic field will penetrate into a superconductor. Here  $d$  again is the thickness of the barrier (1–6, 31).

In the case where we apply a static magnetic field to the junction along the direction of the plane of the barrier, we will suppress the Josephson current. If we meet the criterion that the junction is small compared with the Josephson penetration depth, this suppression of the critical current will take the form

$$\frac{I_c}{I_{c0}} = \left| \frac{\sin \pi(\Phi/\Phi_0)}{\pi(\Phi/\Phi_0)} \right|$$

where  $I_{c0}$  is the zero-field critical current. For Pb, Nb, NbSn,  $\lambda \sim 50$  nm to 100 nm; for NbN,  $\lambda \sim 300$  nm; and for YBaCuO  $\lambda_{\parallel} \sim 30$  nm and  $\lambda_{\perp} \sim 200$  nm. In the last case, parallel and perpendicular refer to the directions along and perpendicular to the CuO planes.

If a junction is not small by this definition, the current distribution will not be uniform and the suppression of the critical current will not follow the simple  $\sin x/x$  behavior noted above. Indeed, the geometry of the junction can be controlled to produce a critical-current behavior parametrically tailored to specific applications, as in the case with the use of junctions as elements for digital applications. In addition, the spatial



**Figure 4.** Shown is the current–voltage characteristic of a nominally SIS junction created by using a scanning tunneling microscope tip to peel up a portion of the surface of a high-temperature superconducting BiSrCaCuO (2212) film and hold it in proximity to the underlying material. Although not an ideal characteristic, this demonstrates the ability to achieve (and provides a means to study in detail) tunneling in the high-temperature superconductor materials. From Miyakawa et al., Ref. 32.

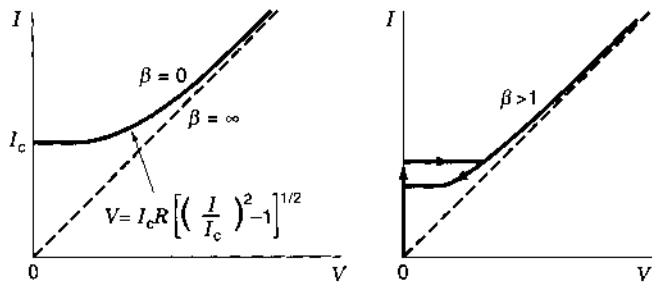
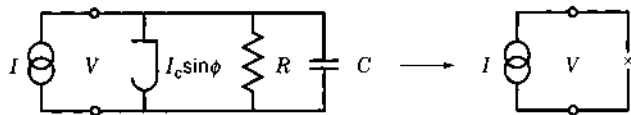
nonuniformity of the barriers of individual junctions can be diagnosed through a deconvolution of critical current versus applied magnetic field characteristics (3).

As far as applications are concerned, this general phenomenology implies that a large penetration depth is generally undesirable. A material such as NbN, which has a larger energy gap than Nb (and from this standpoint represents a superior material) suffers from this limitation. However, this can be in part engineered around by creating hybrid NbN/Nb layer pairs in which a balance of the higher critical temperature of NbN (14 K as opposed to 9.25 K for Nb) and lower penetration depth of Nb is reached. Another important example of the use of hybrid materials systems also involves Nb as Nb/Al bilayers; these are used to produce high-quality tunnel junctions as is discussed later (see section entitled “Digital Applications.”)

Finally we note that SIS junctions have been made with high-temperature superconductor materials as well. As shown in Fig. 4, such devices have been created with the use of a scanning tunneling microscope tip which is “crashed” into the surface of a BiSrCaCuO crystal (32). As the tip is pulled back up, a junction is created between superconducting material remaining on the tip and the underlying crystal. While hardly a practical device, with less-than-perfect current–voltage characteristics, this nonetheless illustrates the potential for creating devices from high-temperature superconductor material and has provided a valuable vehicle for fundamental studies of the system.

#### THE RESISTIVELY SHUNTED JUNCTION (RSJ) MODEL

The application of Josephson junctions generally requires control of the degree of junction hysteresis. One way of achieving this is to shunt the junction with an external resistor. This usually means the use of a thin film of Au or AuCu alloy deposited across the SIS junction to produce a shunt in the range of  $\sim 1 \Omega$ . In some cases, junctions can be produced with internal microscopic resistive links in the barrier layer to achieve the same effect. More sophisticated approaches, involving tuning the properties of tunnel barriers to achieve a conductivity near



**Figure 5.** The resistively shunted junction (RSJ) model for tunneling in SIS systems. This diagram illustrates that the degree of hysteresis in a junction with a given resistance and capacitance is governed by the magnitude of the parameter  $\beta = (2e/h)I_c R^2 C$ . The parameter can be viewed as setting the degree of damping in the circuit.

the metal-insulator transition, have also been proposed and are discussed further in "High-Temperature Superconductor Ramp Junctions."

Figure 5 shows the electrical equivalent circuit of a shunted Josephson junction with a total resistance  $R$  biased by a current source (shunted with an external resistor to produce a resistance far lower than the original tunneling resistance of the junction). Represented in this manner there are three current paths: the Josephson current, the ohmic current, and the displacement current from the junction capacitance. The total current (for a system driven by a current source) is thus

$$I = I_c \sin \phi + \frac{V}{R} + C \frac{dV}{dt}$$

Along with the Josephson relation

$$V = \frac{\hbar}{2e} \frac{d\phi}{dt}$$

this can be rewritten as

$$\frac{I}{I_c} = \sin \phi - \omega_c^{-1} \frac{d\phi}{dt} + \omega_p^{-2} \frac{d^2 \phi}{dt^2}$$

where  $\omega_c = (2e/h)I_c R$  represents the upper operational frequency of the junction and  $\omega_p = (2eI_c/C)^{1/2}$  is the "plasma" or lower-bound propagation frequency of the system. It is customary to define the Stewart-McCumber parameter

$$\beta = \left(\frac{\omega_c}{\omega_p}\right)^2 = \frac{2e}{\hbar} I_c R^2 C$$

The dependence of the current-voltage characteristics of a Josephson junction on  $\beta$  is also depicted in Fig. 5. Small and large values of  $\beta$  represent the high- and low-damping limits, respectively. The most desirable operating regime for many

applications is  $\beta \sim 1$ . As noted by Likharev (4), a useful way of parameterizing  $\beta$  is as

$$\beta = \frac{2e}{\hbar} (I_c R)^2 \frac{\epsilon_r \epsilon_0}{d} J_c^{-1}$$

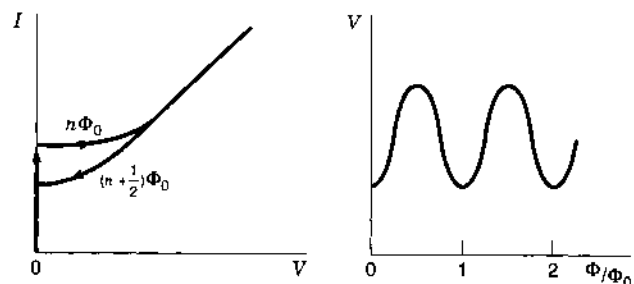
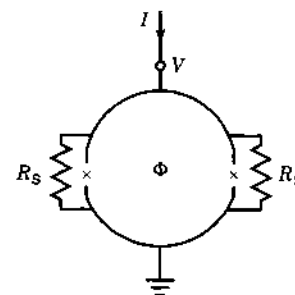
where again  $J_c = I_c/A$ ,  $A$  is the junction area ( $W^2$  in Fig. 3),  $\epsilon_r$  is the dielectric constant of the junction barrier material,  $\epsilon_0 = 8.85 \times 10^{-12}$  F/m, and  $d$  is the barrier thickness. Improved performance at high frequencies generally implies smaller areas and a concomitant increase in  $J_c$  to achieve  $\beta \sim 1$  (33). Note that in principle

$$I_c R = \frac{\pi}{4} \frac{\Delta_1 + \Delta_2}{e}$$

is an intrinsic property of the superconductors. Its actual value, however, can be lower than the indicated theoretical result due to gap suppression at film surfaces and other effects.

**SUPERCONDUCTING QUANTUM INTERFERENCE DEVICES**

If we place two Josephson junctions in a loop (Fig. 6) we form a SQUID, in this case a so-called dc SQUID. To produce a device intended for the measurement of flux, the junctions are shunted, as discussed previously, to suppress junction hysteresis. The flux associated with an applied magnetic field generates a circulating current in the device, which suppresses the Josephson current. An applied flux thus modulates the voltage across the device with a period equal to the magnetic flux quantum  $\Phi_0$ . The result of this is a device that is very sensitive to magnetic flux. SQUIDS are sensitive to flux to magnitudes  $\Phi_{\text{applied}} \ll \Phi_0 = h/2e = 2.068 \times 10^{-15}$  Wb. Practical SQUIDS are



**Figure 6.** Shown is a sketch of a dc SQUID (superconducting quantum interference device) comprising a loop with two Josephson junctions. Shunt resistances are used to control the damping parameter ( $\beta$ ) of the junctions. The current-voltage characteristics and voltage-flux characteristics are shown for an applied flux  $\Phi$ .

capable of measuring fields to the 10 fT magnetic field range, or  $\sim 10^{-34}$  J·Hz $^{-1}$  at 10 $^5$  Hz (34, 35), making them by far the most sensitive device in existence for measuring magnetic fields and flux. Field versions of SQUIDs can usually achieve an order of magnitude more sensitivity than flux gate magnetometers.

SQUIDs have a variety of applications, including commercial biomedical SQUID arrays that noninvasively monitor and map heart and brain function. Other uses include nondestructive evaluation (NDE) (e.g., crack detection in metals), oceanic anomaly detection, gravity-wave antennas, earthquake monitoring, magnetic monopole detection, dark-matter searches, and other fundamental scientific investigations (6, 36, 37).

A key factor limiting the performance of SQUIDs is  $1/f$  noise. This stems from shallow trapping of both fluxoids in the thin-film elements of the SQUID and tunneling electrons in the barriers of the Josephson junctions (shot noise). The observed  $1/f$  noise typically has a knee in the 1 Hz to 10 Hz range with an associated spectral flux noise density  $\approx 10^{-10}$   $\Phi_0^2/\text{Hz}^{-1}$  at 0.1 Hz. Optimizing SQUID performance is especially important for biomedical applications since operation down to the 0.1 Hz range is necessary (38).

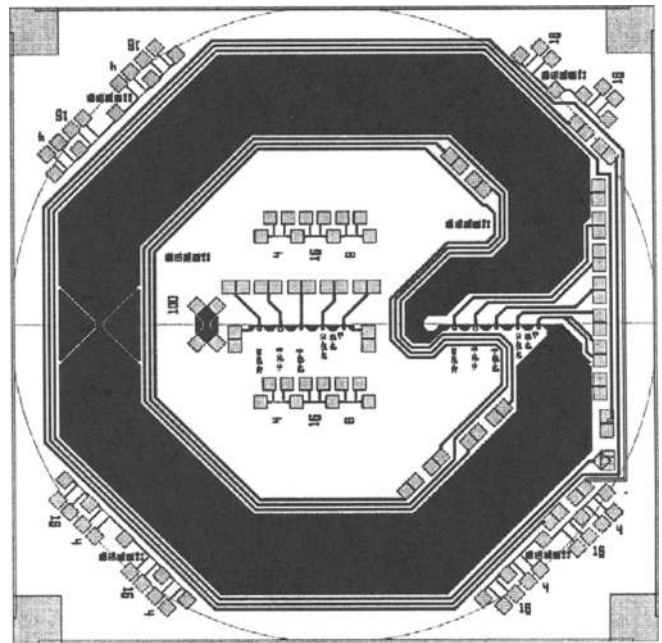
The flux noise energy of a SQUID can be written as (39)

$$\varepsilon \sim 16k_B T \left( \frac{LC}{\beta} \right)^{1/2}$$

which generally suggests the need for devices with the smallest inductance and capacitance, operating at the lowest temperatures. The use of low-temperature superconductors (Pb, Pb-alloys, Nb, and so on) generally means operation at 4.2 K, (the boiling point of liquid helium at 1 atm) and temperatures less than 80 K for the high-temperature superconductor materials. A method successfully used to reduce inductance is the washer loop (39). Because of the Meissner effect associated with superconductivity, flux transformers can be used to collect and couple flux from a larger, external superconducting loop and present it to the SQUID using a multiturn thin-film coil.

High-temperature superconductors have also been relatively successfully explored for SQUID applications. An approach to creating Josephson junctions in high-temperature superconductor materials is to simply make a step in a substrate to create a step-edge junction. Material grown across the edge acquires a defect that creates a weakly linked Josephson junction, with characteristics similar to a resistively shunted SIS system. It was realized that the actual angle of the step or ramp was important in consistently creating junctions with desired characteristics (40–42). Along the same lines, junctions made with silicon-on-sapphire (SOS) substrates have demonstrated RSJ characteristics (43, 44) and were used to make the first SQUIDs operating at 77 K using step-edge techniques (45).

A different approach to the use of high-temperature superconductor materials is the use of bicrystal substrates. Here, two single crystals (for example, of SrTiO $_3$ ) are fused together at a predetermined angle to create an off-axis interface. Junctions are made by depositing a thin film of high-temperature superconductor material across this disruptive interface. Bicrystal work in particular has consistently shown RSJ-type current-voltage characteristics and critical currents that are reproducibly correlated to the angular displacement of the substrate crystal lattices. These bicrystal junctions have been successfully incorporated into YBaCuO (46–48) and BaSrCaCuO



**Figure 7.** SQUID structure from IBM. Most of the structure seen in the figure is the superconducting loop containing a “flux dam,” visible as the triangular restriction on the left, to reduce device noise by limiting circulating currents in the flux loop lines. The Josephson junctions incorporated into the structure are step-edge devices. The overall size of the SQUID is 1 × 1 cm. From Sun et al., Ref. 53.

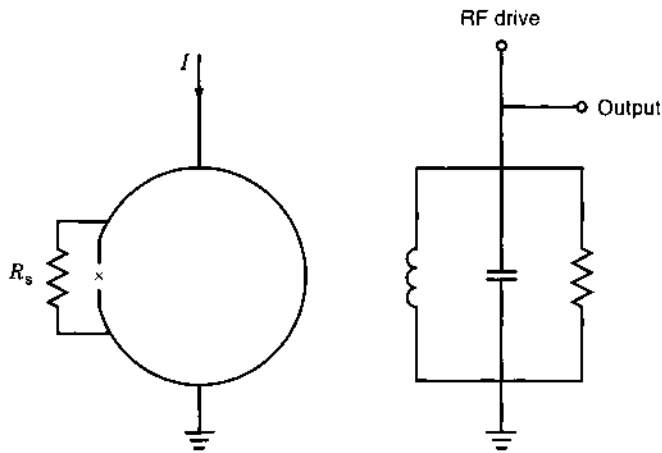
SQUIDs (49, 50). Bicrystal junctions of MgO (51) have also been discussed for use in SQUIDs targeted for nondestructive evaluation systems. Noise studies (52) suggest that bicrystal and step-edge techniques generally produce devices of comparable quality. However, although the techniques are suitable for few-device applications such as SQUIDs, they are not appropriate for medium- or large-scale integration.

The practical implementation of high-temperature SQUID technology is shown in Fig. 7. The work is from Koch’s group at IBM (53). Here junctions are formatted into a large loop containing *flux dams*, one of which is shown to the left as a constriction in the loop. The constriction creates a fuselike link to help eliminate noise-generating current loops.

Beside the double-junction dc SQUID, there is also the single-junction RF SQUID, depicted in Fig. 8. With this device, flux quantization in the ring, along with the corresponding variation of the Josephson current with flux, produce a variation in the inductance of the SQUID loop. This change in loop inductance is coupled to and thus shifts the resonant frequency of an external, RF-driven tank circuit. Because of its relative simplicity and the requirement for only a single, shunted Josephson junction, this scheme was initially and has continued to be the device of choice for many commercial applications, notably for nondestructive evaluation and SQUID voltmeters.

We note finally that, irrespective of the type of SQUID employed, applications such as biomedical diagnostics typically require operation within the confines of a carefully shielded magnetic enclosure. However, for such applications such as nondestructive evaluation field work, SQUID gradiometers can be employed that are more immune from background magnetic fields and can be optimized for unshielded performance.





**Figure 8.** Shown the schematic of an RF SQUID. This type of SQUID uses a single (shunted) Josephson junction in a loop. Applied external flux ( $\Phi$ ) acts to alter the impedance of the loop, which is inductively coupled to and shifts the frequency of an RF driven tank circuit.

This implies a compromise between field sensitivity and voltage modulation characteristics. For example, SQUIDS for non-destructive evaluation (54) are cited as having a flux noise at 1 Hz of  $134 \mu\Phi_0\text{Hz}^{-1/2}$  and  $60 \mu\Phi_0\text{Hz}^{-1/2}$  at 10 Hz, qualifying them as prototype commercial systems (55).

In addition to the measurement of flux, SQUIDS can also be configured to perform other functions. This includes the measurement of ultras small voltages. SQUID voltmeters are commercially available for measurements in the 1 pV range—limited by noise to a range of  $\sim 10^{-10} \text{ V/Hz}^{-1/2}$  at 100  $\Omega$ . Radio frequency SQUIDS have also been employed as low-noise amplifiers to  $\sim 100$  MHz.

## MIXING AND DETECTION

The zero-voltage branch of a Josephson junction (Fig. 3) can respond at frequencies up to  $f \sim (4\Delta/h) \sim (e/h)I_c R \gtrsim 1$  THz for low-temperature superconductors and  $\gtrsim 10$  THz for high-temperature materials. As a result high-frequency Josephson mixing has been extensively studied. Unfortunately, heterodyne mixing using the zero-voltage branch of the Josephson tunneling characteristic is seriously degraded by the appearance of excess noise due to noise down conversion and the dynamic impedance related to the ac Josephson effect (56).

However, the extreme nonlinearity of the quasi-particle portion of the current–voltage characteristic, as the tunneling current turns on for voltages just above the sum-gap voltage, has been successfully exploited for mixing (56–58) in the same manner as high-frequency diodes are used. For superconducting mixing, the Josephson current (zero-voltage curve) becomes a nuisance and can be suppressed by applying a small magnetic field. Such devices are generally referred to as SIS mixers, with maximum operating frequencies the same as noted previously.

SIS mixers are typically employed when uncompromising low-noise performance is required (such as with radio astronomy) and where very low ( $\sim 1 \mu\text{W}$ ) local-oscillator power is required. These mixers are typically used in the  $\sim 40$  GHz to 1 THz region, with single sideband (SSB) noise temperatures at 100 GHz of  $\sim 4$  K. Overall noise temperatures for  $\sim 40$  GHz to 1 THz are typically within an order of magnitude of the quantum

noise limit of  $hf/k_B$ . This currently exceeds the performance of high-electron-mobility transistor (HEMT) devices.

One critical element in mixing is junction quality. The magnitude of the tunneling current for voltages below the turn-on voltage for quasi-particle tunneling, that is, voltages below the sum-gap voltage, is a critical factor in mixer noise performance. Leakage conduction in this regime needs to be as small as possible and is limited in theory only by thermal excitations, but in practice can be dominated by imperfections in the tunnel barrier itself. For this reason, junctions based on oxidized tin and other soft metals were initially employed because of their extremely low so-called subgap conductance due to the high quality of the Sn-oxide barrier. However metals such as Sn are not physically robust with respect to thermal cycling (from their  $\sim 1$  K operating temperature to room temperature).

Mixing requires a very high quality dielectric barrier integrated into superconductive elements with the highest possible energy gap (and hence critical temperature) to provide the highest operating frequency. Matching requirements also mean that junction resistance be in the 20  $\Omega$  to 100  $\Omega$  range and that  $1 < 2\pi RCf < 10$ . Practical compromises have resulted in the use of hybrid systems such as Nb-based electrodes, aluminum-oxide barriers, and Pb-alloy counterelectrodes to form Nb–Al<sub>2</sub>O<sub>3</sub>–Pb-alloy junctions. This is a good union between the thermal stability of Nb and the high-quality dielectric properties of Al<sub>2</sub>O<sub>3</sub>. Nb–Al<sub>2</sub>O<sub>3</sub>–Nb junctions have also been adopted for mixing applications. The upper operational limit for Nb-based junctions is  $\sim 1.3$  THz. Junctions based on NbN (with a critical temperature of  $\sim 14$  K compared with the 9.25 K transition temperature of niobium) have also been made, which have an operational mixing limit of  $\sim 2.5$  THz. Early NbN work explored both NbN–MgO–NbN (59, 60), and NbN–AlN–NbN systems (61–63). Recent NbN work has produced devices with double sideband receiver noise temperatures of 1450 K at 600 GHz and 2800 K at 950 GHz (64).

Mixing experiments have also been conducted with high-temperature superconductor materials. For example, bicrystal silicon-based junctions have been produced for this purpose (65). Bicrystal results on silicon and sapphire (66) appear to be comparable to those achieved with SrTiO<sub>3</sub> in terms of the overall nature of their current–voltage characteristics, but without the disadvantage of the large dielectric effects of SrTiO<sub>3</sub>. Both Shapiro steps (67) and harmonic mixing of a 1.6 THz signal have been observed (68) in such systems. However, high-temperature superconductor tunnel systems that have demonstrably fulfilled the stringent requirements for quasi-particle mixing applications have yet to be produced. These applications will require (non-RSJ-like) high-quality, thin-film, SIS type junctions with low subgap conductance.

## DIGITAL APPLICATIONS

Researchers with an interest in digital applications such as ultrahigh-speed analog-to-digital conversion (ADC) have been attracted by the fundamental properties of Josephson devices. Josephson junctions are fast, with  $\sim 1$  ps switching times; have low switching power,  $P \sim I_c^2 R \sim 0.1 \mu\text{W}$ ; and can be matched to the impedance of modern microstrip technology (69).

The most extensive effort to use Josephson devices to create a prototype superconducting computer was the much-discussed program at IBM. Although the program was not successful in

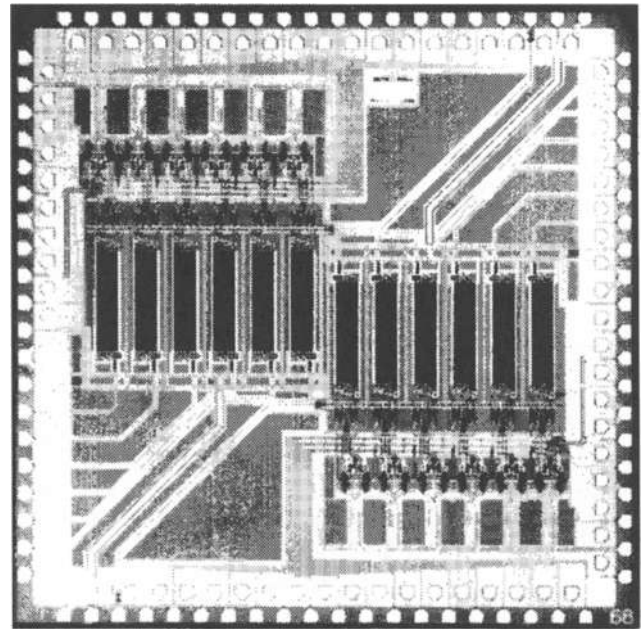
achieving its ultimate goal, a great deal was learned and much progress was made (70). Lead-alloy technology that produced junctions with reproducible characteristics and that could be thermally cycled was used. The logic elements were based on latching circuitry with underdamped junctions ( $\omega_p \ll \omega_c$ ) such that the binary state was defined by the voltage state of a Josephson junction. For example, a junction can be switched from a zero-voltage state to a nonzero voltage state along the load-line of the device. As discussed by Likharev (69), this type of approach has fast set times ( $\sim 1$  ps) but slow reset times ( $\sim 10$  ns). Other difficulties with this arrangement include problems in achieving uniform device clocking across an entire logic array and high latching power ( $\sim 1 \mu\text{W}$ ) per gate. More recent work with single-flux-quantum (SFQ) logic (referring to logic based on the presence of single fluxoids in SQUID-based logic elements) has switched to Nb-based junctions (71–74) in a variety of formats.

A new logic type, which is fundamentally different from latching-type logic, has also been introduced. Called RSFQ (resistive/rapid single flux quantum) logic (69, 75), it is a hybrid logic family, wherein junctions are configured to have both logic and (dynamic) memory functions and logic operations are performed on pulses originating from traditional SFQ devices. This approach addresses to some extent a major problem that was encountered in the IBM experiment: matching SFQ logic and memory elements. The approach also recovers much of the fundamentally fast switching times and low switching power of Josephson devices (the latter since junctions are open for only a small part of a clock period). Switching powers of  $\sim 10^{-7}$  W/gate ( $10^{-18}$  J/bit) implies that packing densities of  $10^7$  gates/cm<sup>2</sup> are possible at speeds to  $\sim 500$  GHz with  $1 \mu\text{m}$  linewidths and at greater speeds for narrower lines (4).

In present practice, much of the effort is not in computing per se but high-speed digital processing, especially ADC [and agile programmable voltage standards (76, 77)] with RSFQ as a favored digital logic format (78, 79). For example, 1024-bit shift registers have been operated at up to 20 GHz (80).

The current application of Josephson technology is still largely based on niobium. Niobium technology is mature and now appears to be the standard for digital Josephson applications. Niobium based junctions are physically robust and amenable to large-scale integration as a result of the important innovation of replacing the native oxide that grows on niobium, which generally forms a poor quality barrier with a relatively large dielectric constant ( $140 \text{ fF}/\mu\text{m}^2$ ) (81, 82). The idea is that the niobium base electrode is capped in situ with a thin aluminum film ( $\sim 1$  nm to 8 nm thick), which is oxidized to completion to form a surface-layer junction and in the process prevents the growth of NbOx (9, 83, 84). This scheme successfully combines the relatively high critical temperature,  $T_c$ , of the (robust) Nb underlayer (9.25 K) with the unsurpassed qualities of Al<sub>2</sub>O<sub>3</sub>, which is an excellent barrier material with a relatively low dielectric constant ( $60 \text{ fF}/\mu\text{m}^2$ ) (84) compared with Nb oxide. A large-scale process for creating many such junctions with a simple anodization process was developed by Kroger et al. (85, 86). Likharev has reviewed the complexities of junction fabrication for digital and other applications (4).

A good example of contemporary large-scale Josephson technology is the work at HYPRES Inc., which produces large-scale integrated Nb-based Josephson tunnel junction circuitry using RSFQ logic. A recently manufactured large-scale integration ADC is shown in Fig. 9. Target applications for such sys-



**Figure 9.** Modern analog/digital converter from HYPRES, Inc. using niobium technology. The chip is a 6-bit flash ADC with a 32-word shift register memory and operates at 16 GS/s.

tems would include high-resolution ADCs for radar and time-to-digital converters (TDCs) to measure the timing of events in high-energy and nuclear physics experiments. These systems have demonstrated flash ADC with 6-bit resolution, in the 1 GHz to 10 GHz operating range (87).

Another potential application for the fast switching speed of Josephson junctions is cross-bar and related switching matrices for switching between processors and memory (88, 89). This is important to high-speed telecommunications and computation applications.

We finally note that other approaches based on quantum-limited-logic have been proposed and may be competitive in some areas. One of these is single-electron logic (SEL), which is based on charging effects in ultrasmall capacitance tunnel junctions (and not the Josephson effect). SEL is operationally similar to RSFQ logic, but the former is based on the presence or absence of single electrons rather than magnetic fluxoids (loosely speaking, bits based on single electrons with charge  $e$  as opposed to single fluxoids with flux  $\Phi_0$ ) (4, 69).

## HIGH-TEMPERATURE SUPERCONDUCTOR SYSTEMS

The appearance of high-temperature superconductor materials has spurred a broad-based effort to produce Josephson junctions for digital and other applications. Some of these approaches were discussed previously in connection with SQUIDs. Beside this work, other approaches have been taken to produce a successful junction format. Certainly all the high-speed applications that are envisioned for niobium-based Josephson technology such as flash ADC, wideband ADCs, transient-event digitizers, and crowbar switches could be envisioned for high-temperature superconductor materials as well. Barring thermodynamic considerations, operation at higher temperatures would generally be an advantage. The biggest hurdle now for high-temperature superconductor digital

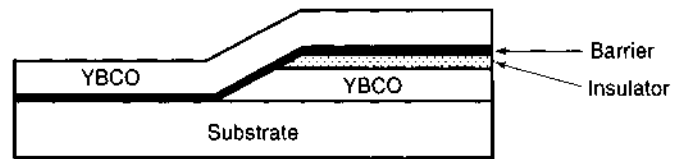
applications is the uniformity of the Josephson critical current density (Josephson critical current per unit junction area) for junctions across a single chip and chip-to-chip critical current uniformity. A maximum variation of a few percentage points in critical current density is probably necessary for large-scale applications (90), and this has yet to be achieved. At present, achievable spreads in critical current density are more in the vicinity of (or perhaps somewhat less than) 10 percent. This single consideration is a dominant factor governing which basic device format will be successful for high-temperature superconducting materials, a number of which have been and are currently being explored in an attempt to find a suitable candidate to meet this criterion.

We again note that high-temperature superconductors can be viewed as quasi-two-dimensional materials—a stack of Josephson coupled two-dimensional superconducting layers. This basic physical structure, combined with a propensity for naturally occurring grain boundaries—which also act as Josephson weak links—has led to the investigation of intrinsic Josephson effects (91–99). Some of the devices based on naturally occurring, intergranular weak links have exhibited good critical-current and normal-state characteristics. However, the use of naturally occurring weak links to create junctions per se has typically lead to poor reproducibility. Focused ion beam (FIB) techniques have also been used in an attempt to artificially induce defects at given locations by inducing substrate damage (100).

A successful approach to creating SNS type junctions in preselected locations is focused electron beam (FEB) writing. Here, an electron beam is directed at a location on a superconducting thin film, creating a narrow, damaged line of material. The material then acts as a normal metal to create an HTS/N/HTS (HTS, high-temperature superconductor) device (101–106). The technique can produce RSJ-like microwave behavior (107, 108). An example is the work at Cambridge (104), where a computer-controlled electron beam (350 KeV at 400 pA) is swept across narrow lines of YBaCuO. This process has produced junctions with an SNS character (17) with long-term room-temperature stability. One drawback of the large-scale use of FEB is its slow materials processing speed, since each device must be individually electron-beam written.

### High-Temperature Superconductor Ramp Junctions

The idea of creating a step in a substrate has also been extended to make SNS and SIS type structures where the normal and insulating materials are separately introduced films as opposed to defect-modified high-temperature superconductor material. Originally, SNS junctions used normal metals such as gold; however, such devices appear to have been dominated by interface resistance. More recently, both cobalt-doped YBaCuO and gallium-doped PrBaCaO have been used as generic barrier materials in YBCO/barrier/YBCO systems, selected because they are physically compatible with YBaCuO and their conductance properties can be tuned with doping level. Cobalt-doped YBaCuO has a relatively low resistance, as opposed to gallium-doped PrBaCuO, and thus imposes a somewhat limited device operating temperature range because the material becomes superconducting below some temperature (109), although relatively narrow critical current density spreads (~12 percent variation) have been observed with the material (110, 111). Such junctions have also been employed in SQUIDs (112).



**Figure 10.** Sketch of the ramp junction geometry used with high-temperature superconductor materials, shown here with YBaCuO films.

Gallium-doped PrBaCuO tunnel barriers tend to produce relatively high resistance barriers and must be operated below 77 K. Nevertheless, the material can produce junctions with values of  $I_c R$  in a technologically useful range (~1 mV) and with independently adjustable critical current density and resistance (113), which is important for engineering considerations. In one view of tunneling in PrBaCuO, supercurrent (zero-voltage) conduction occurs via direct tunneling through the barrier whereas quasi-particle (nonzero-voltage) conduction occurs via resonant tunneling channels within the barrier (114). The fundamentals of this issue were also addressed, with device applications in mind, by work with amorphous silicon barriers (115).

HTS/I/HTS tunnel junctions using PrBaCuO barriers in a ramp format have been produced through a variety of techniques (116, 117). A sketch of the standard ramp junction format used with high-temperature superconductor materials is shown in Fig. 10. All tend to produce devices with RSJ-like characteristics [with  $I_c R \sim 1$  mV at 4.2 K, which scales with PrBaCuO doping (118)]. The combination of step-edge substrates and doped barriers has been promising overall, and these devices have been modeled by microscopic theory (17), although the full applicability of standard proximity-effect theory with regard to high-temperature superconductor materials remains an open question (119).

To date, small circuits have been put together with these junctions, which are designed with tolerance to the large spreads that exist at present in junction critical current (120–124). Clearly, again, the challenge is to improve on the reproducibility of critical currents in this or any other device configuration that may present itself.

Related to this work are engineered interface YBaCuO/I/YBaCuO-type junctions. Here, the barriers are produced by relatively subtle damage to the base YBaCuO film and then YBaCuO is simply deposited on top. The suggestion is that the barrier comprises a thin layer of YBaCuO with oxygen disorder or deficiency perhaps a few nanometers thick (125). Other groups have been experimenting with this and related techniques (126).

In summary, a variety of promising approaches have been explored with high-temperature superconductor materials to produce junctions for SQUID and digital applications. The primary goal is to produce junctions that are parametrically reproducible, and the secondary goal is to produce a thin-film barrier compatible with the high-temperature superconductor materials with transport properties close to the metal–insulator transition. The latter would allow the junction to be self-shunting and thus allow for a broader use of the devices. This balancing act is challenging in light of the problems of material compatibility with high-temperature superconductor systems and unanswered questions regarding the physics of barriers

(especially high-temperature superconductor-compatible materials) near the metal–insulator transition.

### High-Temperature Superconductor Applications: Technical Considerations

A great deal of effort has gone into the application of high-temperature superconductor materials for SQUIDs and SFQ/RSFQ logic circuitry. Such circuits need nonhysteretic junctions with  $I_c \sim 50 \mu\text{A}$  to  $500 \mu\text{A}$  and  $R \sim 1 \Omega$  for  $10 < T < 77 \text{ K}$ , which is achieved with low-temperature superconductors at 4.2 K by shunting. In the case of high-temperature superconductor materials, compatible barrier materials—such as doped PrBaCuO or YBaCuO—typically have far lower resistivities than barriers such as  $\text{Al}_2\text{O}_3$ . Therefore, as noted, these barrier materials have been examined with a view toward obtaining high-temperature superconductor-based junctions in the parametrically desirable region where shunting would be unnecessary (109, 127, 128).

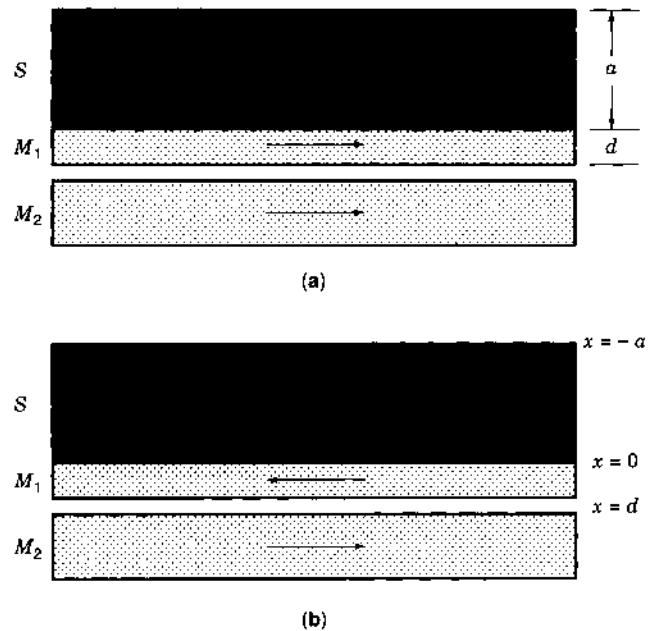
In Likharev's description of SFQ circuits (129) it was shown that for niobium junctions with  $I_c R \sim 0.2 \text{ mV}$ , operation can be in the 100 GHz range. For SFQ to work, one needs series or parallel loops of junctions with inductance  $L$  such that  $(1/2)\Phi_0 < LI_c < (3/2)\Phi_0$ . This limits the acceptable range of  $I_c$  for junctions. A lumped-parameter analysis (130) indicates that at  $\sim 10 \text{ K}$  the minimum  $I_c$  for  $I_c R \sim 0.1 \text{ mV}$  to  $1.0 \text{ mV}$  is  $\sim 0.2 \text{ mA}$  and for high temperatures,  $\sim 60 \text{ K}$  to  $70 \text{ K}$ ,  $I_c$  will need to be  $\sim 1.2 \text{ mA}$ . The analysis also indicates that to avoid latching and ensure sufficiently fast signal propagation, the condition on junction resistance will be that  $R_{\text{max}} \sim 4 \Omega$  for  $\epsilon_r = 40$  with  $1 \mu\text{m}$  lines and  $2 \Omega$  for  $\epsilon_r = 100$ , where  $\epsilon_r$  is the dielectric constant of the material adjacent to the signal lines.

This analysis and the results of other work (131, 132) suggest that devices using PrBaCuO-based barriers can achieve critical currents,  $I_c$ , in the range of  $10^5 \text{ A/cm}^2$  and still be nonhysteretic. The ultimate conclusion is that PrBaCuO barriers can give junctions parametrically compatible with SFQ applications.

### THREE-TERMINAL DEVICES

One overriding characteristic of a Josephson junction is that it is a two-terminal device, more like a forward-biased diode with a hysteretic current-voltage characteristic than a transistor, which has a gate or base terminal. However, there have been a number of proposed and experimentally examined three-terminal superconducting devices. These include hybrid superconductor–semiconductor devices, such as junction and field-effect transistors, and nonequilibrium superconducting devices, such as stacked-junctions and injection-controlled weak links (133). At present none of these devices has found its way into standard use, because they generally have not exhibited above unity gain, although in theory this appears to be possible (134). One inherent problem is the high carrier density of conventional superconducting materials, making it difficult to modulate device transconductance. However, high-temperature superconductor materials, which have far lower carrier densities and perhaps other parametric advantages for certain device configurations (135), provide somewhat more promise for three-terminal devices.

Early work on Josephson FETs per se has included InAlAs/InGaAs HEMT-type configurations, where a gate



**Figure 11.** The figure shows the structure of a proposed magnetic/superconductor memory element. Shown are parallel (a) and antiparallel (b) configurations of magnetization in the M layers. Figure from Sangjun et al., Ref. 170.

controls the magnitude of the Josephson current (136). The overall prospects for Josephson FETs (137), and experimental results with HEMPT devices (138) have also been presented. Generally it appears ultimately possible to create a device with voltage gain, driving loads of  $\sim 100 \Omega$ . FET-type structures have also been created in YBaCuO/barrier/YBaCuO SIS-type systems, where a base layer is used to apply a strong electric field across the barrier region using an external gate, causing a change in surface carrier density and a corresponding modulation of the Josephson current. Examples of this work include YBaCuO/SrTiO<sub>3</sub>/YBaCuO and related structures (139–146) and Au/SrTiO<sub>3</sub>/YBaCuO (NIS) structures (147, 148).

In general, FET structures have to date not shown anything but relatively weak effects, although work on a variety of promising systems continues (149). The work is motivated in part because the devices have the potential to serve as a Josephson-to-semiconductor logic interface. The general problem to be addressed is that Josephson logic output voltages are  $\sim 3 \text{ mV}$  (for low-temperature superconductors) while a CMOS device operates at  $\sim 1 \text{ V}$ . Some work has been specifically directed at this issue (150, 151), including interface electronics combining 4 K Josephson drivers and 77 K semiconducting HEMT amplifiers with a predicted speed of 1 GHz to 3 GHz (152).

A variety of other schemes have also been advanced for both low- and high-temperature systems. These fall under the generic categories of electric-field controlled devices (153–157), quasi-particle injection devices (158, 159), flux-flow transistors (160, 161), and dielectric-base transistors (162–66).

Although these ideas continued to be pursued, an achievable device format with useful gain has yet to emerge unambiguously. Certainly, three-terminal high-temperature superconductor devices will require higher quality junctions and for FETs, dielectric gates compatible with high-temperature superconductor materials and better geometries to provide

improved coupling of electric fields into junctions. Thorough discussions of three-terminal Josephson devices are available in the literature (155, 167–169).

## RELATED SUPERCONDUCTING DEVICES

While not superconducting tunneling devices, several systems have recently been developed that could interface with superconducting devices and logic. One example, illustrated in Fig. 11, is a memory element based on superconductor–magnetic multilayers (170).

There has also recently been considerable interest in the physics and device implications of tunnel junctions containing magnetic elements (171–175), which may also be compatible with and complementary to Josephson circuitry.

## ACKNOWLEDGMENTS

I wish to thank Alan Kleinsasser for many useful comments on the text. I also wish to acknowledge insightful discussions with M. R. Beasley, D. A. Moore, J. M. Rowell, and J. Talvacchio and comments by Alan Kadin. I also gratefully acknowledge support during the preparation of this work by the Department of Energy, Division of Materials Sciences, through grant DE-FG02-88ER45373.

## BIBLIOGRAPHY

1. T. Van Duzer and C. W. Turner, *Principles of Superconductive Devices and Circuits*, New York: Elsevier North Holland, 1981.
2. A. Barone and G. Paterno, *Principles and Applications of the Josephson Effect*, New York: Wiley, 1982.
3. A. Barone and G. Patterson, *Physics and Application of the Josephson Effect*, New York: Wiley, 1982.
4. K. K. Likharev, *Dynamics of Josephson Junctions and Circuits*, New York: Gordon and Breach, 1986.
5. H. Weinstock and M. Nisenoff, *Superconducting Electronics*, Berlin: Springer-Verlag, 1989.
6. J. C. Gallop, *SQUIDS, the Josephson Effects and Superconducting Electronics*, New York: Adam Hilger, 1990.
7. S. T. Ruggiero and D. A. Rudman, *Superconducting Devices*, New York: Academic Press, 1990.
8. T. P. Orlando and K. A. Delin, *Foundations of Applied Superconductivity*, Reading, MA: Addison-Wesley, 1991.
9. S. T. Ruggiero, Artificial tunnel barriers, in S. T. Ruggiero, D. A. Rudman (eds.), *Superconducting Devices*, New York: Academic Press, 1990, pp. 373–390.
10. I. Giaever, Metal-insulator-metal tunneling, in E. Burstein and S. Lundqvist (eds.), *Tunneling Phenomena in Solids*, New York: Plenum, 1969, pp. 19–30.
11. J. G. Simmons, Low-voltage current-voltage relationships of tunnel junctions, *J. Appl. Phys.*, **34**: 238, 1963.
12. J. G. Simmons, Image force in metal-oxide-metal tunnel junctions, in E. Burstein and S. Lundqvist (eds.), *Tunneling Phenomena in Solids*, New York: Plenum, 1969, pp. 135–148.
13. S. M. Sze, *Physics of Semiconductor Devices*, New York: Wiley, 1981.
14. B. D. Josephson, Possible new effects in superconducting tunneling, *Phys. Lett.*, **1**: 251, 1962.
15. B. D. Josephson, The discovery of tunneling supercurrents, *Science*, **184**: 527, 1974.
16. P. W. Anderson and J. M. Rowell, Probable observation of the Josephson superconducting tunneling effect, *Phys. Rev. Lett.*, **10**: 230, 1963.
17. K. A. Delin and W. A. Kleinsasser, Stationary properties of high-critical-temperature proximity effect Josephson junctions, *Supercond. Sci. Technol.*, **9**: 227, 1996.
18. J. Lukens, Josephson arrays as high frequency sources, in S. T. Ruggiero and D. A. Rudman (eds.), *Superconducting Devices*, New York: Academic Press, 1990, pp. 135–167.
19. S. P. Benz, Superconductor-normal-superconductor junctions for programmable voltage standards, *Appl. Phys. Lett.*, **67**: 2714, 1995.
20. C. A. Hamilton, C. J. Burroughs, and S. P. Benz, Josephson voltage standard—a review, *IEEE Trans. Appl. Supercon.*, **7**: 3756, 1997.
21. S. P. Benz, et al., Stable 1 volt programmable voltage standard, *Appl. Phys. Lett.*, **71**: 1866, 1997.
22. L. Fritsch, et al., Superconductor–normal metal–superconductor Josephson junctions with Ti interlayer, *Appl. Phys. Lett.*, **73**: 1583, 1998.
23. I. P. Nevirkovets, Fabrication, dc characteristics of vertically stacked SIS-type structures for use as low-temperature detectors, *Supercond. Sci. Technol.*, **8**: 575, 1995.
24. A. V. Ustinov, H. Kohlstedt, and C. Heiden, Coherent flux-flow in vertically stacked long Josephson tunnel junctions, *IEEE Trans. Appl. Superconduct.*, **5**: 2743, 1995.
25. A. Shoji, T. Kikuchi, and H. Yamamori, Uniformity of critical currents of vertically stacked Josephson junctions, in *Extended Abstr. 5th Int. Workshop High-Temp. Supercond. Electron Devices*, Tokyo, **FED-157**: 157, 1997.
26. M. G. Blamire, et al., Characteristics of vertically-stacked planar tunnel junction structures, *IEEE Trans. Magn.*, **MAG-25**: 1135, 1989.
27. I. P. Nevirkovets, L. P. Stryzhko, and A. V. Poladich, Photon-assisted tunneling in stacked tunnel structures by millimeter wave irradiation, *Physica B*, **194–196**: 2395, 1994.
28. S. Lomatch, E. D. Rippert, and J. B. Ketterson, Multilayer Josephson junction flux quantum devices, *IEEE Trans. Appl. Supercond.*, **5**: 3147, 1995.
29. M. Scheuerman, J. T. Chen, and J. J. Chang, Interaction between microwaves and a single vortex in a long Josephson tunnel junction, *J. Appl. Phys.*, **54**: 3286, 1983.
30. S. Yoshikawa, et al., Intrinsic Josephson junctions of  $Tl_2Ba_2CaCuO_x$  thin films for high frequency devices, *Physica C*, **293**: 44, 1997.
31. M. Tinkham, *Introduction to Superconductivity*, 2nd ed., New York: McGraw-Hill, 1996.
32. N. Miyakawa, et al., Strong dependence of the superconducting gap on oxygen doping from tunneling measurements on  $Bi_2Sr_2CaCu_2O_{7-\delta}$ , *Phys. Rev. Lett.*, **80**: 157, 1998.
33. A. W. Kleinsasser and R. A. Buhrman, High-quality submicron niobium tunnel junctions with reactive-ion-beam oxidation, *Appl. Phys. Lett.*, **37**: 841, 1980.
34. D. J. van Harlingen, R. H. Koch, and J. Clarke, De SQUID near the quantum noise limit, *Physica (Utrecht)*, **B108**: 1083, 1981.
35. D. J. van Harlingen, R. H. Koch, and J. Clarke, Superconducting quantum interference device with very low magnetic flux noise energy, *Appl. Phys. Lett.*, **41**: 197, 1982.
36. Y. Zhang, et al., Applications of high-temperature SQUIDS, *Appl. Supercond.*, **3**: 367, 1995.
37. F. Ludwig, et al., Multilayer magnetometers based on high- $T_c$  SQUIDS, *Appl. Supercond.*, **3**: 383, 1995.

38. G. L. Romani, The use of SQUIDS in the study of biomagnetic fields, in H.-D. Hahlbohm and H. Lubbig (eds.), *SQUID '85*, Berlin: de Gruyter, 1985.
39. J. Clarke, SQUIDS: Principles, noise and applications, in S. T. Ruggiero and D. A. Rudman (eds.), *Superconducting Devices*, New York: Academic Press, 1990, pp. 51–99.
40. J. Gao, et al., Controlled preparation of all high- $T_c$  SNS-type edge junctions and dc SQUIDS, *Physica C*, **171**: 126, 1990.
41. C. L. Jia, et al., Microstructure of epitaxial  $YBa_2Cu_3O_{7-x}$  films on step-edge SrTiO, *Physica C*, **175**: 545, 1991.
42. R. B. Laibowitz, et al., All high  $T_c$  edge junctions and SQUIDS, *Appl. Phys. Lett.*, **56**: 686, 1990.
43. Y. A. Boikov, et al., Biepitaxial Josephson junctions with high critical current density based on  $YBa_2/Cu_3/O_{7-x}$  films on silicon on sapphire, *J. Appl. Phys.*, **77**: 1654, 1995.
44. B. N. Hunt, et al., High- $T_c$  edge-geometry SNS weak link on silicon-on-sapphire substrates, *Physica C*, **230**: 141, 1994.
45. M. J. Burns, et al., Demonstration of  $YBa_2Cu_3O_{7-x}$  and complementary metal-oxide-semiconductor device fabrication on the same sapphire substrate, *Appl. Phys. Lett.*, **63**: 1282, 1993.
46. D. Koelle, et al., High performance dc SQUID magnetometers with single layer  $YBa_2/Cu_3/O_{7-x}$  flux transformers, *Appl. Phys. Lett.*, **63**: 3630, 1993.
47. R. Cantor, et al., Low-noise, single-layer  $YBa_2/Cu_3/O_{7-x}$  dc SQUID magnetometers at 77 K, *IEEE Trans. Appl. Supercond.*, **5**: 2927, 1995.
48. F. Dillno, V. N. Giyantsev, and M. Siegel, Performance of  $YBa_2/Cu_3/O_{7-x}$  direct current SQUIDS with high-resistance, *Appl. Phys. Lett.*, **69**: 1948, 1996.
49. T. Amrein, et al., Thin film  $Bi_2Sr_2CaCu_2O_{8+x}$  Josephson junctions and direct current superconducting quantum interference devices on (001) SrTiO<sub>3</sub> bicrystals, *Appl. Phys. Lett.*, **63**: 1978, 1993.
50. T. Takami, et al., Dc superconducting quantum interference devices with BiSrCaCuO bicrystal grain boundary junctions at 77 K, *Jpn. J. Appl. Phys.*, **35**: L391, 1996.
51. C. Carr, et al., Electromagnetic nondestructive evaluation: Moving HTS SQUIDS, including field nulling and dual frequency measurements, *IEEE Trans. Appl. Supercond.*, **7**: 3275, 1997.
52. L. Hao, J. C. Macfarlane, and C. M. Pegrum, Excess noise in  $YBa_2Cu_3O_7$  thin film grain boundary Josephson junctions and devices, *Supercond. Sci. Technol.*, **9**: 678, 1996.
53. J. Z. Sun, et al., Improved process for high- $T_c$  superconducting step-edge junctions, *Appl. Phys. Lett.*, **63**: 1561, 1993.
54. A. Cochran, et al., Experimental results in non-destructive evaluation with HTS SQUIDS, in *Proc. EUCAS '95*, Institute Physics Conf., **2**: 1511, 1995.
55. L. N. Morgan, et al., Electromagnetic nondestructive evaluation with simple HTS SQUIDS: Measurements and modeling, *IEEE Trans. Appl. Supercond.*, **5**: 3127, 1995.
56. Q. Hu and P. L. Richards, Quasiparticle mixers and detectors, in S. T. Ruggiero and D. A. Rudman (eds.), *Superconducting Devices*, New York: Academic Press, 1990, pp. 169–196.
57. J. R. Tucker and M. J. Feldman, Quantum detection at millimeter wavelengths, *Rev. Mod. Phys.*, **57**: 1055, 1985.
58. P. L. Richards and Q. Hu, Superconducting components for infrared and millimeter-wave receivers, *Proc. IEEE*, **77**: 1233, 1989.
59. J. A. Stern, et al., NbN/MgO/NbN SIS tunnel junctions for sub mm wave mixers, *IEEE Trans. Magn.*, **25**: 1054, 1989.
60. W. R. McGrath, et al., Performance of NbN superconductive tunnel junctions as SIS mixers at 205 GHz, *IEEE Trans. Magn.*, **27**: 2650, 1991.
61. W. Zhen, A. Kawakami, and Y. Uzawa, NbN/AlN/NbN tunnel junctions with high current density up to 54 kA/cm<sup>2</sup>, *Appl. Phys. Lett.*, **70**: 114, 1997.
62. Y. Uzawa, et al., Submillimeter wave responses in NbN/AlN/NbN tunnel junctions, *Appl. Phys. Lett.*, **66**: 1992, 1995.
63. Y. Uzawa, W. Zhen, and A. Kawakami, Quasi-optical submillimeter-wave mixers with NbN/AlN/NbN tunnel junctions, *Appl. Phys. Lett.*, **69**: 2435, 1996.
64. P. Dielman, et al., Shot noise in NbN SIS junctions suitable for THz radiation detection, in *Appl. Superconductivity, Proc. of EUCAS 1997 3rd. Eur. Conf. Appl. Superconductivity*, **1**: 421, 1997.
65. J. Chen, et al.,  $YBa_2Cu_3O_{7-x}$  angle grain boundary junction on Si bicrystal substrate, *Jp. J. Appl. Phys.*, **30**: 1964, 1991.
66. P. Seidel, S. Linzen, and F. Schmidl, High- $T_c$  thin films and Josephson junctions on silicon and sapphire substrates, in *Extended Abstr. 5th Int. Workshop High-Temp. Supercond. Electron Devices*, Tokyo, **FED-157**: 178, 1997.
67. P. A. Rosenthal and E. N. Grossman, Terahertz Shapiro steps in high temperature SNS Josephson junctions, *IEEE Trans. Microw. Theory Tech.*, **42**: 547, 1994.
68. K. Nakajima et al., THz response of YBCO grain boundary junctions, in *Extended Abstr. 5th Int. Workshop High-Temp. Supercond. Electron Devices*, Tokyo, **FED-157**: 185, 1997.
69. K. K. Likharev, V. K. Semenow, and A. B. Zorin, New possibilities for superconductor devices, in S. T. Ruggiero and D. A. Rudman (eds.), *Superconducting Devices*, New York: Academic Press, 1990, pp. 1–49.
70. Collection of articles, *IBM J. Res. Develop.*, **24**, 1980.
71. S. Kotani et al., Feasibility of an ultra-high-speed Josephson multiplier, *IEEE J. Solid-State Circuits.*, **SC-22**: 98, 1987.
72. H. Hayakawa, Computing, in S. T. Ruggiero, D. A. Rudman (eds.), *Superconducting Devices*, New York: Academic Press, 1990, pp. 101–134.
73. T. van Duzer and G. Lee, Digital signal processing, in S. T. Ruggiero and D. A. Rudman (eds.), *Superconducting Devices*, New York: Academic Press, 1990, pp. 197–225.
74. R. S. Withers, Wideband analog signal processing, in S. T. Ruggiero and D. A. Rudman (eds.), *Superconducting Devices*, New York: Academic Press, 1990, pp. 227–272.
75. K. K. Likharev, O. A. Mukhanov, and V. K. Semenov, Resistive single flux quantum logic for the Josephson-junction digital technology, in H.-D. Hahlbohm and H. Lubbig (eds.) *SQUID '85*, Berlin: de Gruyter, 1985, pp. 1103–1108.
76. S. P. Benz, C. J. Burroughs, and C. A. Hamilton, Operating margins for a pulse-driven programmable voltage standard, *IEEE Trans. Appl. Supercond.*, **7**: 2653, 1997.
77. S. P. Benz and C. A. Hamilton, A pulse-driven programmable Josephson voltage standard, *Appl. Phys. Lett.*, **68**: 3171, 1996.
78. K. K. Likharev and V. K. Semenov, RSFQ logic/memory family: A new Josephson-junction technology for sub-terahertz-clock-frequency digital systems, *IEEE Trans. Appl. Supercond.*, **1**: 3, 1991.
79. K. K. Likharev, Ultrafast superconducting electronics: RSFQ technology roadmap, *Proc. Low Temp. Phys. Conf. LT-21*, 1996.
80. O. A. Mukhanov, RSFQ 1024-bit shift register for acquisition memory, *IEEE Trans. Appl. Supercond.*, **3**: 3102, 1993.
81. J. H. Magerlein, Specific capacitance of Josephson tunnel junctions, *IEEE Trans. Magn.*, **17**: 286, 1981.
82. R. F. Broom et al., Niobium oxide-barrier tunnel junction, *IEEE Trans. Elect. Dev.*, **ED-27**: 1998, 1980.
83. E. L. Wolf, et al., Proximity electron tunneling spectroscopy I. Experiments on Nb, *J. Low Temp. Phys.*, **40**: 19, 1980.
84. M. Gurvitch, M. A. Washington, and H. A. Huggins, High quality refractory tunnel junctions using thin aluminum layers, *Appl. Phys. Lett.*, **42**: 472, 1983.

85. H. Kroger, L. N. Smith, and D. W. Jillie, Selective niobium anodization process for fabricating Josephson tunnel junctions, *Appl. Phys. Lett.*, **39**: 280, 1981.
86. D. Jillie et al., All-refractor, Josephson logic circuits, *IEEE J. Sol. State Circuits*, **SC-18**: 173, 1983.
87. P. Bradley, A 6-bit Josephson flash A/D converter with GHz input bandwidth, *IEEE Trans. Appl. Supercond.*, **3**: 2550, 1993.
88. T. Sterling, P. Messina, and P. H. Smith, *Enabling Technologies for Petaflops Computing*, Cambridge, MA, MIT Press, 1996.
89. N. B. Dubash, P.-F. Yuh, and V. V. Brozenets, SFQ data communication switch, *IEEE Trans. Appl. Supercond.*, **7**: 2681, 1997.
90. D. L. Miller, J. X. Przybysz, and J. H. Kang, Margins and yields of SFQ circuits in HTS materials, *IEEE Trans. Appl. Supercond.*, **3**: 2728, 1993.
91. Y. Ishimaru et al., Observation of boundary Josephson junction with d-wave pairing characteristics, *Jpn. J. Appl. Phys.*, **34**: L1532, 1995.
92. R. Kleiner, et al., Intrinsic Josephson effects in high- $T_c$  superconductors, *Phys. Rev.*, **B49**: 1327, 1994.
93. R. Kleiner, et al., Dynamic behavior of Josephson-coupled layered structures, *Phys. Rev.*, **B50**: 3942, 1994.
94. A. Irie and G. Oya, Flux flow behavior in the intrinsic Josephson junctions in  $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_y$  single crystals, *Physica C*, **235**: 3277, 1994.
95. Y. I. Latyshev, J. E. Nevelskaya, and P. Monceau, Dimensional crossover for intrinsic dc Josephson effect in  $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_8$  2212 single crystal whiskers, *Phys. Rev. Lett.*, **77**: 932, 1996.
96. A. Yurgens, et al., Strong temperature dependence of the c-axis gap parameter of  $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$  intrinsic Josephson junctions, *Phys. Rev.*, **B53**: R8887, 1996.
97. A. Irie, M. Sakakibara, and G. Oya, Growth and tunneling properties of  $(\text{Bi,Pb})_2\text{Sr}_2\text{CaCu}_2\text{O}_y$  single crystals, *IEICE Trans. Electron*, **E77-C**: 1191, 1994.
98. A. Irie and G. Oya, Microwave response of intrinsic Josephson junctions in BSCCO single crystals, *IEEE Trans. Appl. Supercond.*, **5**: 3267, 1995.
99. A. Irie, M. Iwama, and G. Oya, Novel microwave-induced steps of intrinsic Josephson junctions in mesa-shaped BSCCO single crystals, *Supercond. Sci. Technol.*, **9**: A14, 1996.
100. C. Neumann et al., Fabrication of high  $I_c R_n$  YBCO-Josephson-junctions on MgO-substrates using a focused-ion-beam system, *Physica*, **C210**: 138, 1993.
101. A. J. Pauza et al., High- $T_c$  Josephson junctions by electron beam irradiation, *IEEE Trans. Appl. Supercond.*, **3**: 2405, 1993.
102. S. K. Tolpygo, et al., Normal-state properties and Josephson effects in HTS weak links produced by electron beam, *IEEE Trans. Appl. Supercond.*, **5**: 2521, 1995.
103. B. A. Davidson et al., Superconductor-normal-superconductor behavior of Josephson junctions scribed in  $\text{Y}_1\text{Ba}_2\text{Cu}_3\text{O}_{7-\delta}$  by a high-brightness electron source, *Appl. Phys. Lett.*, **68**: 3811, 1996.
104. D. F. Moore et al., Asymmetric YBaCuO Interferometers and SQUIDS made with focused electronbeam irradiation junctions, *IEEE Trans. Appl. Supercond.*, **7**: 2494, 1997.
105. W. E. Booij et al., Electrodynamics of closely coupled YBCO junctions, *IEEE Trans. Appl. Supercond.*, **7**: 3025, 1997.
106. A. J. Pauza et al., Electron-beam damaged high-temperature superconductor Josephson junctions, *J. Appl. Phys.*, **82**: 5612, 1997.
107. S.-J. Kim et al., Successful fabrication of bicrystal Si substrates for  $\text{YBa}_2\text{Cu}_3\text{O}_{7-y}$  Josephson junctions, *Cryogenics*, **35**: 901, 1995.
108. S.-J. Kim et al., Microwave induced steps of YBCO weak link damaged by focused electron beam irradiation, in *Extended Abstr. 5th Int. Workshop on High-Temp. Supercond. Electron Devices*, Tokyo, **FED-157**, 182, 1997.
109. L. Antognazza et al., Proximity effect in  $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}/\text{YBa}_2\text{Cu}_{1-x}\text{Co}_x\text{O}_3\text{O}_{7-\delta}/\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$  junctions: From the clean limit to the dirty limit with pair breaking, *Phys. Rev.*, **B51**: 8560, 1995.
110. W. H. Mallison et al., A multilayer  $\text{YBa}_2\text{Cu}_3\text{O}_x$  Josephson junction process for digital circuit applications, *Appl. Phys. Lett.*, **68**: 3808, 1996.
111. H. Sato et al., Improvement in parameter spreads of YBaCuO/PrBaCuO/YBaCuO trilayer junctions, *Jpn. J. Appl. Phys.*, **35**: L1411, 1996.
112. Q. X. Jia, et al., Directly coupled direct current superconducting quantum interference devices based on ramp-edge Ag: YBCO/PBCO/Ag: YBCO junctions, *Appl. Phys. Lett.*, **72**: 3068, 1998.
113. M. A. J. Verhoeven, High- $T_c$  superconducting ramp-type junctions., PhD thesis, University of Twente, The Netherlands, 1996.
114. M. A. J. Verhoeven, et al., Ramp-type junction parameter control by Ga doping of  $\text{PrBa}_2\text{Cu}_3\text{O}_{7-\delta}$  barriers, *Appl. Phys. Lett.*, **69**: 848, 1996.
115. S. J. Bending and M. R. Beasley, Transport processes via localized states in thin a-Si tunnel barriers, *Phys. Rev. Lett.*, **55**: 324, 1985.
116. M. A. J. Verhoeven, G. J. Gerritsma, and H. Rogalla, Ramp-type junctions with very thin PBCO barriers, in *Proc. EUCAS '95 Conf.*, Edinburgh, 1995, p. 1395.
117. J. B. Barner and A. W. Kleinsasser, High- $T_c$  superconductor-normal-superconductor junctions with polyimide-passivated ambient-temperature edge formation, *IEEE Trans. Appl. Supercond.*, **7**: 2502, 1997.
118. H. Myoren et al., Terahertz response of high- $T_c$  ramp-type junctions on MgO, in *Extended Abstr. 5th Int. Workshop High-Temp. Supercond. Electron Devices*, Tokyo, **FED-157**: 188, 1997.
119. E. Demler et al., Proximity effect and Josephson coupling in the SO(5) theory of high- $T_c$  superconductivity, *Phys. Rev. Lett.*, **80**: 2917, 1998.
120. M. A. J. Verhoeven et al., Ramp type HTS Josephson junctions with PrBaCuGaO barriers, *IEEE Trans. Appl. Supercond.*, **5**: 2095, 1995.
121. A. E. Lee, J. A. Luine, and C. L. Pettiette-Hall, Co-doped-YBCO SNS junctions for superconductive integrated circuits, *IEEE Trans. Appl. Supercond.*, **5**: 2980, 1995.
122. M. G. Forrester et al., A single flux quantum shift register operating at 65 K, *IEEE Trans. Appl. Supercond.*, **5**: 3401, 1995.
123. V. K. Kaplunenko, Voltage divider based on submicron slits in a high  $T_c$  superconducting film and two bicrystal grain boundaries, *Appl. Phys. Lett.*, **67**: 282, 1995.
124. S. J. Berkowitz, Int. Workshop on Superconductivity, High-temperature superconducting circuits using SNS long junctions transistors, *High Temperature Supercond. Electron. Fund. Appl.*, Int. Superconductivity Technol. Centre, Tokyo Japan, 1996, p. 40.
125. B. H. Moeckly, et al., Interface-engineering high- $T_c$  Josephson junctions, *Appl. Supercond.*, **6**: 317, 1998.
126. J. Talvacchio, private communication.
127. K. Char, HTS SNS Josephson junctions: Interfaces and mechanisms, *Physica C*, **282-287**: 419, 1997.
128. L. Antognazza, K. Char, and T. H. Geballe, High- $T_c$  edge junctions with YPrBaCuCoO barrier layers near the metal-insulator transition, *Appl. Phys. Lett.*, **70**: 3152, 1997.
129. S. V. Polonsky et al., New RSGQ circuits (Josephson junction digital devices), *IEEE Trans. Appl. Supercond.*, **3**: 2566, 1993.
130. J. Yoshida et al., High- $T_c$  Josephson junction technology for digital applications, in *Extended Abstr. 5th Int. Workshop High-Temp. Supercond. Electron Devices*, Tokyo, **FED-157**: 133, 1997.
131. J. Yoshida and T. Nagano, Tunneling and hopping conduction via

- localized states in thin PBCO barriers, *Phys. Rev.*, **B55**: 11860, 1997.
132. T. Nagano, et al., Thin film growth for fabrication of YBaCuO/PrBaCuCoO/YBaCuO ramp-edge junctions, in *Extended Abstr. 5th Int. Workshop High-Temp. Supercond. Electron Devices*, Tokyo **FED-157**: 233, 1997.
133. A. W. Kleinsasser and W. J. Gallagher, Three-terminal devices, in S. T. Ruggiero and D. A. Rudman (eds.), *Superconducting Devices*, New York: Academic Press, 1990, pp. 325–372.
134. K. K. Likharev et al., Josephson junction with lateral injection as a vortex transistor, *IEEE Trans. Magn.*, **15**: 420, 1979.
135. W. E. Booij, et al., Supercurrents in closely spaced high  $T_c$  Josephson junctions, in *Proc. ATS-Workshop '96*, Twente, The Netherlands, 1996, p. 6.
136. T. Akazaki et al., A Josephson field effect transistor using an InAs-inserted-channel  $\text{In}_{0.52}\text{Al}_{0.48}/\text{As}/\text{In}_{0.53}/\text{Ga}_{0.47}\text{As}$  inverted modulation-doped structure, *Appl. Phys. Lett.*, **68**: 418, 1996.
137. A. W. Kleinsasser and T. N. Jackson, Prospects for proximity effect superconducting FETs, *IEEE Trans. Magn.*, **MAG-25**: 1274, 1989.
138. H. Takayanagi and T. Akazaki, Superconducting three-terminal devices using HEMT structures, in *Extended Abstr. 5th Int. Workshop High-Temp. Supercond. Electron Devices*, Tokyo, **FED-157**: 153, 1997.
139. J. Mannhart, et al., Electric field effect on superconducting YBaCuO films. *Z. Phys.*, **B83**: 307, 1991.
140. X. X. Xi, et al., Electric field effects in high  $T_c$  superconducting ultrathin YBaCuO films, *Appl. Phys. Lett.*, **59**: 3470, 1992.
141. K. Matsui, et al., Analysis on electric field effect in  $\text{Al}/\text{SrTiO}_3/\text{YBa}_2\text{Cu}_3\text{O}_y$  structure, *Jpn. J. Appl. Phys.*, **31**: L1342, 1992.
142. K. Joosse, et al., Multilayers for high- $T_c$  superconducting electric field effect devices, *J. Alloys Comp.*, **195**: 723, 1993.
143. T. Nakamura et al., Study on a superconducting channel transistor, *Sumitomo Elect. Tech. Rev.*, No. 38, 35, 1994.
144. R. Schneider and R. Auer, Temporal response of a high- $T_c$  superconducting field effect transistor, *Appl. Phys. Lett.*, **67**: 2075, 1995.
145. T. Nakamura, H. Tokuda, and M. Iiyama, Electric field effects of  $\text{SrTiO}_3/\text{YBaCuO}$  multilayers grown by all-MBE, *IEEE Trans. Appl. Supercond.*, **5**: 2875, 1995.
146. T. Fujii, et al., Metal-insulator-superconductor field-effect-transistor using  $\text{SrTiO}_3/\text{YBaCuO}$  heteroepitaxial films, *Jpn. J. Appl. Phys.*, **31**: L612, 1992.
147. T. Saito, et al., Fabrication of a high- $T_c$  superconducting field effect transistor by ion beam sputtering, *IEEE Trans. Appl. Supercond.*, **7**: 3528, 1997.
148. T. Goto, et al., Properties of high- $T_c$  field effect devices with ion beam sputtered  $\text{SrTiO}$  Film, in *Extended Abstr. 5th Int. Workshop High-Temp. Supercond. Electron Devices*, Tokyo, **FED-157**: 159, 1997.
149. T. Nakamura, H. Inada, and M. Iiyama, Superconducting field effect transistor of YBCO/ $\text{SrTiO}_3$ /YBCO tri-layer structures, in *Extended Abstr. 5th Int. Workshop High-Temp. Supercond. Electron Devices*, Tokyo, **FED-157**: 123, 1997.
150. U. Ghoshal, et al., CMOS amplifier designs for Josephson-CMOS interface circuits, *IEEE Trans. Appl. Supercond.*, **5**: 2640, 1995.
151. D. Crupta et al., Optimization of hybrid JJ/CMOS memory operating temperatures, *IEEE Trans. Appl. Supercond.*, **7**: 3307, 1997.
152. K. Hikosaka, et al., Josephson-semiconductor high-speed interface circuit using 77 K analog HEMT amplifiers, in *Extended Abstr. 5th Int. Workshop High-Temp. Supercond. Electron Devices*, Tokyo, **FED-157**: 55, 1997.
153. J. Mannhart, High- $T_c$  transistors, *Supercond. Sci. Technol.*, **9**: 49, 1996.
154. J. Chen et al., Possible three-terminal device with YBCO angle grain boundary, *IEEE Trans. Appl. Supercond.*, **1**: 102, 1991.
155. J. Mannhart et al., Large electric field effects in  $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$  films containing weak links, *Appl. Phys. Lett.*, **62**: 630, 1993.
156. Z. G. Ivanov, Field effect transistor based on a bi-crystal grain boundary Josephson junction, *IEEE Trans. Appl. Supercond.*, **3**: 2925, 1993.
157. B. Mayer, J. Mannhart, and H. Hilgenkamp, Electric field controllable Josephson junctions of high quality in high- $T_c$  superconductors, *Appl. Phys. Lett.*, **68**: 3031, 1996.
158. I. Iguchi, N. Nukui, and K. Lee, Dynamic Cooper-pair breaking by tunnel injection of quasiparticles into a high- $T_c$   $\text{YBa}_2\text{Cu}_3\text{O}_7$  superconductor, *Phys. Rev.*, **B50**: 457, 1994.
159. E. Kume, H. Hiroyuki, and I. Iguchi, Quasi-particle injection effect of YBCO thin films using double-injector junctions having two-holes, in *Extended Abstr. 5th Int. Workshop High-Temp. Supercond. Electron Devices*, Tokyo, **FED-157**: 167, 1997.
160. D. J. Frank, M. J. Brady, and A. Davidson, A new superconducting-base transistor, *IEEE Trans. Magn.*, **MAG-21**: 721, 1985.
161. T. Yamada, et al., (Ba, Rb)BiO<sub>3</sub> base and  $\text{YBa}_2\text{Cu}_3\text{O}_x$  flux flow transistors, in *Extended Abstr. 5th Int. Workshop High-Temp. Supercond. Electron Devices*, Tokyo, **FED-157**: 142, 1997.
162. H. Tamura, A. Yoshida, and S. Hasuo, Transistor action based on field-effect controlled current injection into an insulator/ $\text{SrTiO}_3$  interface, *Appl. Phys. Lett.*, **59**: 298, 1991.
163. A. Yoshida, et al., Dielectric-base transistor using  $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}/\text{NdGaO}_3/\text{SrTiO}_3$  heterostructures, *J. Appl. Phys.*, **71**: 5284, 1992.
164. T. Hato et al., Improved emitter-base junction with  $\text{In}_2\text{O}_3$  in dielectric-base transistor *Jpn J. Appl. Phys.*, **34**: 6379, 1995.
165. C. Yoshida, et al., Electric field effect in  $\text{LaTiO}_3/\text{SrTiO}_3$  heterostructure, *Jpn. J. Appl. Phys.*, **35**: 5691, 1996.
166. T. Hato, et al., Dielectric-base transistor with doped channel, *Appl. Phys. Lett.*, **70**: 2900, 1997.
167. A. W. Kleinsasser and W. J. Gallagher, Three-terminal devices, in S. T. Ruggiero and D. A. Rudman (eds.), *Superconducting Devices*, **372**. New York: Academic Press, 1990, pp. 325–372.
168. A. W. Kleinsasser, H. Weinstock, and R. W. Ralston (eds.), *Superconducting Devices*, Dordrecht: Kluwer Academic, 1993, p. 249.
169. J. Mannhart, Changes in the superconducting properties of high- $T_c$  cuprates produced by applied electric fields, *Mod. Phys. Lett.*, **B6**: 555, 1992.
170. O. Sangjun, D. Youm, and M. R. Beasley, A superconductive magnetoresistive memory element using controlled exchange interaction, *Appl. Phys. Lett.*, **71**: 2376, 1997.
171. J. S. Moodera, J. Nowak, and J. M. van de Veedonk, Interference magnetism and spin wave scattering in ferromagnetic-insulator-ferromagnet tunnel junctions, *Phys. Rev. Lett.*, **80**: 2941, 1998.
172. P. K. Wong, J. E. Evetts, and M. G. Blamire, High conductance small area magnetoresistive tunnel junctions, *Appl. Phys. Lett.*, **73**: 384, 1998.
173. S. Sankar, A. E. Berkowitz, and D. J. Smith, Spin-dependent tunneling in discontinuous Co-SiO<sub>2</sub> magnetic tunnel junctions, *Appl. Phys. Lett.*, **73**: 535, 1998.
174. J. Z. Sun, et al., Temperature and bias dependence of magnetoresistance in doped manganite thin film trilayer junctions, *Appl. Phys. Lett.*, **73**: 1008, 1998.
175. K. Inomata and Y. Saito, Spin-dependent tunneling through layered ferromagnetic nanoparticles, *Appl. Phys. Lett.*, **73**: 1143, 1998.



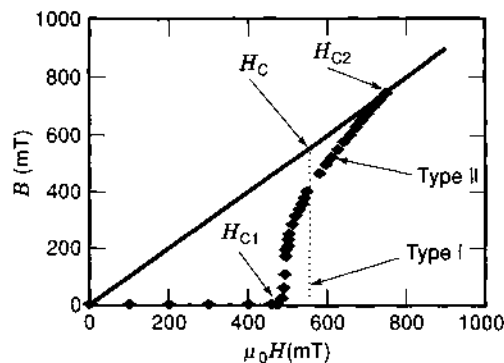
## SUPERCONDUCTIVITY: TYPE I AND II

Superconductors differ from normal metals in some truly remarkable ways. In a normal metal, conduction electrons behave as individual particles that move through the material, interacting with each other and with the array of ions that form the background lattice. When a voltage is applied, the electrons carry current, scatter randomly, show electrical resistance, and obey Ohm's law. Electrons in a Pb wire at 8 K, for example, will be scattered from impurity atoms and dissipate energy. This electron-scattering process can be modeled by treating the electrons as single particles subject to the Coulomb and exchange forces of the neighboring electrons. The individual-particle picture works well, and one can discuss one electron without concern for the events happening to electrons that are many atomic spacings away.

In a superconducting metal, this individual-particle picture does not hold. Rather, the motion of one electron is highly correlated with other electrons in the metal over very large distances. Again using Pb as an example, when a sample is cooled below the superconducting transition temperature  $T_c$  of 7.25 K (1), the electrical resistance drops by more than a factor of  $10^{15}$  in a temperature interval of a few millikelvins. In the course of this transition the motion of the electrons transforms from a single-particle picture to a highly correlated picture. These correlations extend over macroscopic distances and the electrical resistance vanishes. During the superconducting transition, the wave function of the electrons develops a certain rigidity, and the electrons tend to move as a giant molecule rather than as individual particles.

This rigidity is reflected in a special feature of superconductivity called *phase locking*. Electrons, like all elementary particles, have both a particle and a wave character and the wave properties are described by a wave function  $\psi = Ae^{i\phi}$  with both an amplitude  $A$  and a phase  $\phi$ . In the course of the superconducting transition, the phase  $\phi$  of any one superfluid electron becomes tightly locked to the phase of all the neighboring superfluid electrons. This rigidity and phase locking of the wave function leads to most of the remarkable properties of superconductors. It leads to the suppression of scattering, to zero electrical resistance, to the expulsion of magnetic flux from the interior of a superconductor, and to all of the marvelous aspects of quantum interference effects that are so crucial to microelectronic devices. It is important to understand how this rigidity and phase locking come about.

There are two very broad classes of superconductors that are distinguished from one another by the way magnetic flux is distributed inside the material. In one class, called type I, the magnetic flux is totally excluded from the interior of a bulk sample, and the magnetic induction  $B$  is zero everywhere except for a thin layer that is a few hundred nanometers thick near the surface. This thin layer is called the penetration depth  $\lambda$ , and the  $B = 0$  state is called the Meissner state. In this thin layer, surface currents flow to cancel the applied field and give  $B = 0$  in the interior. It is useful to recall here that  $B$  in the interior is the sum of the applied field,  $\mu_0 H$ , plus the field due to circulating currents in the material,  $\mu_0 M$ , where  $H$  is the magnetic field and  $M$  is the volume magnetization. When the external magnetic field reaches a certain critical magnitude called the thermodynamic critical field  $H_c$ , the superconducting



**Figure 1.**  $B$  vs  $\mu_0 H$  plot to illustrate the defining characteristics of type I and type II superconductors.

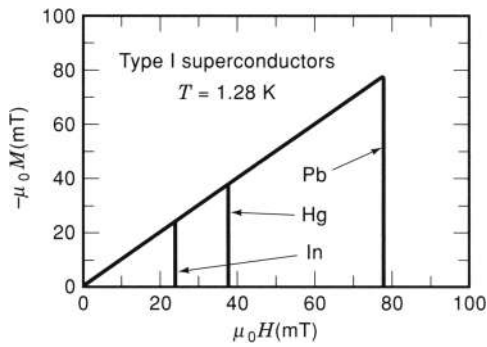
state collapses and the material reverts to the normal state at higher fields. This is illustrated by the dashed curve in Fig. 1.

In the second class of materials called type II, as you increase the magnetic field from zero, the flux is totally excluded at first, just as for type I material. Then, at some characteristic field called  $H_{c1}$ , the magnetic flux begins to enter the material in the form of tiny vortices of magnetic flux, called flux quanta  $\Phi_0$ . Each quantum carries  $2.07 \times 10^{-15}$  Wb of magnetic flux. With further increases in  $H$ , the quantized vortices flood into the material until an upper critical field,  $H_{c2}$ , is reached where the sample again reverts to the normal state. These lower and upper critical fields are illustrated on the diamond points of Fig. 1.

The purpose of this article is to review the microscopic origins of these two types of magnetic behavior for superconductors. We will discuss the features that distinguish type I from type II materials, and describe the way in which type I materials can be transformed into type II materials by alloying. Discussion will include both the classical superconductors and the cuprate superconductors that were discovered in 1986 (2) to have extraordinarily high  $T_c$ . After taking a brief look at a few typical examples of type I and type II behavior, we will review some features of the superconducting wave function that leads to the remarkable properties of superconductors. Then we will return to a more detailed discussion of more subtle aspects of type I and type II behavior. Various practical methods to measure the characteristic fields of  $H_c$ ,  $H_{c1}$ , and  $H_{c2}$  will be explored. Throughout, there will be an attempt to provide useful equations for computation, even though the derivation of the relations is beyond the scope of this presentation. The goal here is to discuss the behavior of these two classes of superconductors using simple physical pictures wherever possible.

### TYPICAL EXAMPLES

Superconductivity is a very common phenomenon. Approximately half the elements in the periodic table and a large number of intermetallic compounds show the effect. Often, it is necessary to go to rather low temperatures and in some cases it is necessary to apply high pressure to transform the material into a superconductor. Although there are exceptions, pure elemental metals tend to be type I superconductors and alloys tend to be type II superconductors.



**Figure 2.** Magnetization curves for Pb, Hg, and In at 1.28 K showing typical values of  $H_c$ .

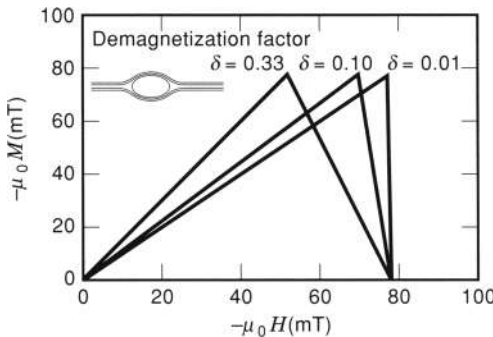
**Low- $T_c$  Type I Superconductors**

In many classical superconductors, such as high-purity Pb, Hg, and In (1,3), the magnetization curves,  $M$  versus  $H$ , exhibit a Meissner ( $B = 0$ ) behavior up to some critical field at which the magnetic flux suddenly collapses into the sample as shown in Fig. 2 for 1.28 K data. Generally speaking, very pure metals that have  $s$ -band or  $p$ -band conduction electrons at the Fermi surface will show this so-called type I behavior. Very pure metals that have a relatively high Fermi velocity  $v_F$  and a relatively low  $T_c$  tend to be type I superconductors.

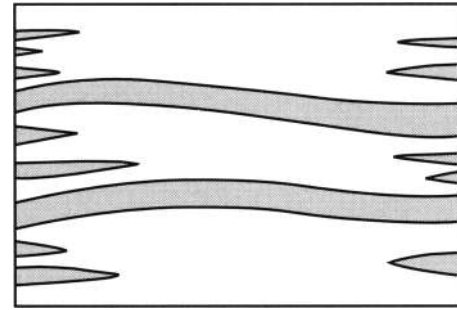
A more detailed look at the magnetization curves reveals that the collapse of flux into a bulk sample occurs over a finite magnetic field interval. When the flux begins to enter the sample, the slope of the line is governed by demagnetizing effects (4), that is, effects associated with the shape of the sample. As shown in Fig. 3, a magnetic field applied to a superconductor in the Meissner state has a larger field at the “equator” of the ellipse than the applied field  $H_a$ . For this elliptical sample, flux begins to enter the sample at  $H_a = (1 - \delta)H_c$  and the transition is complete at  $H_c$ .  $\delta$  is called the demagnetizing factor. Because the free-energy difference between the superconducting and normal state is directly related to  $H_c$ , it is called the thermodynamic critical field curve.

**Intermediate State**

Between  $(1 - \delta)H_c$  and  $H_c$ , the material breaks up into an intermediate state in which both Meissner regions and normal re-



**Figure 3.** Change in the shape of the magnetization curves for a given material as the demagnetizing factor increases. The inset shows an elliptical sample in the Meissner state with the field at the equator larger than the applied field.

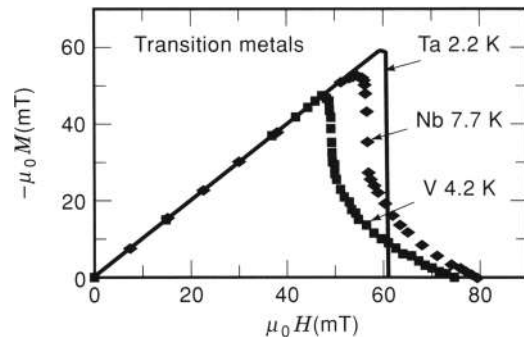


**Figure 4.** A typical magnetic flux distribution in the intermediate state of a type I superconductor as lamella of normal-state regions (shaded area) collapse into the interior. This is for a flat plate with  $H$  perpendicular to the plate.

gions coexist in the sample. When the screening currents that are flowing in the penetration depth region near the surface reach a critical value, normal regions will nucleate at the surface and propagate into the interior of the sample. Usually this happens when the kinetic energy of the electrons in the circulating screening current are a significant fraction of characteristic superconducting energy,  $k_B T_c$ . Here,  $k_B$  is the Boltzmann constant. The detailed shape of the Meissner regions in the intermediate state depends on structure of the material and the shape of the sample. Normal regions nucleate at the surface and often propagate into the interior in the form of lamellar regions as shown in Fig. 4. For a long slender sample with the magnetic field parallel to the long axis of the sample,  $\delta$  is close to zero and the region of the intermediate state is narrow. This behavior is illustrated by the data in Fig. 2. For a sphere,  $\delta$  is  $\frac{1}{3}$ , and the magnetization curve is illustrated in Fig. 3. For a flat plate with the magnetic field perpendicular to the plane of the plate,  $\delta$  is close to 1, and the sample enters the intermediate state at fields far below  $H_c$ . Shoenberg, p. 103 (5), gives a detailed discussion of the intermediate state for spheres and Tinkham, p. 25 (6), gives a detailed discussion for flat slabs.

**Low- $T_c$  Type II Superconductors and the Vortex State**

Magnetization curves for three different transition metals are shown in Fig. 5. Very pure Ta is a type I material (7) with an abrupt collapse of the superconducting state similar to Pb. The other two very pure transition metals, Nb and V, show a much broader transition to the normal state and are called type II



**Figure 5.** Magnetization curves for three transition metals to illustrate both type I (Ta) and type II (Nb and V) behavior.

superconductors (8,9). When very pure metals show type II behavior, they are called intrinsic type II superconductors. The interval between  $H_{c1}$  and  $H_{c2}$  is called the vortex state because in this magnetic field interval, the sample fills with vortices, each carrying one quantum of flux,  $\Phi_0$ . In all of these intrinsic type II materials, the conduction electrons at the Fermi surface are mostly  $d$  band in character, so the Fermi velocity is relatively low. In addition,  $T_c$  is relatively high. The magnetization data in three samples shown here are highly reversible, and therefore the first flux entry is very close to the field at which a vortex is thermodynamically stable in the material,  $H_{c1}$ . For the data shown in Fig. 5, a temperature is chosen so that  $H_c$  is in the vicinity of 50 to 60 mT. This permits an easy comparison of the shapes of the curves. As the temperature is decreased, the ratio of  $H_{c2}/H_c$  typically rises by about 50%, and the slope of the magnetization at  $H_{c2}$  always decreases. Among the transition metals, very pure Ta (9) is an exception in that it is type I. A very small amount of impurity, however, will transform it to a type II material very similar to Nb and V.

**Low- $T_c$  Alloys**

To understand the difference between type I and type II superconductors on a microscopic scale, it is essential to know that there is a characteristic distance, called the coherence distance,  $\xi$ , over which the superconducting wave function can change. In a superconductor, the basic charge-carrying unit in the system is a highly correlated pair of electrons called the Cooper pair. The minimum distance in which the density of superconducting electrons,  $n_s$ , can change from the value in the bulk superconductor to zero in a normal metal is roughly the size of these Cooper pairs. The coherence distance, or the size of the Cooper pairs, is a very important property of a superconductor and typically varies from about 1  $\mu\text{m}$  in Al to as small as 2 nm in the cuprate superconductors. In the transition metals such as Nb and V  $\xi \sim 30$  nm and in the conventional type I materials such as Sn, Pb, Hg, and In  $\xi \sim 500$  nm.

The critical factor governing the shape of the magnetization curves of superconductors is the ratio of the magnetic field penetration depth to the coherence distance,  $\lambda/\xi$ . This ratio is sufficiently important that it is given a symbol of its own,

$$\kappa \approx \lambda/\xi \tag{1}$$

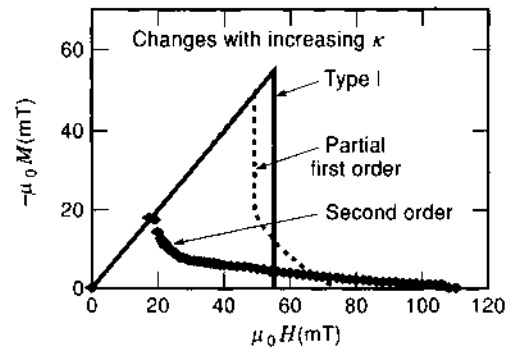
Small- $\kappa$  materials are type I, and large- $\kappa$  materials are called type II. The transition from type I to type II behavior occurs at  $\kappa \approx 1/\sqrt{2} = 0.707$ . In a type I superconductor,  $H_c$  is directly related to the free-energy difference between the superconducting and normal state by the area under the magnetization curve,

$$G_n - G_s = - \int_0^{H_c} \mu_0 M dH = \mu_0 H_c^2 / 2 \tag{2}$$

In type II materials,  $H_c$  retains this definition. The ratio  $\kappa$  from Eq. (1) also is related to the characteristic critical fields on the magnetization curve via

$$\kappa = H_{c2} / \sqrt{2} H_c \tag{3}$$

The range of the correlations among the electrons and hence  $\xi$  in a superconductor can be reduced by shortening the normal-state mean free path of the electrons,  $l$ . Hence, one should ex-

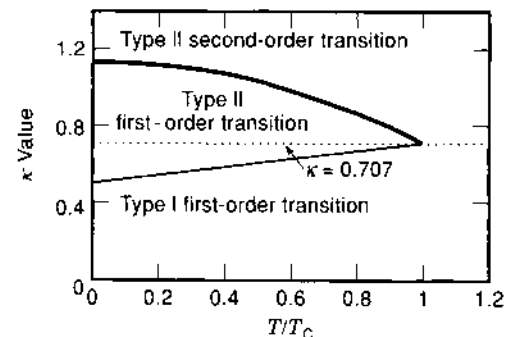


**Figure 6.** Transformation from type I to type II behavior as  $l$  is gradually increased. For  $\kappa$  just greater than 0.707, there is an attractive interaction between vortices and a first-order transition at  $H_{c1}$ . This illustrates the change in shape of magnetization curves as impurities are added.

pect that alloying would reduce both  $l$  and  $\xi$  and would cause a transition from type I to type II behavior if  $\lambda$  remains roughly constant. This was shown to be true, for example, by Kumpf (10), who demonstrated that pure Pb with  $\kappa \sim 0.4$  could be converted to a type II material with  $\kappa \sim 0.84$  with the addition of 2.0 at. % Tl. For this Pb system, it requires a resistivity of  $\rho_n \sim 2 \mu\Omega \cdot \text{cm}$  to change Pb from a type I to a type II superconductor. To convert pure Ta from type I to type II, it requires alloying until  $\rho_n \sim 0.5 \mu\Omega \cdot \text{cm}$ .

The sketch in Fig. 6 shows the changes that occur in the magnetization curves as  $\kappa$  is gradually increased for a material similar to Ta. For  $\kappa$  substantially less than 0.707, the magnetization curves are typical type I with a complete first-order transition to the normal state at  $H_c$ . A first-order transition has a latent heat, but a second-order transition has no latent heat. For  $\kappa$  comparable to  $1/\sqrt{2}$ , there is a partial first-order transition at  $H_{c1}$  with a tail in the magnetization extending out to  $H_{c2}$  as shown by the dashed line in Fig. 6. For  $\kappa$  substantially larger than  $\sqrt{2}$ , the phase transition is second order at  $H_{c1}$ . As shown by Auer and Ullmaier (7), the range of  $\kappa$  showing the type II behavior with a first-order transition at  $H_{c1}$  depends on temperature. Figure 7 is a sketch showing roughly the expected behavior. The horizontal line on Fig. 7 is  $\kappa = 1/\sqrt{2}$ .

Because  $l$  is closely connected to the normal-state resistivity  $\rho_n$ ,  $\kappa$  can be related to  $\rho_n$  and the  $\kappa$  value for the pure metal,  $\kappa_0$ ,



**Figure 7.** The transformation from type I to type II behavior on a  $\kappa$  vs.  $T/T_c$  plot. This qualitatively shows the boundaries of these three different types of phase transitions.

by the useful relation

$$\kappa = \kappa_0 + 7.5 \times 10^5 \gamma^{1/2} \rho_n \quad (4)$$

where  $\gamma$  is the electronic specific constant and  $\rho_n$  is the resistivity. If you use units where  $\gamma$  is in  $\text{erg/cm}^3 \cdot \text{K}^2$  and  $\rho_n$  is in  $\Omega \cdot \text{cm}$ , then the constant is  $7.5 \times 10^5$ . For the extreme dirty or short mean-free-path limit, Hake (11) showed that Ti-16 at. %Mo samples with  $\rho \sim 100 \mu\Omega \cdot \text{cm}$  can be nearly reversible with a  $\kappa$  value of 66.

Among the high-purity  $s-p$  band metals,  $\kappa$  ranges from 0.01 for Al, to 0.15 for Sn, to 0.4 for Pb. Among the high-purity transition-metal superconductors,  $\kappa$  ranges from 0.36 for Ta, to 0.78 for Nb, to 0.90 for V. Nitrogen impurities are particularly good to show the transition from type I to type II behavior because they go into the lattice of Ta as statistically distributed interstitial atoms. Hence, N decreases the electronic mean free path without forming clusters that would substantially increase the pinning of vortices and irreversibility effects. As shown by Auer and Ullmaier (7), Ta is transformed from a type I to a type II superconductor when the sample residual resistivity  $\rho_0 = 585 \mu\Omega \cdot \text{cm}$  and attains a  $\kappa$  value of about 1.5 when  $\rho_0 \sim 2000 \mu\Omega \cdot \text{cm}$ .

### High- $T_c$ Materials

The high-temperature superconductors (HTS) such as  $\text{La}_{1.85}\text{Sr}_{0.15}\text{CuO}_4$ , La(214), are qualitatively different from all of the metals that had been studied before. They are different because the normal state is created by doping an insulator. They also are different because they tend to be rather anisotropic with the charge carriers moving most easily in the  $a-b$  planes of the  $\text{CuO}_2$  sheets. The parent cuprate for La(214) with no Sr doping,  $\text{La}_2\text{CuO}_4$  is not a superconductor and is not a metal. Rather, it is an antiferromagnetic insulator. If, however, part of the trivalent La ions are replaced by divalent Sr ions, holes are created in the copper oxide planes and the material becomes an anisotropic normal metal and a superconductor with  $T_c$  of about 42 K. In these oxide conductors, the carrier mobility along the copper oxide planes is much higher than it is perpendicular to the copper oxide planes, so both the normal conductivity and the superconductivity are quite anisotropic.

In the superconducting state, the magnetization data for the cuprate superconductors show type II behavior with very high  $\kappa$  values, commonly about  $\kappa \sim 100$ . Hence,  $H_{c2}$  is over 100 times larger than  $H_c$ . For many of these cuprates, the penetration depth is about 200 nm, and the coherence distance along the copper oxide planes is about 2 nm. Along the  $c$  axis, the coherence distance is even smaller at  $\sim 0.5$  nm.

### PHYSICAL PICTURES FOR TYPE I AND TYPE II PHENOMENA

Superconductivity is rather special in the field of condensed-matter physics because it is a manifestation of quantum mechanics on a macroscopic scale. If one induces a supercurrent to flow in a superconducting ring, the circulating charge carriers obey the Bohr-Sommerfeld quantization condition even if the diameter of the ring is hundreds of micrometers. In a single atom, such as the hydrogen atom, the angular momentum of the circulating electron is quantized because the wave function of the electron must be single valued going around the atom by  $2\pi$ . This quantization of angular momentum in the hydrogen

atom is then reflected in a quantization of the magnetic moment in units of Bohr magnetons. The same kind of quantization occurs for superconducting electrons circulating in a large ring. Because the electrons are phase locked, the wave function must be single valued and the Bohr-Sommerfeld quantization condition,  $\oint p dq = n$ , is obeyed. Here,  $p$  is the momentum,  $dq$  is the path element, and  $n$  is an integer. This creates quantized circulating currents, and the resulting flux also is quantized in units of  $\Phi_0 = h/2e$ .

Even though quantum mechanics is fundamental to understanding superconductivity, there are some simple pictures that enable the beginner to visualize where the electrons or holes are and how they interact. The goal of this section is to present some of the vocabulary and ideas of the Bardeen-Cooper-Schrieffer (BCS) (12) theory in a way that one can picture the basic behavior of type I and type II superconductors.

#### Idea 1: All Superconductors Look Alike

There is a great deal of similarity in the physical properties of superconductors. If the thermodynamic critical field is plotted as a function of temperature or the superconducting energy gap in the excitation spectrum,  $\Delta$ , as a function of temperature, the curves have the same functional form and the values scale with the transition temperature. The ratio of  $H_c(T=0)/T_c$  is always about 10 mT/K, and the ratio of  $\Delta(T=0)/k_B T_c \sim 1.8$  for the low  $T_c$  superconductors. Because superconductors are so similar, there is a reasonable expectation that the details of the metal are not terribly important in the basic explanation of superconductivity. A rather general and simple theory may explain the effect.

#### Idea 2: BCS Theory

In the development of a theory of superconductivity, BCS constructed the superconducting ground state by taking special combinations of normal-state wave functions. At the superconducting transition, the space dependence of the wave functions do not change, but rather the occupation probability of a given state changes. To create the superconducting wave function, the electrons are allowed to exchange phonons, thus scattering around the Fermi surface and coherently mixing the normal-state wave functions. Instead of the random occupation of states at the Fermi surface that occurs in a normal metal, the system can gain energy via phonon exchange if the electronic states are occupied in pairs. This pairing increases the amount of phonon exchange that can occur and each phonon exchange lowers the energy a bit. To define terms more explicitly, pair occupation means that if the state with momentum  $k$  is occupied, then the state with momentum  $-k$  also is occupied. Similarly, if the state with  $k$  is empty, then the state with  $-k$  also is empty. The pairs are chosen to have equal and opposite momentum because, by symmetry, this choice gives the maximum number of final states for the electron-phonon scattering and maximizes the phonon exchange. As the metal undergoes the superconducting transition, it is the probability of occupation that changes. It changes from a random occupation to a pair occupation.

In BCS theory, three essential variables are used. First, the normal-state density of states at the Fermi surface,  $N(0)$ , is a measure of the number of electrons participating. The number of electrons that can take advantage of phonon exchange is determined by  $N(0)$ . Second, the Debye energy  $\omega_D$  is a

measure of the range of phonons available for the electron-phonon exchange process. Third, the strength of the electron-phonon interaction,  $V$ , is a measure of the amount of energy the electrons gain in a phonon exchange. Within the theory, the superconducting transition temperature  $T_c$  will be calculated to be

$$k_B T_c \approx \hbar \omega_D e^{-1/M(0)V} \tag{5}$$

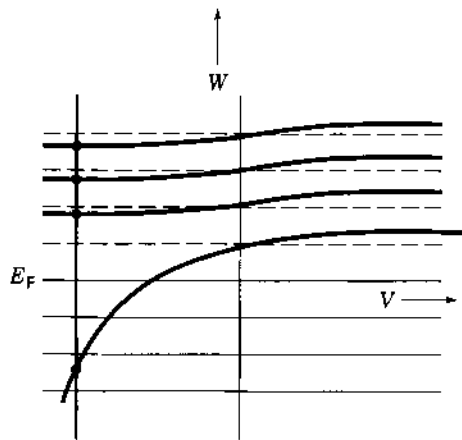
This provides a connection between the transition temperature and the three variables in the theory. If, in addition, one works out the minimum energy to create an excitation out of the superfluid ground state by disrupting one pair of electrons, at  $T = 0$ , this turns out to be

$$\Delta_0 \approx 1.8 k_B T_c \tag{6}$$

Disrupting a pair simply means preventing that pair of states from participating in the coherent phonon exchange. This minimum energy to create an excitation in a superconductor,  $\Delta$ , is often called the order parameter or the energy gap, and typically it is about 1 meV for a  $T_c$  of 7 K. The value of  $\Delta$  decreases from the  $T = 0$  value,  $\Delta_0$ , as the temperature rises and goes to zero at  $T_c$ . Our goal here is to provide a physical picture to go with these equations.

**Idea 3: The Cooper-Pair Problem**

A first step in the development of the BCS theory was the Cooper-pair problem (13). Cooper showed that there is a fundamental instability in an electron gas if one introduces an attractive interaction that scatters electrons around the Fermi surface. In this problem, you start with one pair of electrons outside a Fermi sea. Without the electron-phonon interaction, the energies available to the electrons are the usual free-electron energies above the Fermi energy  $E_F$ , as shown by the dotted lines in Fig. 8. The solid horizontal lines are filled states below the Fermi sea. When the electron-phonon interaction is introduced, the new energies for the perturbed system,  $W$ , are represented by the heavy solid lines. Here they are plotted as a function of the strength of the interaction,  $V$ . Note that if  $V$



**Figure 8.** Energy-level diagram for the Cooper-pair problem. Here  $W$  is the new energy of the pair state and  $V$  is the strength of the perturbing interaction. Note that for large and negative  $V$ , one state falls far below the Fermi energy.

is positive or repulsive, all the energies are pushed up a bit above the unperturbed dotted line values. Note also that if  $V$  is negative or attractive, all of the energies are pushed down a bit except the level just above  $E_F$ . This level is pushed far below  $E_F$ . The splitoff of this state far below  $E_F$  means that a free-electron gas with random occupation of states is unstable to the formation of pairs when an attractive interaction is introduced. The energetically most favored combination of electrons for this phonon exchange occurs when the pair has zero center-of-mass momentum (electrons have equal and opposite momentum), hence, giving the lowest energy. Experiments have shown that the pairs of electrons also are known to have opposite spin. Presumably this antiparallel spin arrangement arises from a quantum-mechanical effect called exchange forces. Electrons with opposite spins are closer together than electrons with parallel spins so they can take better advantage of the electron-phonon interaction. Hence, pairs with opposite spin and momentum have the lower energy. If some other type of interaction were to cause pairing, then a parallel spin arrangement is possible.

A special feature of the Cooper pair problem is that the normal-state wave functions that are being mixed by the electron-phonon interaction all add in phase at some point in real space to form a wave packet. This packet is the so-called Cooper pair and by the summation of occupied states, it has a real-space extent of the coherence distance,

$$\xi = \hbar v_F / \pi \Delta_0 = 0.18 \hbar v_F / k_B T_c \tag{7}$$

Equation (7) shows that the coherence distance, or the size of the Cooper pairs, is directly related to the ratio of  $v_F/T_c$ . Because the  $s - p$  band metals tend to have relatively high  $v_F$  and low  $T_c$ , they would be expected to have large coherence distances and hence tend to be type I superconductors. The basic charge-carrying unit in the circulating current around a superconducting vortex in a type II material in the vortex state is the Cooper pair.

**Idea 4: BCS Theory and Pair-Pair Correlations**

The Cooper-pair problem is only part of the picture because it describes just one pair of electrons outside a Fermi sea. The total problem must deal with all of the electrons and must deal with their correlated motion. If typical numbers are put into Eq. (7), Al has  $\xi \sim 1600$  nm, Sn has  $\xi \sim 230$  nm, Pb has  $\xi \sim 83$  nm, and Nb has  $\xi \sim 40$  nm. Given the normal density of electrons in a metal, there are thousands to millions of pairs occupying the space of any single pair. Hence there is a great deal of pair-pair overlap. These highly interacting and overlapping Cooper pairs provide a physical picture for the origin of phase locking over macroscopic distances.

The high-temperature superconductors are a very special case because the coherence distance is so short. Typically,  $\xi \sim 2$  nm so the overlap of Cooper pairs is much less in this class of materials and the phase locking is less strong. For HTS, the number of pairs overlapping any one pair is measured in dozens rather than thousands to millions. This leads to a less rigid vortex lattice and a greater susceptibility to flux creep.

The BCS problem is similar in some ways to a variational calculation in an undergraduate quantum-mechanics course. One guesses a form for the wave function with the pair occupation number as an adjustable parameter. Next one writes an

expression for the free energy of the system in terms of this trial wave function and minimizes the energy as a function of pair occupation. In both the Cooper and BCS problems, a paired trial wave function is selected in order to maximize the final states available for electron-phonon scattering. One result of the theory is that it is energetically favorable to mix normal-state wave functions within an energy band that is  $\Delta$  wide near  $E_F$ . All of the pairs share the same states so there is an energy advantage for correlated motion of many pairs to make most efficient use of the states available for electron-phonon exchange. Pair-pair correlations are critical because they provide the fundamental mechanism to propagate phase coherence over long distances.

To create an excitation from the superconducting ground state, one of the pair states is simply disrupted. If a pair state is broken, say by injecting an electron into one of the two states of a pair in the metal, then other pairs cannot use that channel for phonon exchange. That pair state is removed from the coherent phonon exchange for all of the other electrons. This raises the energy of all the electrons in the ground state. In computing the ground-state energy and the excitation energies, one has to give up the single-particle picture and go to a highly correlated many-paired electron picture in which the disruption of one pair changes the energy of all the pairs in the ground state.

#### Idea 5: Phase Locking and Rigidity of the Wave Function

A central feature of superconductivity is that the electrons phase lock into a many-electron ground-state wave function that has a substantial amount of rigidity. If the system is disturbed, it responds as a giant unit rather than responding as individual particles. In the Cooper problem, a pair of electrons is formed by mixing normal-state wave functions so that they all add in phase at some point in space. In the BCS problem, pair-pair correlations play a central role because all of the pairs are using the same states for electron-phonon scattering. This leads to a highly correlated ground-state superfluid that extends over macroscopic distances. The ground state is somewhat like a giant macromolecule in that the superfluid is phase locked and responds as a rigid unit. A common analog in chemistry would be a benzene molecule in which the electrons in this benzene ring respond as a rigid unit to small stimuli. In a superconducting Pb wire 1 m long, the electrons at one end are phase locked to the electrons at the other end.

Pairs, of course, can be disrupted by many processes: by thermal ( $k_B T$ ) excitations, by electromagnetic absorption, or by electron injection. In all of these cases the disrupted electrons or excitations behave just like normal-state electrons. The ground-state electrons can be thought of as a superfluid with density  $n_s$ , with a condensation energy or pairing energy  $\Delta$  per pair. The excitations out of the ground state can be thought of as a normal fluid with density  $n_n$ . In this framework the sum of these probabilities must add to one,  $n_s + n_n = 1$ . This picture is quite analogous to the two-fluid model originally proposed by London (14). As the temperature rises, the incoherent phonon scattering becomes larger and overwhelms the coherent phonon exchange and the material reverts to the normal state.

#### Zero Electrical Resistance

Electrical resistance in a material such as Pb disappears below the superconducting transition temperature because a significant fraction of the electrons transform into a coherent, phase-locked, superfluid ground state having considerable rigidity in

the wave function. In a normal metal, such as Pb above  $T_c$ , the electrons are relatively uncorrelated and behave as individual particles. Electrical resistance arises because individual electrons scatter from impurity atoms, phonons, or some other excitation in the metal as random or thermal events. In a superconductor, electrons are locked together in a giant rigid wave function and individual scattering is essentially eliminated. The reason that the transition is so sharp is that there is such overkill in the pair-pair correlation. With millions of pairs overlapping the region of any one pair, one can get substantial pair-pair correlation even if only 0.1% of the electrons are in the superfluid state. This can occur within a temperature interval of a millikelvin.

#### Meissner Flux Exclusion in Type I Materials

In thinking about the ability of a type I superconductor to push magnetic flux out of the interior to give  $B = 0$ , it is important to remember that the free-energy difference per atom to push the flux out is very small. The value of  $G_n - G_s$  given by Eq. (2) can be evaluated to be about  $10^{-8}$  eV per atom. This is tiny compared with most electronic processes, which are about 1 eV. Hence, it does not take much energy to exclude the flux.

There are several competing energies in this problem. If a magnetic field penetrates the interior of a sample where the paired superfluid density is high, then there is a tendency for the spins to align with the field and thus break pairs. This would raise the free energy. This rise in free energy would be prevented if circulating currents are created at the surface of the sample to cancel the magnetic field to give  $B = 0$ . There is, of course, a kinetic energy cost in creating these circulating supercurrents. For small  $\kappa$  and  $H < H_{c1}$ , it is energetically more favorable for the material to develop a supercurrent at the surface and provide a  $B = 0$  condition in the interior than it is to lose the condensation energy associated with breaking pairs as the spins align with the applied magnetic field. If, however,  $\kappa$  is greater than  $1/\sqrt{2}$  and  $H > H_{c1}$ , it is energetically favorable for quantized vortices to enter the material.

The time dependence of the response of the electrons to an applied magnetic field is governed by the plasma frequency. With the application of an external field, the superconducting electrons respond as a coherent plasma, and the screening length for magnetic fields is given by

$$\lambda^2 = m^* / \mu_0 n_s e^2 \quad (8)$$

where  $m$  is an effective mass and  $e$  is the charge on the electron. For very pure metals,  $\lambda \sim 50$  nm and for the HTS materials,  $\lambda \sim 170$  nm. Because  $\lambda$  is governed by the superfluid density, it does not vary from superconductor to superconductor as strongly as  $\xi$ , which is governed by the ratio of  $v_F / k_B T_c$ .

#### Ginzburg-Landau Equations

Very early in the development of the theory of superconductivity, Ginzburg and Landau (15) developed a very general and yet very powerful formulation of the problem. They assumed that there is an order parameter that behaves much like a wave function, and they further assumed that this wave function  $\psi$  is related to the local density of superconducting electrons by  $|\psi|^2 \sim n_s$ . The free energy is then written as the sum of a kinetic

energy term, a potential energy term, and a magnetic term:

$$G_s - G_n = \frac{1}{2m} \left| \left( \frac{\hbar}{i} \nabla - eA \right) \psi \right|^2 + \alpha |\psi|^2 + \frac{\beta}{2} |\psi|^4 + \frac{\hbar^2}{8\pi} \quad (9)$$

Here, the potential energy term is a power-series expansion in  $|\psi|^2$  and  $\hbar^2/8\pi$  is the magnetic energy term. Minimizing this free energy with respect to the order parameter gives the famous Ginzburg–Landau (GL) equation:

$$\alpha \psi + \beta |\psi|^2 \psi + \frac{1}{2m^*} \left( \frac{\hbar}{i} \nabla - eA \right)^2 \psi = 0 \quad (10)$$

Excellent discussions of solutions of these equations are given by Fetter and Hohenberg (16), Tinkham (6), and de Gennes (17).

A central factor determining whether a material is a type I or type II superconductor is the boundary energy between a superconducting region and a normal region. If a superconductor–normal-metal boundary has a positive surface energy, then the system will minimize the boundary area and a Meissner solution will be expected. If the superconductor–normal-metal boundary has a negative surface energy, then it will be expected that the material will be unstable against a breakup into small domains in order to maximize the amount of superconductor–normal-metal boundary area. In the original paper (15), it was shown by numerical calculations that for Eq. (10), the onset of negative surface energy occurs when the ratio of two characteristic lengths, the penetration depth and the coherence distance, is equal to 0.707 or

$$\kappa \approx 1/\sqrt{2} \quad (11)$$

This criterion then specifies whether a superconductor is type I or type II.

### Vortex State

Building on the GL result, Abrikosov (18) predicted in 1957 that flux entering a type II superconductor is in the form of quantized vortices. A quantized vortex with  $\Phi_0$  of flux is the minimum unit of flux that can enter. He showed that if  $\kappa > 1/\sqrt{2}$ , then a regular array of flux tubes or vortices would form in the interior of the material. The stable state is a triangular array so the area of a flux tube would be

$$\frac{\sqrt{3}}{2} a_0^2 = \frac{\Phi_0}{B} \quad (12)$$

where  $a_0$  is the lattice spacing of the triangular lattice. More detailed calculation shows that vortices are first stable in a superconductor at

$$H_{c1} = \frac{H_c}{\sqrt{2}\kappa} \ln \kappa \quad (13)$$

If there are no surface barriers to flux entry, vortices will nucleate and move into the interior above  $H_{c1}$ . As the applied field increases, the vortices crowd closer together. At high magnetic field, when the cores of the vortices begin to overlap, the sample

goes normal. This occurs at

$$H_{c2} = \frac{\Phi_0}{2\pi\xi^2} = \sqrt{2}\kappa H_c \quad (14)$$

Combining Eq. (13) with Eq. (14) gives a convenient rule of thumb that  $H_{c1}H_{c2} \approx H_c^2 \ln \kappa$ .

### Large- $\kappa$ Case: Repulsive Interaction Between Vortices

For the case of  $\kappa \gg 1/\sqrt{2}$ , the Ginsburg–Landau equations have relatively simple solutions and there are rather good physical pictures to describe the behavior of vortices. A reasonably good model to describe the vortex is a normal core of radius  $\xi$  with circulating currents sufficient to give one quantum,  $\Phi_0$ , of flux. The superconducting order parameter  $\Delta$  is zero in the center of the core and rises toward the bulk value with a characteristic length  $\xi$ . To estimate the free energy per unit length of the vortex,  $g_v$ , the free energy per unit volume,  $\mu_0 H_c^2/2$ , can be multiplied by the area of the core,  $\pi\xi^2$ . A more accurate calculation (6) gives

$$g_v \approx (\mu_0 H_c^2/2)(4\pi\xi^2) \ln \kappa \quad (15)$$

The magnetic field around the vortex is given by (6)

$$h(r) = \frac{\Phi_0}{2\pi\lambda^2} K_0(r/\lambda) \quad (16)$$

where  $K_0$  is the zeroth-order Hankel function of imaginary argument. The force on a quantized vortex in the presence of a current density  $J$  is (6)

$$f_p = J\Phi_0 \quad (17)$$

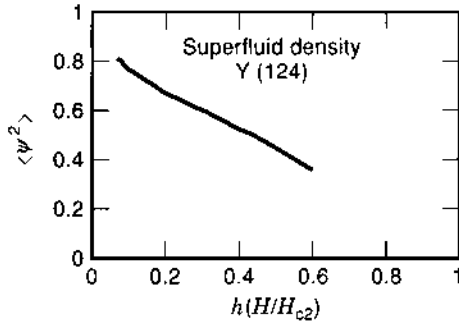
where  $J$  can be either the current density caused by the circulating current of other vortices or a transport current density that is externally applied. In this large- $\kappa$  regime, the force between vortices is repulsive and the Abrikosov theory (18) shows that the vortices will arrange themselves in a triangular array. As the field rises above  $H_{c1}$ , the vortices flow into the interior under the influence of the magnetic pressure of the applied field.

The high- $T_c$  cuprate superconductors normally have a coherence distance in the  $a - b$  plane on the order of  $\xi_{ab} \sim 2$  nm and a penetration depth on the order of  $\lambda \sim 200$  nm so the  $\kappa$  values are in the range of  $\kappa \sim 100$ . In this extreme type II limit, the size of the Cooper pairs is very small compared to the vortex size and several simplifying assumptions can be made in the development of models to describe the magnetization curves.

The Hao and Clem model (19) for the magnetization curves of high- $T_c$  superconductors is typical of approaches that can be used. They developed a variational method in which the trial wave function is written as

$$\psi = \frac{\rho}{\sqrt{\rho^2 + \xi_0^2}} \psi_\infty \quad (18)$$

where  $\rho$  is the distance from the core of the vortex, and both  $\xi_v$  and  $\psi_\infty$  are adjustable variables in the trial wave function. They start with the free energy including the core energy and minimize with respect to the trial variables of  $\xi_v$  and  $\psi_\infty$ . The



**Figure 9.** Superfluid density as a function of magnetic field for Y(123). These data show that the superfluid density falls linearly with magnetic field and approaches zero as  $H$  goes to  $H_{c2}$ .

magnetization curves are found to scale to a universal function on a  $\mu_0 M/\sqrt{2}H_c$  vs  $H/\sqrt{2}H_c$  plot. Both  $Y_1Ba_2Cu_3O_{7-\delta}$  and  $Y_1Ba_2Cu_4O_{8-\delta}$  are found to obey the Hao-Clem model very well (19). In addition, it was found that the superfluid density averaged over one vortex unit cell falls linearly as  $H/H_{c2}$  for fields greater than  $0.3H_{c2}$  over the entire range where there is thermodynamic reversibility and measurements can be made (20). Figure 9 shows that  $n_s \sim |\psi|^2$  is linear in  $H/H_{c2}$  for this high- $T_c$  Y(123) material.

### Surface Superconductivity

The surface of the superconductor modifies the potential for the superfluid electrons, and superconductivity will persist to fields above  $H_{c2}$  in a narrow layer near the surface. With the applied field parallel to the flat surface (17), Ginzburg-Landau theory with plane surface boundary conditions predicts that a superconducting layer about a coherence distance thick will be present up to  $H_{c3} = 1.7H_{c2}$ . Experiment verifies that this basic idea is correct. For the case of pure Nb, the measured ratio of  $H_{c3}/H_{c2}$  for pure Nb varies from 1.78 at low temperatures to 1.70 at  $T/T_c = 0.9$  (21). At higher temperatures,  $H_{c3}/H_{c2}$  approaches 1.0 at  $T_c$ . Altering of the surface condition, say, by a normal metal coating will suppress or destroy this surface state.

### Small- $\kappa$ Case: Attractive Interaction Between Vortices

For the case in which  $\xi$  is comparable to  $\lambda$ , there can be overlap of the vortex cores and there can be an attractive interaction between vortices. This effect was established experimentally by Essmann and Trauble (22) with experiments in which they decorated the surface of Nb with Fe spheres about 4 nm in diameter. To perform these experiments, typically an array of vortices is trapped in a coin-shaped Nb sample by applying a magnetic field above  $H_{c1}$  and then turning the field off. A “smoke” of Fe was then created by evaporating Fe metal in an atmosphere of a few Torr of He gas. The tiny Fe particles that are created follow the flux lines to the point on the surface where the core of the vortex emerges. Once the Fe particles touch the Nb, they stick very strongly. To image the vortex lattice, the Fe is stripped off the surface by a graphite replication technique and viewed in a transmission electron microscope. For the initial flux entry in the first-order transition of  $H_{c1}$ , illustrated by the dashed curve of Fig. 6, it is found that there is a two-phase region. There are clusters of a few hundred vortices all on a triangular lattice and

separated by about 200 nm. Between these clusters there are Meissner-like or vortex-free regions. In this two-phase region, the spacing of the vortices is independent of magnetic field. As the magnetic field increases, the sample fills with vortices having 200 nm spacing. The magnetization curve in this region is linear with a slope governed by the demagnetizing factor. An abrupt change in the slope of the magnetization curve occurs when the sample is just filled with vortices at the lattice spacing governed by the attractive interaction between vortices. This characteristic field is denoted by  $B_0$ . At higher magnetic fields, the vortices are pushed closer together and the magnetization curve is similar to the repulsive force case for high- $\kappa$  materials. Auer and Ullmaier (7) performed a very systematic study of these same effects, as  $\kappa$  in Ta is systematically increased from type I to type II behavior by alloying with N.

To analyze quantitatively the transition from type I to type II behavior by alloying, it is useful to focus on the connection between the coherence distance and the normal-state electronic mean free path  $l$ . With small additions of impurity,  $\xi$  is given by

$$1/\xi \approx 1/\xi_0 + 1/l \quad (19)$$

where  $\xi_0$  is the intrinsic or clean limit of  $\xi$ . At higher impurity concentrations,  $l$  becomes comparable to  $\xi_0$  and the diffusion limit is more appropriate. In this regime

$$\xi = \sqrt{\xi_0 l} \quad (20)$$

### Experiments to Determine $H_{c1}$ and $H_{c2}$

Many factors can lead to errors in determining both  $H_{c1}$  and  $H_{c2}$ . The lowest field for which a vortex is thermodynamically stable in a superconductor is defined to be  $H_{c1}$ . It is a difficult quantity to measure because thermodynamic equilibrium is not achieved easily at fields close to  $H_{c1}$ . The most common error arises from the presence of surface barriers to flux entry into the sample. For a cylindrical superconductor with a radius of curvature much larger than both  $\lambda$  and  $\xi$  (17, p. 79), the applied field must be larger than  $H_{c1}$  before a vortex will nucleate. As the field increases from zero, Meissner screening currents flow in the superconductor within a distance  $\lambda$  of the surface. When a vortex starts to nucleate at the surface and move into the interior of the sample, an image vortex develops to pull the vortex back toward the surface. The competition between the image force pulling the vortex toward the surface and the Meissner currents pushing the vortex into the interior creates a surface barrier to flux entry. Often the flux-entry field is comparable to  $H_c$ . To overcome this effect, the surface needs to be rough on the scale of the penetration depth or the surface needs to be coated in some way to suppress the surface barrier to zero.

The highest field for which a vortex is thermodynamically stable in a superconductor is defined to be  $H_{c2}$ . For a sample that obeys the Ginzburg-Landau theory, there is a sharp change in slope at the second-order phase transition of a reversible magnetization curve that identifies  $H_{c2}$ . This is the most reliable measurement of  $H_{c2}$ . Measurements of the electrical resistivity is less reliable because it depends on the motion or depinning of vortices and often is not a measure of the point of thermodynamic stability of a vortex. For a clean classical



superconductor such as Nb, the electrical resistance goes to zero at very nearly the same field that the magnetization goes linearly to zero, and both resistivity and magnetization methods can be used to determine  $H_{c2}$ . For high- $T_c$  materials, the situation is more complicated. Because  $\xi$  is so small, fluctuation effects smear out the normal-metal to superconductor transition so that it is typically 3 K wide. Both the magnetization curves and the electrical resistivity are greatly rounded even for a very perfect sample. For this case, it is better to move to lower fields at which fluctuations are negligible and use a fit of the  $M$  vs  $H$  data to the Hao-Clem theory to determine  $H_{c2}$ .

## BIBLIOGRAPHY

1. D. L. Decker, D. E. Mapother, and R. W. Shaw, Critical field measurements on superconducting lead isotopes, *Phys. Rev.*, **112**: 1888–1897, 1958.
2. G. Bednorz and K. A. Muller, Possible high  $T_c$  superconductivity in the Ba-La-Cu system, *Z. Phys. B*, **64**: 189–197, 1986.
3. D. K. Finnemore and D. E. Mapother, Superconducting properties of tin, indium, and mercury below 1 K, *Phys. Rev.*, **140**: A507–A518, 1965.
4. R. B. Goldfarb, Demagnetizing factors. In J. Evetts (ed.), *Concise Encyclopedia of Magnetic and Superconducting Materials*, Oxford: Pergamon, 1992.
5. D. Shoenberg, *Superconductivity*, Cambridge: Cambridge University Press, 1960.
6. M. Tinkham, *Introduction to Superconductivity*, 2nd ed., New York: McGraw-Hill, 1996.
7. J. Auer and H. Ullmaier, Magnetic behavior of type II superconductors with small Ginsburg-Landau parameter, *Phys. Rev.*, **7**: 136–145, 1973.
8. D. K. Finnemore, T. F. Stromberg, and C. A. Swenson, Superconducting properties of high purity Nb, *Phys. Rev.*, **149**: 231–243, 1966.
9. J. J. Wollan et al., Phase transition at  $H_{c1}$  for superconducting Nb and V, *Phys. Rev.*, **10**: 1874–1880, 1974.
10. U. Kumpf, Magnetization curves for type II superconductors with small Ginsburg-Landau parameter, *Phys. Status Solidi*, **44**: 829–843, 1971.
11. R. R. Hake, Mixed state paramagnetism in high field type II superconductors, *Phys. Rev. Lett.*, **15**: 865–868, 1965.
12. J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Theory of superconductivity, *Phys. Rev.*, **108**: 1175–1204, 1957.
13. L. N. Cooper, Bound electron pairs in a degenerate Fermi gas, *Phys. Rev.*, **104**: 1189–1190, 1956.
14. F. London, *Superfluids*, New York: Wiley, 1950, Vol. 1.
15. V. L. Ginsburg and L. D. Landau, Regarding a theory of superconductivity, *Zh. Expt. Teor. Fiz.*, **20**: 1064–1072, 1950; **20**: 1064–1077, 1950.
16. A. L. Fetter and P. C. Hohenberg, Theory of type II superconductors. In R. D. Parks (ed.), *Superconductivity*, Marcel Dekker: New York, 1969.
17. P. G. de Gennes, *Superconductivity of Metals and Alloys*, New York: W. A. Benjamin, 1966.
18. A. A. Abrikosov, Magnetic properties of superconductors, *Zh. Expt. Teor. Fiz.*, **32**: 1442–1450, 1957 (*Sov. Phys. JETP*, **5**: 1174–1182, 1957).
19. Z. Hao and J. R. Clem, Limitations of the London model for the reversible magnetization of type II superconductors, *Phys. Rev. Lett.*, **67**: 2371–2373, 1991.
20. J. Sok et al., Reversible magnetization, critical fields, and vortex structure in grain aligned YBa<sub>2</sub>Cu<sub>4</sub>O<sub>8</sub>, *Phys. Rev.*, **51**: 6035–6040, 1992.
21. J. E. Ostenson, J. R. Hopkins, and D. K. Finnemore, Surface superconductivity in Nb and V, *Physica*, **55**: 502–506, 1971.
22. U. Essmann and H. Trauble, Direct observation of individual flux lines in type II superconductors, *Phys. Lett.*, **A24**: 526–527, 1967.

D. K. FINNEMORE  
Iowa State University

# SUPERCONDUCTORS

## INTRODUCTION

In February 2000 C. Pao and H. A. Farach tabulated the transition temperatures and structure types of about 200 classical superconducting compounds (1). In this section we focus on those superconductors in commercial production. Some of that production is quite substantial. In the year 2000 the global demand for Nb-Ti for MRI applications alone was approximately 1000 tons (including the dominant copper fraction), or 60,000 km to 70,000 km. The estimated value of that demand was about \$ 50M. In this introduction, two potentially important superconductors not covered in the rest of this section are also introduced.

### NbN Including NbTiN

Nitride and carbonitride superconductors date back to the explorations of Meissner's group in the 1930s. During the Second World War, Aschermann and coworkers found that niobium nitride had a  $T_c$  of 15 K (2), the first time the liquid helium region had been surpassed. As in the case of A15 compounds, the highest critical temperatures are for compounds with numbers of valence electrons per atom closest to 4.7 and 6.7. (See the article on A15 SUPERCONDUCTORS in this section.)

For niobium carbonitride Matthias found a  $T_c$  of 17.8 K (3). The application of these materials in bulk form was hindered by their metastability, which required them to be made at high temperature and then quenched. Successful thin films, however, could be made by a number of deposition routes, including chemical vapor deposition and reactive evaporation as well as reactive sputtering. Cavalier and coworkers at Westinghouse recorded perhaps the highest ever critical current densities by measuring 60 nm NbN whiskers (4). Of more typical thickness are the multilayers produced by Gray and coworkers at Argonne National Laboratory (5). These values closely match the APC Nb:Sn thin film values of Dilscherich at Lawrence Berkeley National Laboratory (6) and approach the YBCO microbridge values of Fulton and coworkers at Los Alamos National Laboratory as shown in Fig. 1. NbN/AIN multilayers and NbN films are extremely radiation-hard. NbN has been a focus of Josephson-junction fabrication research. It is much easier to fabricate digital circuits using NbN than using HTS YBCO. Even though it has a much lower  $T_c$  than YBCO, it can still be operated with a 10 K (small pulse-tube or Stirling cooler) cryocooler, which is preferable to the 4 K cryocooler that would be required for Nb-based junctions. For SIS devices, the larger bandgap (1.2 THz) of NbN than that of Nb makes it a better candidate for frequencies above 700 GHz. High-gain dc SQUIDS have been fabricated using NbN nanobridges for biomedical magnetic measurements (7). NbN has also proved a viable option for instruments designed to measure molecular spectra for astronomical applications as well as in remote sensing of the atmosphere in the terahertz regime (8). Recently there has also been work on NbTiN, which is a promising material for use in the tuning circuits of Nb-based SIS mixers for operating frequencies above the gap frequency of Nb (approximately 700 GHz), achieving low noise at 850 GHz. NbN and NbTiN continue to be a very active area of research.

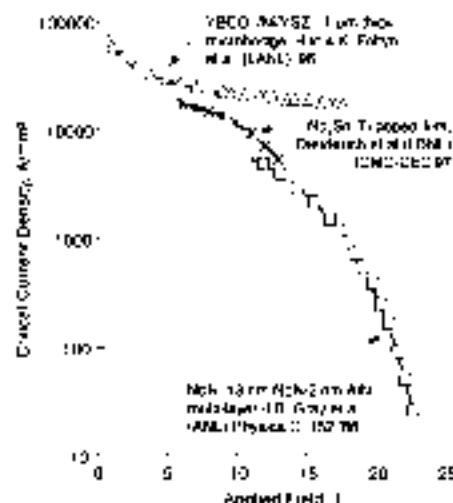


Figure 1. A comparison of the highest recorded Nb:Sn and YBCO critical currents with those of the multilayers of Gray et al.

### Chevreil-Phase Superconductors

Chevreil phases include all compounds with the formula  $Me_1 \times 4Me_2X_8$ , where M is a metal and X is a chalcogen. They are not so much HTS or LTS superconductors as high-field superconductors, with  $PbMo_6S_8$  having an upper critical field of 60 T in bulk samples. In wire form, however, the highest critical fields have been closer to 35 T at 4.2 K, and the wires have low critical currents (9). Nevertheless, they have a unique application to high-field studies and are under active study by a number of research groups.

## BIBLIOGRAPHY

1. C. P. Pao and H. A. Farach, Tabulations and correlations for transition temperatures of classical superconductors, *J. Superconductivity*, **13**(1): 47-60, 2000.
2. G. Aschermann, et al., Supraleitfähigkeit Verbindungen mit extrem hohen Sprungtemperaturen (NbH und Nb), *Phys. Z.* **42**: 349-60, 1941.
3. B. T. Matthias, Transition temperatures of superconductors, *Phys. Rev.* **92**: 74-876, 1953.
4. R. Cavalier et al., Dimensional effects on current and field properties in NbN films, *IEEE Trans. Magn.*, **17**: 573-576, 1991.
5. K. E. Gray, et al., Experimental study of the ultimate limit of flux pinning and critical currents in superconductors, *Physica C*, **152**: 445-455, 1988.
6. D. B. Dietrich, et al., High critical current densities in Nb:Sn films with engineered microstructures: artificial pinning microstructures, *Adv. Cryo. Eng. (Materials)* **44b**: 951-958, 1998.
7. A. Irie, K. Hamazaki, and T. Yamashita, Fabrication of high gain DC-SQUID magnetometer with NbN nanobridges, *Electronics Commun. Japan (Electronics)* **73**(7): 88-96, 1990.
8. E. Gerschl, et al., NbN hot electron bolometric mixer with intrinsic receiver noise temperature of less than five times the quantum noise limit, *2000 IEEE MTT-S International Microwave Symposium*, Vol. 2, pp. 1907-19, 2000.
9. N. Chagnon, et al., Irreversibility line and granularity in Chevreil phase superconducting wires, *J. Appl. Phys.* **84**: 2181-2192, 1998.

## A15 SUPERCONDUCTORS

When discussing technical superconductors one distinguishes between high-temperature superconductors and low-temperature superconductors. The characteristic temperature is the critical temperature,  $T_c$ . Above this temperature superconductivity ceases. When thinking about "high" temperatures one typically addresses temperatures around the point of liquification of nitrogen (77 K); when considering "low" temperatures, values in the region of the boiling point of liquid helium (4.2 K) are of interest.

The low-temperature superconductors can be further subdivided. For example, there are solid-solution superconductors and  $\beta$ -tungsten superconductors. In this article, we consider the latter. Another name for  $\beta$ -tungsten superconductors is A-15 superconductors, a classification according to the crystal structure of the material.

Additional main parameters to characterize superconductors are the upper critical magnetic field,  $B_{c2}$ , and the critical density,  $J_c$ . Again, superconductivity disappears for values of the magnetic field  $B$  or the current density  $J$  above their critical values.

Solid-solution superconductors, for example, NbTi, are limited to be used up to a magnetic field of about 9 T. Some superconductors of the  $\beta$ -tungsten group have the potential to reach magnetic fields well above 20 T. To take advantage of this potential, many different production methods of wires and tapes have been established. As all A-15 compounds are very brittle, the processing to final dimension takes place in a more ductile state and mainly in combination with substrate materials for electrical and mechanical stabilization. At final dimensions, or even after winding to coils, the superconducting layers are formed during a defined heat treatment.

Many procedures have been developed in order to increase  $B_{c2}$  and  $J_c$  by adding elements for alloying the basic metals. Because of poor ductility and differences in thermal expansion coefficients of the materials in such a composite, stress and strain needs to be carefully controlled during handling and in the final application, for example, high-field magnets of different sizes.

This name, used for the class of materials addressed by the title, is based on the  $\beta$ -form of the element tungsten (W), first reported by Hartmann et al. in 1931 (1). This modification is taken as a representative of the A15 structure. As one was not entirely sure whether or not this structure occurs in pure W, but only in connection with oxygen, other prototypes have been sought. The testing of the elements has not been fruitful, and so the lattice of a binary intermetallic compound of chromium (Cr) and silicon (Si), that is, Cr<sub>3</sub>Si, became a commonly used indicator. In the following text, the term A15 will be utilized. In the unit cell of a binary compound A<sub>3</sub>B, which shows the A15 structure, the ions are ordered in the following manner (see Fig. 1). The B ions are arranged in such a way as to build up a body-centered cubic (bcc) sublattice. The A ions form non-intersecting chains in the (100)-, (010)-, and (001)-directions on the faces of the unit cell.

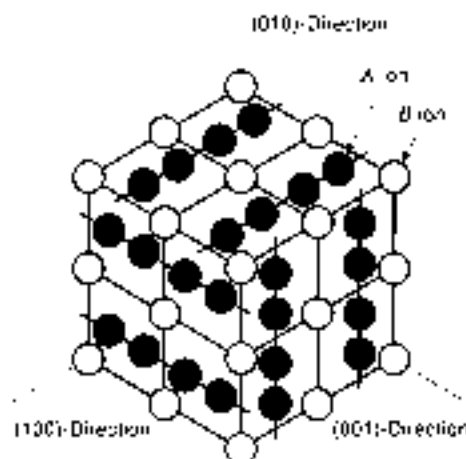


Figure 1. Part of an A15 lattice emphasizing the bcc sublattice of the B ions and the nonintersecting linear chains of the A ions.

In this group of superconductors, however, the A ions commonly are niobium (Nb) or vanadium (V), and the B ions are, for example, recruited from tin (Sn), aluminum (Al), germanium (Ge), or gallium (Ga), without trying to be exhaustive. Thus, typical representatives are materials such as Nb<sub>3</sub>Sn, Nb<sub>3</sub>Al, Nb<sub>3</sub>Ge, or V<sub>3</sub>Ga, to name only a few. Other members, together with their transition temperature  $T_c$ , can be found in Table 1 (2-4).

The explanation of why a crystal structure is applied in order to characterize a class of superconductors, in spite of the fact that it is not entirely possible to assign superconductivity to a certain combination of qualities, becomes evident noting the following: Until the so-called high  $T_c$  superconductors were discovered, the record for the maximum  $T_c$  has always been held by members of the A15 family.

All metallic materials being superconducting in the range from 14 K to 23 K have the crystal structure of the A15s. The high superconducting transition temperature  $T_c$  for V<sub>3</sub>Si and Nb<sub>3</sub>Sn has been discovered by Hardy and Hulm (5) in 1953 and Matthias et al. (6) in 1954. A15 materials are expected not to be superconducting above 25 K (7), due to the increasing instability of their structure related to the electron-phonon interaction.

Nevertheless, there are compounds showing the A15 structure which do not have a remarkably high  $T_c$ . For example, the intermetallic compound containing Nb and osmium (Os), that is, Nb<sub>3</sub>Os, has rather a low one:  $T_c = 1$  K (8), whereas the compound of V with cobalt (Co), that is, V<sub>3</sub>Co, does not show superconductivity at all, down to a temperature of  $T = 0.015$  K (9). That fact made the search for an explanation for the qualities of those materials even more difficult.

In general, it is possible to describe the physical phenomena relevant for the A15s, for example, the high superconducting transition temperature, which is also related to the BCS formula (9), by one-dimensionality, partially localized states near to or at the Fermi level, elastic softness, and strong electron-phonon coupling.

Table 1. Transition Temperatures (in K) of Some A15 Superconductors (2-4)

| A15   | Cr <sub>3</sub> Si | Mo <sub>3</sub> Ir | Nb <sub>3</sub> Al | Nb <sub>3</sub> Au | Nb <sub>3</sub> Ga | Nb <sub>3</sub> Ge | Nb <sub>3</sub> Sn | Ta <sub>3</sub> Sn | Ti <sub>3</sub> Ir | V <sub>3</sub> Al | V <sub>3</sub> Ga | V <sub>3</sub> Ge | V <sub>3</sub> In | V <sub>3</sub> Si |
|-------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| $T_c$ | ~12                | 8.5                | 19.1               | 11.9               | 20.7               | 23.2               | 18.2               | 5.8                | 4.3                | 9.6               | 14.6              | 6                 | 14                | 17.1              |

In the A15s, a martensitic diffusion, characterized by an atomic motion during the transformation at low temperatures, not exceeding the distance of an unit cell's size, takes place. Such martensitic transformations are quite common in solids. The most famous example is an alloy of iron (Fe) and carbon (C), the FeC transformation to the Fe<sub>3</sub>C, the *bcc*-iron or martensite. The martensitic transformation in A15, above  $T_m$  (martensitic phase transition temperature) is from cubic A15 to a more tetragonal structure at a few tenths of a kelvin below this temperature. This transformation has been observed first at V<sub>3</sub>Si and Nb<sub>3</sub>Sn by Barterman and Barrett<sup>10</sup> by x-ray diffractum.

Further investigations have been made by susceptibility and NMR measurement of Knight Shift and nuclear spin lattice relaxation rate (11). Results are that, in the superconducting state, the A15s have the behavior of type-II superconductors. It is remarkable that, the higher the transition temperature, the higher the temperature dependency of the susceptibility becomes.

The development of technically usable A15s was correlated to the research work done for conductors having a high critical temperature  $T_c$ , and being adequate for the production of high-field magnets.

Early progress for commercial use has been achieved with powder metallurgical systems, by filling Nb and Sn powder into a Nb tube, compacting the powder, and drawing the entire piece to the final wire diameter. Other methods targeting the production of tapes, which was possible by passing tapes of the substrate material, for example, V or Nb, through a bath of molten Ga or Sn, respectively. Additional experiments used the same principal approach, by producing pancake-like tapes, made by rolling together already alloyed tapes. Chemical vapor deposition (CVD) and cathodic sputtering are being used, as well.

The advantage of A15 conductors at high  $T_c$  and high  $B_{c2}$ , the upper critical field (To be precise, the upper critical field would be  $H_{c2}$ , which is related to the upper critical flux density  $B_{c2}$  in the following way:  $B_{c2} = \mu_0 H_{c2}$ . But the common combination used in the superconducting society is the one above.) is faced by the disadvantage of the brittle intermetallic A15 compounds. That applies for conductors having reduced mechanical tension in the diffusion layer. In the case of tapes this is given, due to the smallness of the distance of the neutral phase to the diffusion layer, resulting in bending stresses low enough to wind magnets after the A15 forming heat treatments.

Tapes, having naturally a large width-to-thickness ratio, are unsuitable for rapidly excited magnets. The A15 layer causes instability because of flux jumps originating from the currents perpendicular to the superconducting layer. Fulfilling the demand of intrinsic stability is one of the important criteria for modern superconducting magnets, and can only be achieved by dividing the superconductive parts into very small portions.

First ideas for the production of fine filaments in a substrate have been laid down in the British patent GB No. 1203292 of 1966 (12). Nevertheless, the advantages of the properties of tapes (Nb<sub>3</sub>Sn, V<sub>3</sub>Ga) for the design and construction of pancake coil magnets has been used on several high-field magnets. A magnet of 17.5 T (13.5 T with Nb<sub>3</sub>Sn plus 4.0 T with V<sub>3</sub>Ga), being constructed by InterMagnetics General Corporation (IGC), has been installed at the National Research Institute for Metals (NRI) in Japan in 1975 (13).

In the case of fine filaments, materials in a ductile state are drawn down to final dimensions. The A15 compound is then

formed by a heat treatment, using the solid-state diffusion. This method of solid-state diffusion has first been applied by Tachikawa for a V<sub>3</sub>Ga wire (14). The A15 compound is made via a heat treatment at approximately 700 °C, which enables the Ga, solved in the Cu, to diffuse into the V filaments, forming V<sub>3</sub>Ga.

As the problem of brittleness still exists, it has now become common to wind the magnets' coils first, having the treatment afterwards. This so-called wind-and-react method is now used for almost all magnets assembled of A15 superconductors. The structures of V<sub>3</sub>Ga and Nb<sub>3</sub>Sn are very similar, but material costs and production procedures are in favor of Nb<sub>3</sub>Sn. It is possible to perform the solid-state diffusion by several methods. Examples are diffusion of Sn from a bronze with a high tin content into Nb filaments, or by an external diffusion of Sn in a Cu-Nb system, with intermediate steps of forming a bronze by the Sn diffusion procedure. The bronze process is applicable, for example, for Nb-Sn, V-Ga, V-Si, V-Ge, Nb-Au, and Nb-Al.

In order to get high critical currents, it is obvious that the reaction of Nb and Sn or V and Ga to Nb<sub>3</sub>Sn or V<sub>3</sub>Ga must be as complete as possible. Different methods have been developed to reach this target, for example, Nb-Sn internal Sn conductors, powder-in-tube, jelly roll, or in situ technologies. Such methods are necessary, as otherwise the amount of Nb<sub>3</sub>Sn in the cross section is limited to the Sn content of the bronze. Compared with conductors produced via the bronze route, these materials are superior, in terms of the critical current density  $J_c$ . Forced by design and fabrication technology, those conductors tend to have filament coupling and high hysteresis losses, caused by relatively large effective filament diameters. These values of the effective filament diameters are significantly better for the bronze route conductors, also showing higher figures for  $T_c$  and  $B_{c2}$ .

Nb<sub>3</sub>Al is one of the about 47 known A15 superconductors, having the third highest  $T_c$  of 19.1 K (see also Table 1), ranking behind only impractical materials like Nb<sub>3</sub>Ge (23 K) and Nb<sub>3</sub>Ga (20.4 K). While having these promising figures, the production of Nb<sub>3</sub>Al turned out to be a difficult task. Different production methods of powder metallurgy have been utilized, as well as the tube and the jelly roll processes.

A15 superconductors are strain sensitive and, while Nb<sub>3</sub>Sn may drop about 50% in  $J_c$  at a strain of 0.05%, the  $J_c$  of Nb<sub>3</sub>Al is reduced by 20% only at a strain up to 0.8% (15). That gave reason for renewed interest in that material, especially for large-scale applications like the International Thermonuclear Experimental Reactor (ITER). For the very big coils of such a machine, the conductors are cabled and welded into a stainless-steel conduit, resulting in additional axial strain.

A significant amount of research work has been done toward powder metallurgical solutions, but no commercially usable process has been developed for Nb<sub>3</sub>Al. The jelly roll and Nb tube processes seem to be more promising to yield high  $J_c$  values at reasonable length. Wires of A15 type superconductors used for the wind-and-react technique need to be insulated with an insulation material like glass, quartz, or ceramic, withstanding the heat-treatment temperature. In the case of Nb<sub>3</sub>Sn, this temperature is at approximately 700 °C, leading to a braided insulation type of E- or S-glass. The higher temperatures, which are necessary to form the Nb<sub>3</sub>Al phase, are another obstacle for a broader use of this material.

While  $Nb_3Sn$  became the main choice of the A15 materials being used in a wide section of high-field magnets, a development leading to higher critical currents and better mechanical values as necessary.

The main factor that is limiting high-magnetic field is the Lorentz force, which increases proportionally to the square of the generated magnetic field. The stresses induced into the conductor and the strains require reinforcement, reducing the overall current density and thus leading to magnets being bigger and, therefore, less efficient. The bronze substrate of A15 superconductors is highly resistive and the stabilization is poor. Cu of very good conductivity has to be added (compare Fig. 2(a) with Fig. 2(b)). Different methods, different with respect to the method of processing the conductors, like internal Sn, jolly roll, and bronze process, are used. In order to prevent the diffusion of B ions, for example, Sn into the Cu, diffusion barriers are necessary. In case of the bronze process conductors, diffusion barriers of Ta or Nb are used.

To influence the superconducting parameters, the critical current density  $J_c$  and the upper critical field  $H_{c2}$ , by getting better heat-treatment conditions and higher pinning forces, third elements, creating ternary and quaternary alloys, for example, (Nb, Ta) $_3$ Sn, (Nb, Ta, Ti) $_3$ Sn, or Ge, Hf, or Al are added. As not only  $J_c$  but also the overall current, usually denoted by  $I_{c,overall}$ , is important, the size of conductors has to be increased

This is possible by producing larger monolithic superconductors, mainly of rectangular shape. Another alternative is the cabling of a number of round wires into flat or round cables. Special attention to possible reasons for the reduction of  $J_c$  by filament size, the principle conductor design, or the cabling process, has to be given.

To fulfil basic requirements of technically reliable superconductors like high transition temperature  $T_c$ , high critical field  $H_{c2}$ , and high critical current density  $J_c$ , many different approaches have been used. This was mainly done by developing production and material treatment methods adequate to overcome the sometimes difficult material parameters of the A15 components and in order to optimize the parameters of the final conductor configurations. As all A15 materials are brittle by nature, which cannot be overcome by any means, it is necessary to use methods of deformation to get tapes or wires being cold-worked prior to forming the A15 layers. Once the layers are formed, the handling of the conductors needs utmost care, in order not to reduce or completely to destroy the superconducting properties. Deformation of conductors with already existing A15 formation is only possible with complicated, not very practical methods, for example, extremely high hydrostatic pressure.

## METHODS OF PRODUCTION

### Surface Diffusion Process

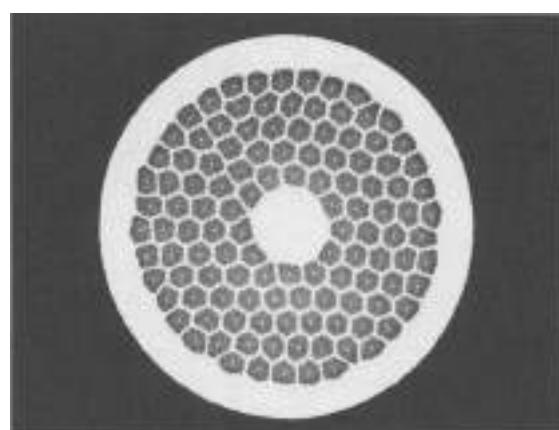
In this process, a tape of V or Nb is dipped into a Cu or Sn bath. The heat treatment then forms  $V_3Cu$  or  $Nb_3Sn$ . It is also possible to hold the bath at the reaction temperature, forming the A15 phase during immersion. To use the better flexibility of tapes and getting the relatively thin layers of the brittle A15 only, other methods, such as sputtering the V or Nb to the core material by cathodic deposition or chemical vapor deposition, or by condensing under high vacuum conditions have been applied. For special MRI magnets, tapes of (Nb, Zr) $_3$ Sn are introduced, operating at 9 K (16).

### Composite Diffusion Process

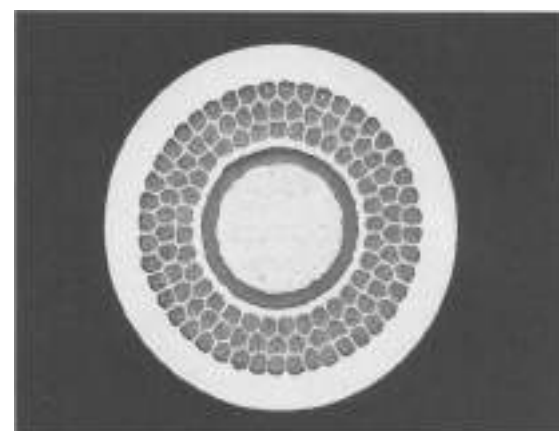
Members of the composite are cold-worked together in final shape. The final heat treatment forms, by solid-state diffusion, the superconducting phase or a microstructure, which also has normal conducting phases. Tapes or even wires made of the components of the superconducting phase, probably in combination with other elements for electrical or mechanical stabilization, or just to improve the workability, are cold-worked together, tapes of Nb and Al or Cu, Sn, and Nb being stacked and rolled. It is also possible to use tapes of the high-melting component with layers of the lower-melting component received by CVD for stacking and rolling. Clad chip extrusion (17), by producing the components as described, chipping and filling them into a Nb-lined Cu can for extrusion and drawing, is another development of this production.

### Powder Metallurgy Process

Matrix material like Cu is mixed with other components, for example, Nb and Sn, or Nb and Al, after the compaction to final shape of wires or tapes. The formation of the superconducting phase takes place on the outside of the powder layers.



(a)



(b)

**Figure 2.** Different cross sections of  $Nb_3Sn$  heat-conductors; (a) with TaCu core for stabilization, and (b) unstabilized (courtesy of Viewcom-Whelan).

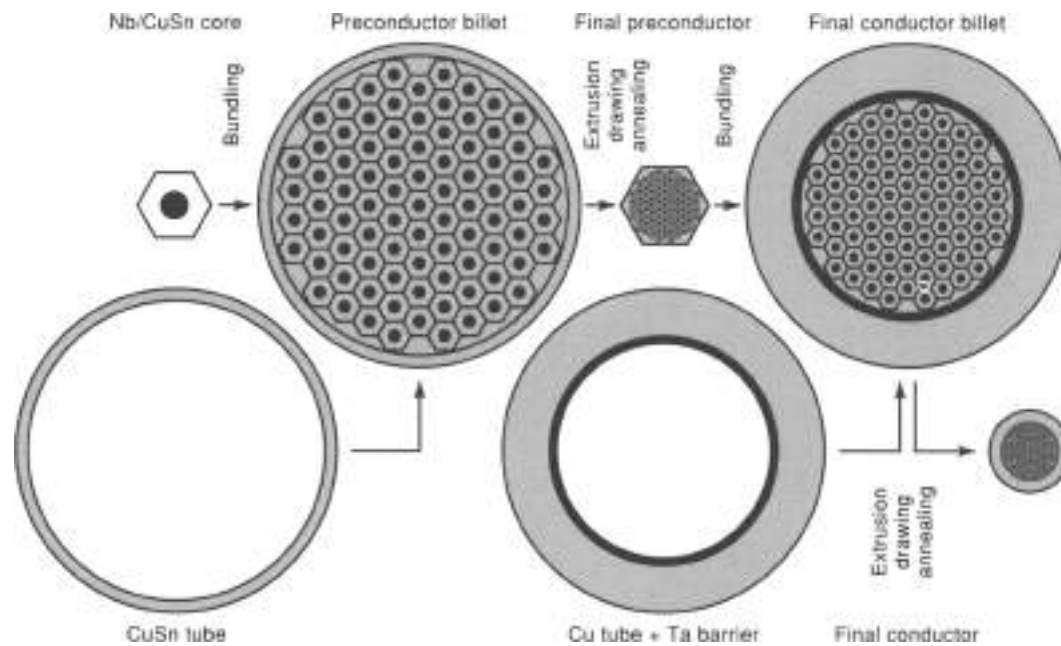


Figure 3. Fabrication of bronze matrix conductors (courtesy of VacuumSchmelze).

Producing a superconducting wire by PM methods has been approved by Kunzler et al. (15) and was the first successful process for A15 wires. The Nb tube has been filled with (Nb, Sn) powder with a ratio of 3 to 1 or a  $(\text{Nb}_3\text{Sn}, \text{Sn})$  powder. Heat-treatment temperatures of 970 °C to 1400 °C are used. More recently, powder of the intermetallic compound  $\text{Nb}_3\text{Sn}_2$ , mixed with Sn powder, is filled into a thin-walled tube of Cu or Cu-based alloy, and reduced in size for compaction. The powder is then placed into a Nb tube within a Cu tube. After that, it is possible to use this combination and stack it again into a Cu can with other matrix parts. Such a billet can be cold-worked to final size without further heat treatment. During the heat treatment, at typically 650 °C to 700 °C during 15 h to 45 h, a  $\text{Nb}_3\text{Sn}$  layer of about 2.5  $\mu\text{m}$  thickness is formed on the inner side of the Nb tube, resulting in filament diameters of 10  $\mu\text{m}$  to 20  $\mu\text{m}$ , depending on the number of cores. The reaction process is, in fact, a two-step process of  $\text{NbSn}_2 + \text{Sn} \rightarrow \text{Nb}_3\text{Sn}$ , from the core, and then forming with the Nb tube and, by depleting of Sn the  $\text{Nb}_3\text{Sn}$ . Similar methods can be used for  $\text{Nb}_3\text{Al}$ , especially by using sintered rods of Nb powder, metal-impregnated to get it self-supporting, due to immersion into Al-Ce or Al-Si baths at temperatures of 600 °C or 580 °C, respectively (19). Reaction treatment at 1700 °C is followed by final annealing, at 750 °C.

#### Bronze Process

This technology is very typical for  $\text{Nb}_3\text{Sn}$  and  $\text{V}_3\text{Ga}$  superconductors. In a substrate of Cu-Sn or Cu-Ga, a number of filaments of Nb or V are inserted. The solubility of Sn and Ga is limited to approximately 8 at.% or 20 at.%, respectively. This solubility restrains the amount of the A15 compound that can be formed by solid-state diffusion. The high-strength component, the bronze, is getting the low-strength component, Nb-Ga, to a laminated flow, but multiple intermediate heat treatments due to work hardening are necessary; see, for example, Table 2 (20). Relatively thin filaments, with a maximum diameter of about 5  $\mu\text{m}$ , are necessary to receive good overall current density. The

Sn reservoir and the diffusion path is the bronze between the filaments, but the decreasing content of Sn during the reaction treatment reduces the speed of diffusion. Heat treatments at about 700 °C, in many different variations, form the intermetallic A15 compound at the interlayer of the bronze to Nb or V. This also creates Kirkendall voids (21), which disturb the diffusion paths, see Fig. 4. To get high filament numbers for conductors of bigger size and due to the limitations in filament diameter, second and third stacking of billets may be necessary. For electrical stabilization, Cu can be embedded in the matrix or in the center of the conductor. An outer shell of Cu is also possible, especially in the case where a larger amount of Cu is needed; see Fig. 3. Diffusion barriers of Ta, Nb, or V are necessary, to protect the Cu during the diffusion process by preventing B ions entering from the bronze into the Cu, which would significantly reduce the conductivity.

#### Internal Sn Process

A Cu jacket, having a Sn rod in the center and Nb cores distributed around it in the remaining cross section, is the starting element of this process. These subelements may be surrounded by Nb or Ta barriers, to protect the stabilizing Cu from diffusion of the Sn. This Cu is the outer tube into which the subelements are inserted. The Cu-Nb-Sn composite is cold-worked to final dimension, without intermediate annealing. The bronze matrix is formed when the conductor already is in its final shape. Therefore, the heat treatment needs two cycles. The first typically starts at 200 °C for 100 h, continues at 375 °C for 24 h, and finishes at 580 °C for 50 h. It is needed for the homogenization of the Cu-Sn bronze, containing different phases according to the phase diagram in Fig. 4 (23). A consecutive second heat treatment of one or more steps is necessary to form the  $\text{Nb}_3\text{Sn}$  layer. The typical temperature is around 725 °C. For a better distribution of the Sn, methods of pinning it rather close to the filaments instead, as a core in the center of the subelements, is used, resulting in reduced heat-treatment time and

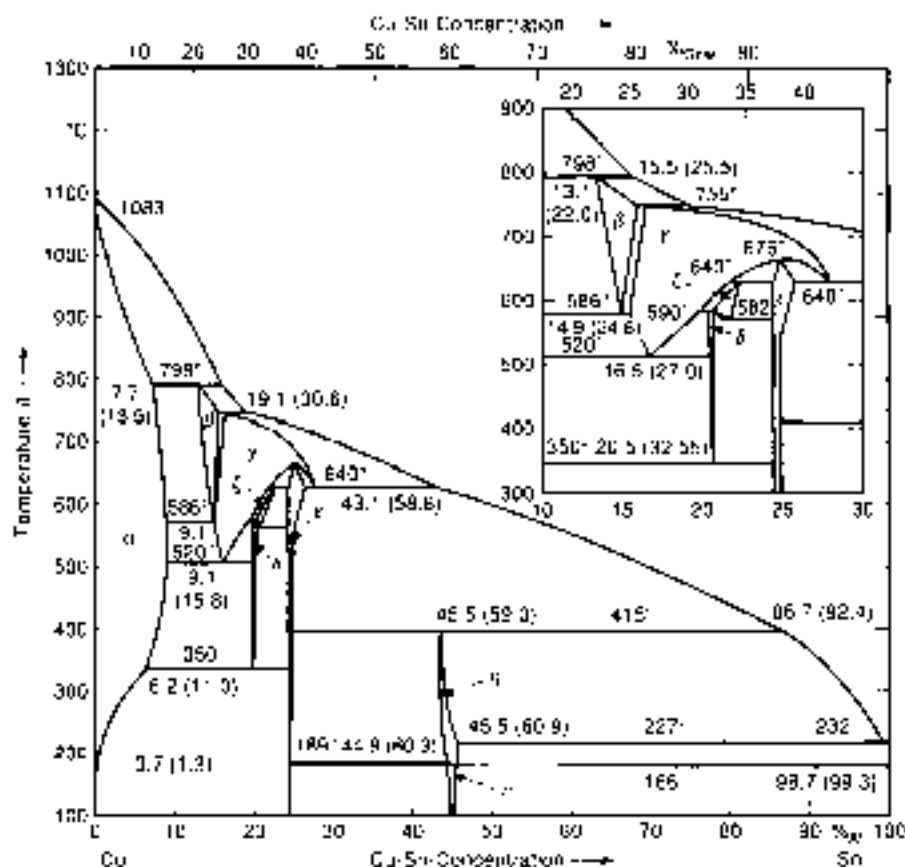


Figure 4. Detailed phase diagram of Cu-Sn (23).

a homogeneous A15 layer distribution over the cross section of the conductor (24).

#### In Situ Process

Casting ingots of two phase Cu-Nb or Cu-V with dendrites dispersed into the matrix of Cu, cold-working, and then coating with Sn or Ga, are the basic steps to start this process (Fig. 5(a)). To reduce the likelihood of wire breakage, a bronze matrix may be used. After dispersing the Sn or Ga throughout the matrix,  $Nb_3Sn$  or  $V_3Ga$  are formed during the heat treatment at the interfaces between the matrix and the filaments. Highly homogeneous Cu-V ingots are difficult to produce by conventional casting methods, due to a large miscibility gap in the liquid region. A continuous arc casting plus Ga coating give the opportunity to optimize the material toward better mechanical values and better current-carrying capacity. Large ingots up to 150 mm in diameter (Fig. 5(b)) have been produced by smelting the Cu-Nb into a CuO mold for solidification. Wires produced thereof have shown, after heat treatment at 800°C for 25 h, good mechanical values in untwisted condition. No degradation is found up to 1.2% strain, but in the case of twisted wires, the degradation is significant (25a). Multistage cables made from fine wires with a diameter of 0.2 mm with small in situ filaments of 0.5  $\mu$ m have been developed for ac applications. Such cables are showing even more improved values in strain resistance (25b).

#### External Diffusion Process

This process is also typically used for  $V_3Ga$  and  $Nb_3Sn$  composites. In a Cu billet with drilled holes, rods of V or Nb are

embedded and cold work is applied to get the final dimensions. Prior to heat treatment, the surface is coated with Ga or Sn having first the diffusion into the matrix, and then furns at the interface between matrix and the cores the  $V_3Ga$  or  $Nb_3Sn$  layers, respectively. The amount of A15 phase produced by this method is limited by the quantity of the B material which can be stored in the coating of the Cu-A wire. This fact dictates the maximum diameter of the wire. The combination of this method with the bronze process, in order to increase the tin content, is also possible.

#### Jelly Roll Process

A foil of Nb with slit meshes and a foil of Cu-Sn bronze are spirally rolled into a cylinder, cold-worked, and heat-treated. The  $Nb_3Sn$  is received at the interface of the Cu-Sn and the Nb. Despite the successful use of the jelly roll process for the  $Nb_3Sn$ , this method was originally developed for  $Nb_3Al$  (26). Foils of Nb and Al have been wrapped around one Cu cylinder and inserted into another Cu cylinder. Wires with a diameter of 0.2 mm receive a heat treatment of several hours at about 850°C, resulting in a mixture of  $Nb_3Al$  and nonreacted Nb. Placing some, for example, nine, of the unreacted jelly rolls into a Cu rod with drilled holes, is a logical way to increase the current by getting a larger cross section. The modified jelly roll technology is performed (27) by wrapping layers of Cu sheets and expanded metal sheets of Nb around a Sn rod. Conductor designs with 54 of such jelly rolls assembled with a barrier of Nb or V are placed in a Cu can for cold-working to final dimensions, providing a high  $J_c$  and large Cu cross-section, likewise. In principle, the modified jelly roll process proceeds like the internal Sn process.

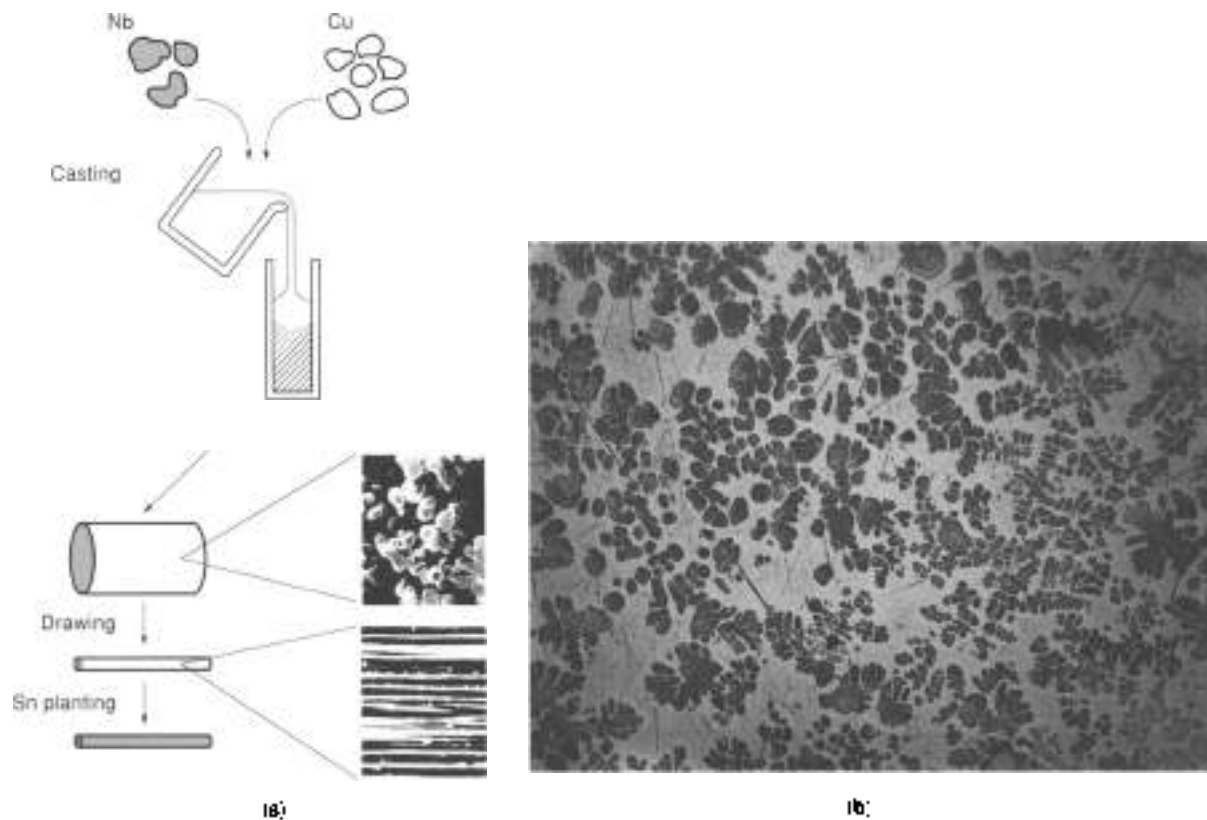


Figure 5. (a) Schematic diagram of in situ process of Cu-Nb composite (courtesy of Fujikura). (b) Transverse section showing Nb dendrite solidified in Cu matrix (courtesy of Fujikura).

#### Tube Process

To form the basic composite, a Cu tube with a Sn or Cu-Sn alloy core is inserted into a Nb tube. This combination is introduced into another Cu tube and then cold-drawn. A number of those single-core elements are stacked for assembling into a Cu can and cold-worked to final dimension. For the production of basic composites of larger conductors, cold working is performed down only to intermediate dimensions, followed by a second assembly. During heat treatment, the Sn diffusion into the matrix takes place first. Then the Nb<sub>3</sub>Sn layer is formed at the interface of the Cu to the inner end of the Nb tubes. This technology has also been used for V<sub>3</sub>Sn wires and tapes, rolled from wires (28).

#### Specialties of the Production of Nb<sub>3</sub>Al

The A15 phase in Nb<sub>3</sub>Al can be obtained from the melting bath using high temperatures, and at reduced temperatures for the reaction between Nb and Nb<sub>3</sub>Al. The  $\sigma$ -phase of Nb<sub>3</sub>Al needs temperatures above 1600 °C. By reacting with Nb at about 1800 °C, A15 is formed. The desired stoichiometry of the A15 phase seems to be stable at high temperatures of > 1770 °C (only 1940 °C in the corner of the phase diagram of Fig. 6 (29)). In general, the high reaction temperature for the conventionally processed materials leads to grain growth. Because many production techniques failed to get stoichiometric Nb<sub>3</sub>Al, numbers of different approaches have been tried, like laser alloying or rapid quenching by melt spinning. Powder metallurgical processes are unfavorable, because of the high oxygen (O) content in the powder, which does not allow the high deformation rate

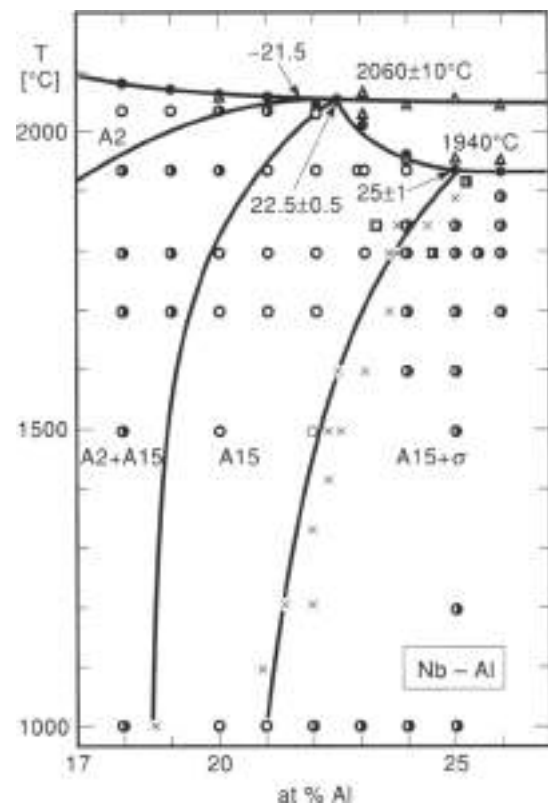


Figure 6. Detailed phase diagram Nb-Al (29)



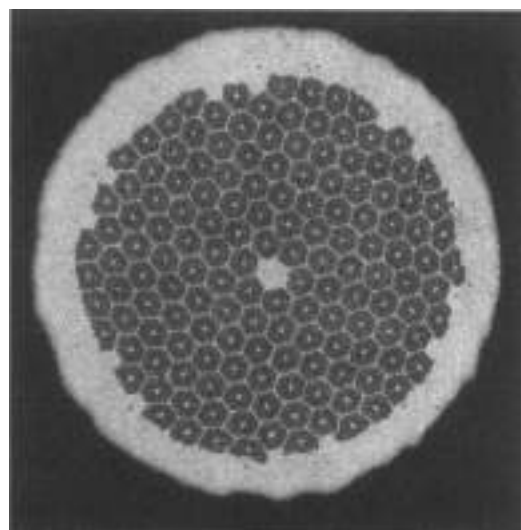


Figure 7. Cross section of Nb-Al composite wire (40).

required. To get reliable conductor lengths, especially for magnets with magnetic fields above 21 T, a distinct rapid-quench process has been established (30). By the jelly roll technology, the wire consisting of several elements made from Nb and Al sheets wrapped around an Nb core is produced, using multi-stacking and extrusion procedures, resulting in a cross-section pattern, as can be seen in Fig. 7. The wire produced this way is subject to ohmic heating at 1900 C to 2000 C, and rapidly but continuously quenched in a bath of molten Ga (melting point: 30 C); see Fig. 8. Such a treatment leads to a supersaturated Nb-Al<sub>2</sub>Sn phase. This metastable phase is transformed by a heat treatment of 700 C to 900 C to microcrystalline A15 of nearly stoichiometric composition. The resistivity ratio RRR in the Nb matrix has a value of about 17, giving reduced concern about bridging of filaments in this kind of conductor. Due to the processing, the wire consists of Nb-Al only, with excess of Nb. For the use at a higher current  $I$ , Cu has to be clad, in order to stabilize the wires.

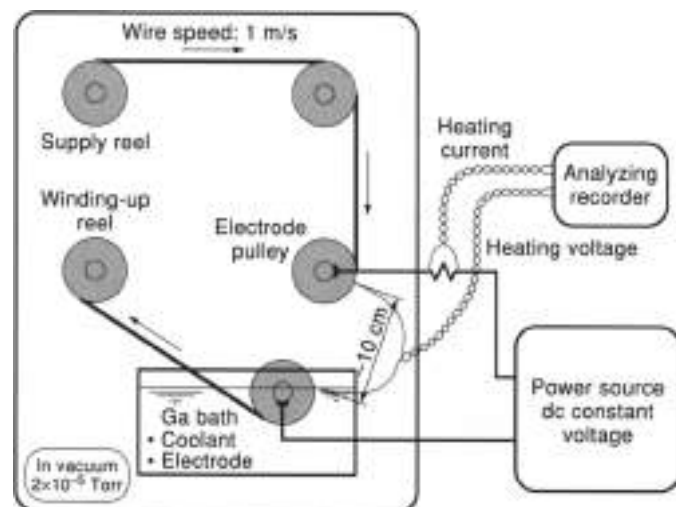


Figure 8. Schematic diagram of rapid-heating, rapid-quenching apparatus (30).

## BRONZE CONDUCTORS

To fabricate high-field magnets, flexible tapes, having the advantage of a small distance from the brittle A15 compound to the normal phase in bending direction, have been used successfully. The large area-to-thickness ratio of the A15 layer leads to instability (flux jumps), especially if magnets have a rapid ramping rate. The solid-state diffusion process, as used for bronze conductors of Nb<sub>3</sub>Sn and V<sub>3</sub>Ga (31), has solved this problem, by dividing the core material into plenty of fine filaments. The formation of A15 layers is principally limited by the amount of Sn and Ga in the bronze. The solubility of Sn in Cu is 8.5 at.% and for Ga in Cu is 20 at.%. Bronze with about 7.5 at.% Sn or about 18 at.% Ga has been used. The diffusion process forms the A15 layer until the equilibrium for a given temperature is established. At a temperature of 620 C to 700 C, the diffusion for Ga ends at a remaining concentration of 14 at.% to 15 at.% Ga in the matrix. Sn diffusion from the bronze proceeds at approximately 700 C to 850 C, leaving a Sn concentration in the bronze of approximately 3 at.% to 4 at.%, see Table 3.

The heat-treatment time and temperature has to be controlled in such a way as to receive an optimum layer thickness, but without increasing too much the grain size. Especially for a long heat treatment of >200 h, the matrix volume has to be increased, in order to provide enough Sn or Ga. Small distances between the filaments seem to be desirable, due to reduced bending strain, but the space between the filaments acts as a diffusion path for the B ions from the conductor periphery, too. Those diffusion paths are reduced in their effective width by the Kirkendall voids (see Fig. 9) (32) caused by the diffusion mechanism during heat treatment. At a given temperature and a constant concentration gap, the quantity of B ions diffusing through a cross section in a given time is proportional to the area of this cross section (Fick's first law). From this follows that the cross section of the cores of Nb or V should be divided in as many portions as feasible, to increase the interlayer between the bronze and the core material. This leads to an increase of the total amount of A15 material, even with reduced heat treatment time. Optimization studies of diffusion treatment versus layer thickness have shown that filament diameters should be in the range of 3  $\mu$ m to 5  $\mu$ m. For conductors with a diameter of 1.5 mm, and taking into account the cross section needed for stabilization and diffusion barrier, approximately 15,000 filaments are necessary. Workability of the component is an essential request to arrive with technically and commercially usable conductors. While the basic components, Nb or V, electron beam- or arc-melted, are high-purity materials of excellent

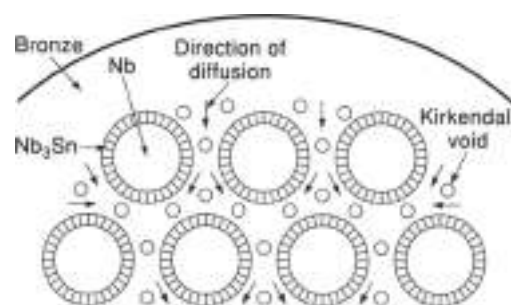


Figure 9. Diffusion in the system tin-bronze with Nb filaments showing the narrowing of the diffusion paths by the Kirkendall voids after (32).

**Table 2. Increase of Hardness of Niobium and Tin Bronze During Deformation (20)**

| Sample                  | Degree of Deformation (%) | Heat Treatment Time (h) | Temperature (°C) | Hardness (N/mm <sup>2</sup> ) |
|-------------------------|---------------------------|-------------------------|------------------|-------------------------------|
| 7.7 at. % Sn            | 0                         | —                       | —                | 1260                          |
|                         | 21                        | —                       | —                | 2230                          |
|                         | 27                        | —                       | —                | 2370                          |
|                         | 21                        | 1                       | 500              | 1940                          |
|                         | 21                        | 5                       | 500              | 1960                          |
|                         | 21                        | 8                       | 500              | 1140                          |
| 9.5 at. % Sn            | 0                         | —                       | —                | 1350                          |
|                         | 21                        | —                       | —                | 2560                          |
|                         | 21                        | 1                       | 400              | 1640                          |
|                         | 21                        | 5                       | 400              | 1370                          |
|                         | 21                        | 9                       | 400              | 1330                          |
| Nb, original<br>Nb Core | 0                         | —                       | —                | 1020                          |
|                         | 25                        | —                       | —                | 1330                          |
|                         | 57                        | —                       | —                | 1480                          |
|                         | 68                        | —                       | —                | 1460                          |
|                         | 75                        | 1                       | 550              | 1530                          |
|                         | 97                        | 1                       | 550              | 1510                          |
|                         | 99.7                      | 1                       | 550              | 2030                          |

ductility, they are sensitive to embrittlement by interstitials of oxygen (O), nitrogen (N), or C. This is especially true for the V, but the more problematic part is the bronze; see Table 2. Normally Cu-Sn bronze contains about 10 wt.% Sn, and as deoxidizer phosphorus (P) is used. In the Nb-Sn system, P prevents the diffusion procedure. The amount of Sn should be as close as possible to the solubility limit of 8.5 at.%. For many years, the technically attainable Sn content was limited to about 8 at.%. Newer processes made homogeneous bronze at 8.5 at.% Sn available (33). The positive influence of the Sn content on  $J_c$  is shown in Fig. 10. During cold work, the hardness of the bronze is increasing rapidly, as shown in Table 2, and a significant number of intermediate heat treatments have to be applied. From the workability point of view, this is certainly a disadvantage of the bronze process.

#### STABILIZATION AND BARRIERS

To have maximum  $J_c$ , the core material has to be converted into A15, theoretically to 100%. Regarding mechanical and electrical stability, a small core of unreacted material, even in fine filaments, has a positive influence on the overall performance of the conductor. For reasons of electrical and thermal stabilization, for example, during the occurrence of a quench, which may result from wire movement in connection with

energy dissipation, it is necessary to have a highly conductive material in the cross section. The bronze itself has a rather low conductivity (specific resistance  $\approx 70$  n $\Omega$ m at a temperature of 4 K), which is important to reduce the alternating current (ac) losses. Experiments have shown that, in case of bronze conductors, a reversed transport reaction can lead to a very small (0.3 at.%) Sn content, compared with the values of Table 3. In this low-Sn "bronze," the value for the specific resistivity is reduced by a factor of 10. This method is not very practical and has, in view of ac losses, disadvantages. It is therefore necessary to increase the electrical stability by designing conductors with Cu included in the cross section. This can be done by a few percent distributed throughout the matrix, up to 20% in the wire center and up to a maximum part of the cross section as an outer shell. The composite needs a barrier to protect the Cu from the diffusion of H ions into the stabilizing part, which would reduce the conductivity of the Cu. Barrier materials fulfilling this task are V, Nb, Ta, or alloys and combinations thereof (34). The use of a Nb barrier seems to be the natural choice, as it fits the material parameters of the complete conductor. The Nb<sub>3</sub>Sn layer formed at the interface of bronze and barrier acts just like a large filament, bringing additional high ac losses. The effect of the barrier materials on the hysteresis losses is shown in Fig. 11(a) and Fig. 11(b). The use of Ta barriers avoids magnetic disturbances, even in cases where high-temperature heat treatments are being used. Furthermore, Ta barriers are

**Table 3. Maximum Area Ratio, Remaining Tin Concentration, and Transition Temperature of Solid-State-Diffused Nb<sub>3</sub>Sn Samples (20)**

| Tin Concentration in Alloy (at. %) | Diffusion Temperature (°C) | Maximum Area Ratio $A_{Nb_3Sn}/A_{total}$ | Remaining Tin Concentration (at. % (calculated)) | Transition Temperature $T_c$ (K) |
|------------------------------------|----------------------------|---|--|----------------------------------|
| 7.7                                | 700                        | 0.15                                      | 5.1  | 17.5                             |
|                                    | 750                        | 0.18                                      | 4.6  | —                                |
|                                    | 750                        | 0.26                                      | 3.3  | 19.1                             |
| 9.5                                | 750                        | 0.30                                      | 3.1  | 18.04                            |

Heat-treatment time: 24 h.

Heat-treatment time: 56 h.

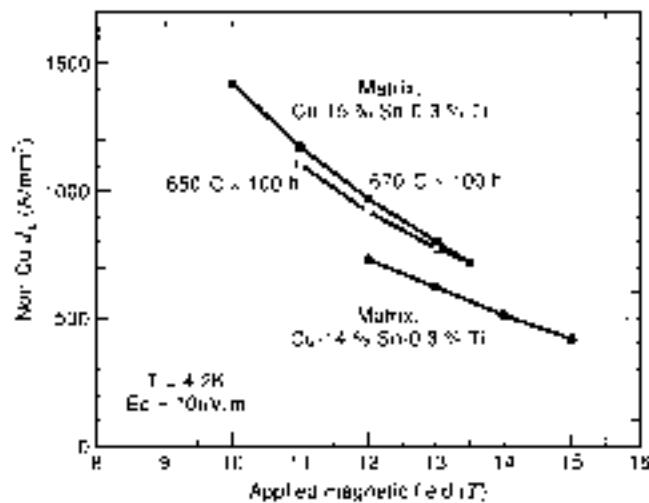


Figure 10. Diagram of non-Cu  $J_c$  versus applied magnetic field  $H$  showing the enhancement of  $J_c$  by the Sn concentration (3%).

effective as reinforcement due to their high Young's modulus and remarkably high strength at 4 K. Access of O must be prevented during any heat treatment in the course of production and the diffusion procedure, as Ta interacts intensively with O. Conductors with peripheral stabilization should have  $\geq 25\%$

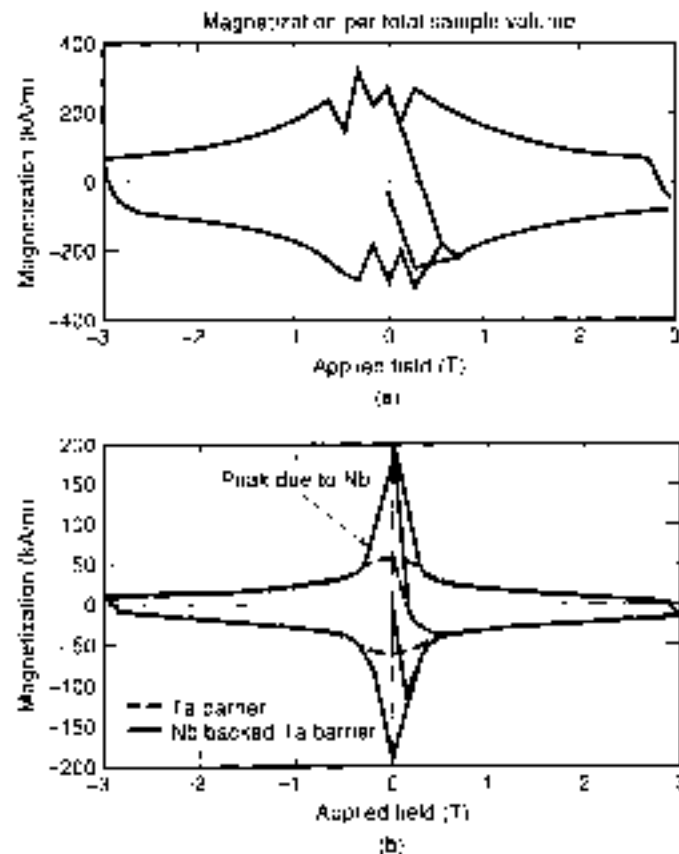
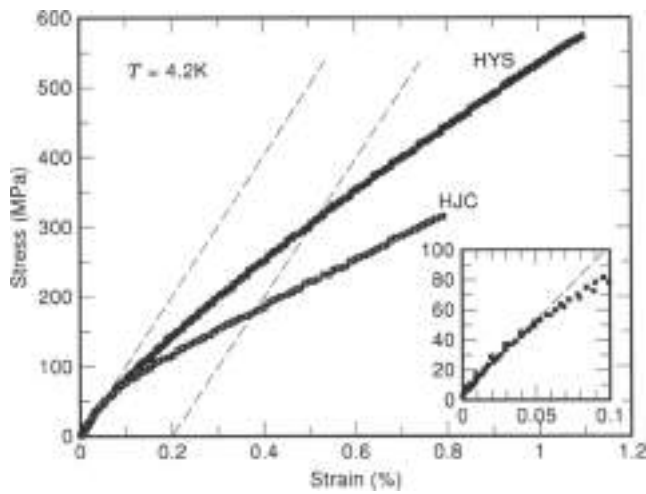


Figure 11. Magnetization curve for material: (a) containing a 100% Nb barrier: The relatively large flux jumps in the low field region are very apparent; and (b) containing an interrupted Nb barrier: No flux jumps in the low field region are observed (24).

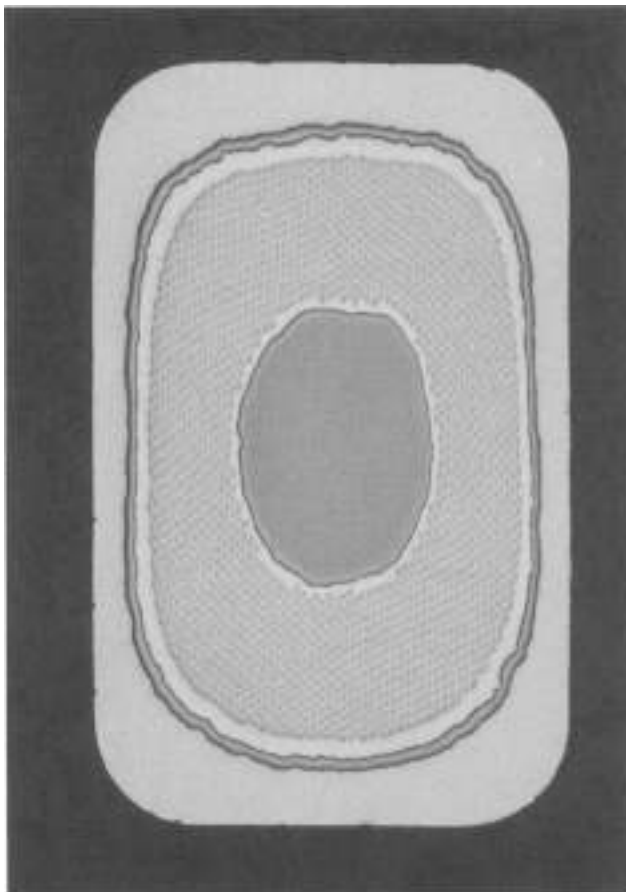
area of Cu, in order not to get a Cu layer  $< 10 \mu\text{m}$ , as O might penetrate into the Ta (32). Barriers of Ta penetrated by O are likely to burst during the final diffusion treatment. Further solutions might be the use of a Ta core, increasing the yield strength significantly by the larger cross section of Ta. Enlarging the Ta portion of the cross section area to  $\geq 10\%$  results in yield strength  $H_{90}$  of  $\sim 250 \text{ MPa}$ . See Fig. 12(a) and Fig. 12(b) (35) for the characteristics and the cross section of a high-yield-strength conductor.

Al may act as a stabilizing material due to its attractive properties: low weight, high specific residual resistivity ratio RRR ( $\sim 1000$ ), high thermal conductivity, and low magnetoresistance. For practical use, restrictions arise from the poor mechanical values, not compatible with the other components of bronze conductors. The melting point of Al is lower than the reaction temperature. Thus, Al is only useable for reacted conductors, giving the need for the react-and-wind technology, which does not implement too high a strain in the conductor during winding, but which is feasible for large magnets only. Metallic and cable conductors can be coextruded together with high-purity Al and reinforcing elements of special steel or cobalt (Co)-based alloys (36). In case a high over-all current  $I_c$  is needed, electrical stabilization is possible by cabling of unstabilized or stabilized wires together with Cu wires. It is essential that those Cu wires are also protected by diffusion barriers, which must have an outer Cu or Cu-Sn layer to avoid O penetration into the Ta. An example of a cable with additional stabilizing Cu-Ta-Cu wires is shown in Fig. 13. Presented cables, soft-soldered to Cu clad tapes of Al, have been used for special laboratory magnets. To build very large coils, as necessary for Tokamak fusion technology like ITER, conductors with critical currents of about 100 kA at 12 T and excellent mechanical values are needed. The present way to fulfill such demands is to produce a round cable of many superconducting wires (ITER:  $\sim 1000$  wires) in a multicabling process, embedded into a stainless-steel jacket. By this method, good cooling conditions are received as well.

A direct relationship exists between  $J_c$  and ac losses: With increasing  $J_c$ , the ac losses are increasing as well. The most important reason for such losses in internal-Sn and jelly roll conductors is filament bridging during the reaction treatment and barrier material like Nb. The increase in volume caused by the diffusion of Sn is about 30% to 40% in the Nb<sub>3</sub>Sn layer, while the outer wire dimension is practically unaffected. The filament bridging results in huge effective filament sizes  $d$ . In powder-in-tube conductors, the filament size depends on the grain size of the Nb<sub>3</sub>Sn powder. As the forming of the Nb<sub>3</sub>Sn layer in the tube process takes place at the interface between the Cu and Nb tubes, large filaments are received. Barriers of Nb behave like large filaments, too. Bronze-processed conductors have the lowest ac losses and smallest  $d$ , as bridging in the bronze matrix is negligible. Bridging is also not a problem for Nb<sub>3</sub>Al conductors. The filament diameter of conductors produced by the jelly roll technique is in the range of 50  $\mu\text{m}$ . The specification of the ITER conductors is, in regard to this fact, divided into two parts—one of high  $J_c$  and high ac losses (HP I), and the other with a low  $J_c$  and low ac losses (HP II); see Table 4. A reduction of ac losses in internal-Sn conductors is possible, by reducing the Sn content, but matching the properties of a bronze process conductor, like the one shown in Fig. 14(a) will be hardly possible, see Fig. 14(b).

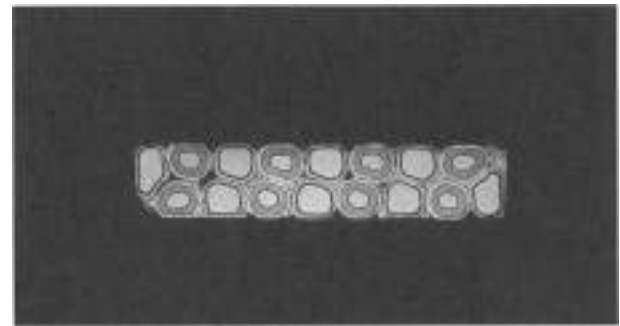


12a



12b

**Figure 12.** Stress-strain characteristics at a temperature of 4.2 K of a high yield strength conductor compared with a high- $J_c$  conductor (35) and cross section of the high yield strength conductor with a Ta core (courtesy of Kobe Steel).



**Figure 13.** Superconducting cable consisting of eight bronze conductors with a TaCu core and eight additional stabilizing wires of CuTaCu (courtesy of Vacuumschmelze).

### Degradation by Cabling

The selection of the jacket material has to take into account that almost no additional strain should be induced into the conductor by the properties of the conductor. The critical current of A15 superconductors is depending on the stress-strain conditions. It is also necessary to consider the behavior of the conductor material during the heat-treatment cycles (react-and-wind) applied for reaction. While cabling seems to be a logical approach to increase the current-carrying capacity, degradation during this process may happen, especially for Rutherford type flat cables. The sensitivity to the deformation which takes place at the edges of flat cables is in dependence to the conductor design and making. Bronze conductors with and without stabilization exhibit a degradation of only 5% and, even as specially enhanced designs, not more than about 10%. Other configurations like internal-Sn have shown a significantly higher degradation of the critical current  $J_c$  (37). Generally, cables of A15 are workable with bronze conductors or with powder-in-tube conductors. In the case where cables made from Nb<sub>3</sub>Al wires of the continuous-quench method are applicable, internal-Sn conductors have potential for improvement (24). Jelly roll conductors of Nb<sub>3</sub>Sn exhibit high sensitivity to strain-induced damage and are thus not suitable (38). For cables consisting of many single superconductor wires, sintering of the wires during heat treatment has to be avoided, and coupling losses must be reduced by the resistivity between them. Cr plating of about 2  $\mu$ m thickness appears to be suitable to fulfill those topics. Problems with the RRR of the stabilizing Cu may arise, due to the heat treatment. Thus measures have to be taken against this effect.

The Cu content of modified jelly roll and internal-Sn conductors is, in the basic design, larger than that of the bronze conductors. Due to the high Sn content in these composites, the  $J_c$ , in particular, the non-Cu  $J_c$ , taking into account the cross-section area without stabilizing portions, is superior as well. The Cu content can be varied up to above 60% (area). However, a minimum of about 30% is needed, in order to have an outer Cu shell, which is necessary for mechanical stability during processing.

### Stress and Strain

The mismatch of thermal expansion coefficients of the conductor components is creating compressive strain in the A15 layer during cooling. The difference from the heat-treatment temperature to 4.2 K is approximately 1000 K. The influence of

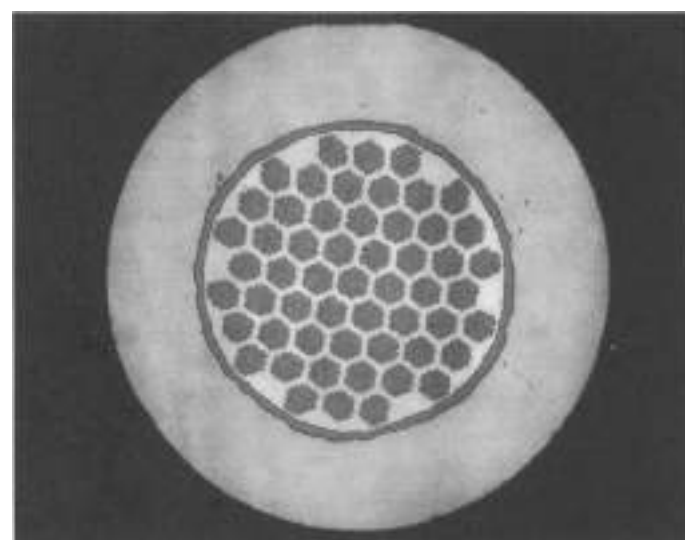
**Table 4. ITER Strands: Test Results and Specifications (38)**

| Technique   | Bronze             | Bronze    | Internal<br>Ta                                  | Internal<br>Ta               | Internal<br>Ta               | HP I<br>spec. 1 | HP II<br>spec. 1 |
|---|--------------------|-----------|---|------------------------------|------------------------------|-----------------|------------------|
| Diameter (excluding Cu-layer) (mm)                          | 0.903              | 0.902     | 0.901   | 0.896                        | 0.892                        | 0.91            | 0.81             |
| Barrier   | Ta                 | Ta        | Ta  | Ta + Nb                      | Ta/Nb                        |                 |                  |
| Thickness barrier (nm)                                      | 10-15              | 6-9       | 5   | 1 + 2 + 3                    | 3                            |                 |                  |
| Twist pitch (mm)  | 8.8                | 18.4      | 18.0  | 9.9                          | 9.3                          | < 10            | < 10             |
| Heat treatment (temperature and time (°C/h))                | 570/220<br>650/175 | 650/240   | 200/6<br>300/18<br>450/28<br>550/180<br>650/240 | 220/175<br>340/96<br>650/180 | 165/120<br>340/72<br>650/200 |                 |                  |
| Filament coated   | partially          | partially | fully   | fully                        | partially                    |                 |                  |
| Cu:Sn:Cu ratio  | 1:39               | 1:39      | 1:59  | 1:38                         | 1:61                         |                 |                  |
| Critical temperature $T_c^*$ at 12 T (K)                    | 10.2               | 10.2      | 9.31  | 8.49                         | 9.31                         |                 |                  |
| Upper critical field $B_c^*$ (T)                            | 26.3               | 27.7      | 24.8  | 25.3                         | 24.6                         |                 |                  |
| Cu thickness (nm)   | 2.1                | 3.2       | 2.6   | 1.7                          | 2.3                          | 2               | 2                |
| RRR (resistivity at 20 K / resistivity at 20 K)             | 150                | 147       | 130   | 80                           | 213                          | > 100           | > 100            |
| Overall strand density (g/cm <sup>3</sup> )                 | 9.33               | 9.14      | 9.04  | 6.98                         | 9.01                         |                 |                  |
| Non-Cu $J_c$ at 12 T, 4.2 K, 0.1 eV/cm (A/cm <sup>2</sup> ) | 550                | 570       | 780   | 710                          | 660                          | > 700           | > 500            |
| Non-Cu hysteresis losses, 3 T cycle at 4.2 K (mJ/cm)        | 94                 | 91        | 136   | 585                          | 589                          | < 100           | < 200            |
| $n$ value at 12 T, 4.2 K, 0.1 eV/cm                         |                    |           |   |                              |                              | > 20            | > 20             |
| Coupling loss time constant (ms)                            | 0.82               | 3.1       | 1.3   | 6.9                          | 6.3                          |                 |                  |

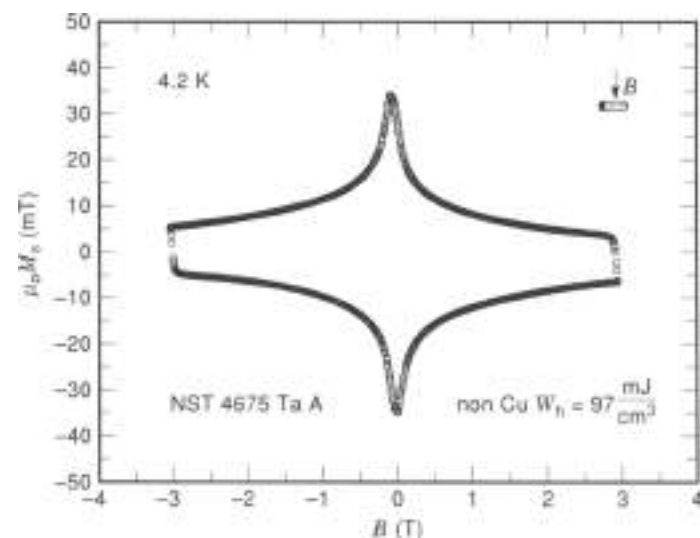
strain on the  $J_c$  values was presented first by Buehler and Levingstein in 1965 (39). The compressive strain influences the  $T_c$  and  $B_{c2}$  values as well. The relative difference in linear expansion,  $\Delta l/l$ , of separated bronze and Nb<sub>3</sub>Sn for  $\Delta T$  of 1000 K, is about 1.05% (32). Depending on the volume-ratio of Nb<sub>3</sub>Sn to the bronze and for a conductor that has good bonding between its elements, the contraction is between 0.00% and 1.05%. This is because the filaments are under tensile stress and the bronze are under compressive stress. Measurements on Nb<sub>3</sub>Sn with removed bronze matrix show a  $T_c$  close to the maximum 18.2 K. The values for prestressed conductors are reduced by about 1 K. Conductor designs which have besides the matrix and the twisted filaments, stabilizing Cu, diffusion

barrier, or reinforcing components, for example, Cu-Nb, are a rather complicated system. For example, regard the following values: A conductor with 22% (area) Cu and 5% Ta has a relative thermal contraction from room temperature down to 4 K of -0.29%, whereas conductors with 33% Cu and 10% Ta have -0.26%.

The critical values decrease by the compression but increase again under axial tension. Maximum  $J_c$  is gained at the strain  $\epsilon_m$ , where the tensile force in the filament is reduced to a minimum. The Young's modulus in the filament area is in the range of 130 GPa and for the bronze between 50 GPa to 80 GPa (40). The value for the bronze depends on the depletion of the Sn. By the diffusion process the Sn content is reduced and Kirkendall



(a)



(b)

**Figure 14.** (a) Cross section of an ITER conductor for the central sole-rod (HP II specification), i.e. low loss with a Cu:non-Cu ratio of 1:5 and 4675 (Nb, Ta) filaments (courtesy of Vaconschmelze). (b) Magnetization curve of the conductor shown in Fig. 14(a), indicating the low hysteresis losses of 97 mJ/cm<sup>3</sup> measurements: Vaconschmelze).

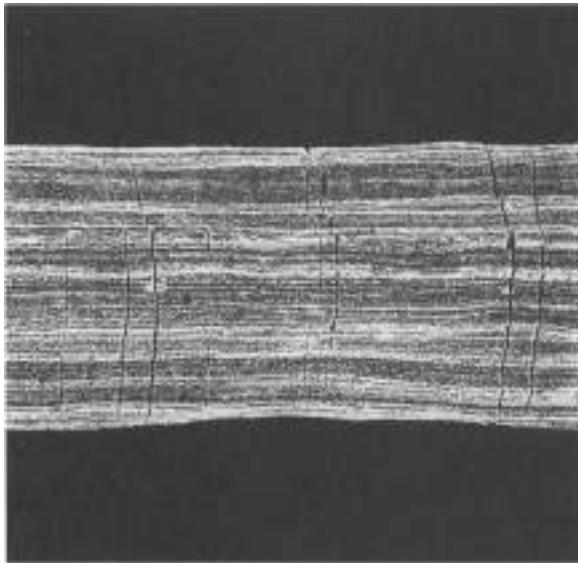


Figure 15. Scanning electron microscope micrograph showing Nb<sub>3</sub>Sn layers after strain of the conductor of 3% (courtesy VacuumSchmelze).

welds are created. These effects, the high temperature for annealing and the length of the heat treatment influencing the grain structure, are responsible for a low yield strength in the bronze. Plastic deformation of the bronze caused by tensile stress, due to mismatch of the thermal expansion coefficients (bronze:  $16 \times 10^{-6} \text{ K}^{-1}$ ; Nb/Nb<sub>3</sub>Sn:  $7 \times 10^{-6} \text{ K}^{-1}$ ), results in reduced differences in the thermal contraction. Bare Nb<sub>3</sub>Sn shows breaks at strain of about 0.2%. Prior to breaking, slip-steps with an angle of 45° can be observed. The pre-compression built up in the conductor has, as a consequence, the drop of  $I_c$ ,  $R_{20}$ , and  $T_c$ . The same pre-compression acts as a mechanical reserve during elongation of the conductor, though strain of 0.7% to 0.8% is possible for the complete composite without irreversible  $I_c$  degradation [42]. This decreasing under compression and increasing under tension is reversible, as long as a critical strain in the filament was not exceeded. If strain was too high,  $I_c$  would not recover completely after the load was released, because of cracks in the filaments; see Fig. 15.

The increased strain tolerance is an important reason for the fact that bronze conductors can be used as technical superconducting wires. Consequently, a thicker layer of Nb<sub>3</sub>Sn leaves a more Sn-depleted bronze with reduced Young's modulus. Possible plastic deformation and the reduced Young's modulus are diminishing the compressive strain of the filaments and, therefore, the degradation of the critical values is smaller. The mechanical reserve against axial stress is smaller, too. The filaments themselves are less strain sensitive if not fully reacted. Besides the thermal contraction during the cooling from about 1000 K to 4 K axial stress is applied to the conductor by several manners. The force used for winding a magnet is giving tension, while the bending force gives rise to tensile and compressive strain above or below the neutral wire axis. In the finished magnet, the Lorentz force  $F = J \times B$  leads to hoop stresses, related to the radius of the winding. Therefore,  $\sigma = J \times B \times r$ , where  $r$  is the radius, leads to very large forces, which makes special reinforcement measures necessary in the magnet or the winding package.

At the stress-compensated state,  $\epsilon_{01}$ , the upper magnetic flux density  $B_{c2}$ , which is strain dependent but always below  $B_{c20}$ ,

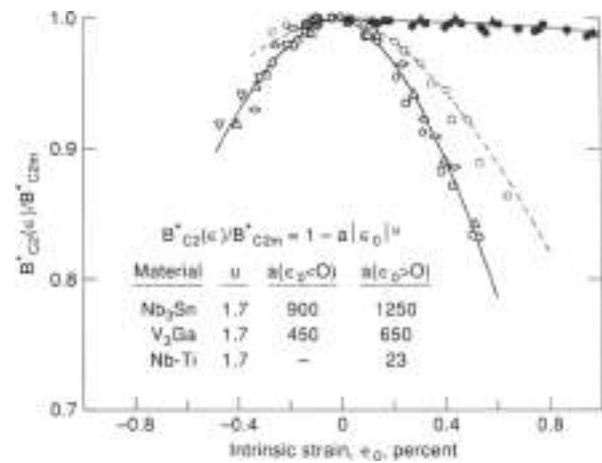


Figure 16. Upper critical field  $B_{c2}$  as a function of intrinsic strain  $\epsilon_0$  [41].

can be calculated from the measured  $I_c$  by using Kramer's law [41]:

At  $\epsilon_{01}$ , the axial strain of the conductor at which  $I_{c,max}$  is reached is in the range of 0.3% to 0.7%, leading to intrinsic strain  $\epsilon_{01} = \epsilon - \epsilon_{01}$ . The maximum critical current is not correlated to the magnetic field. Nevertheless, the difference of  $I_c$  at  $\epsilon$  and  $I_c$  at  $\epsilon_{01}$  strongly depends on the applied magnetic field. At 12 T, between the compressive states  $\epsilon$  and  $\epsilon_{01}$ , the difference in  $I_c$  is about 20% and at 16 T there is a factor of about 2. The detailed effects of intrinsic strain on the critical magnetic field  $B_{c2}$  and the critical current density  $J_c$  can be seen in Fig. 16 and Fig. 17. The influence of the other conductor components besides the filaments and the bronze, like stabilizers or diffusion barriers is, of course, not negligible. A description of  $I_c$ ,  $J_c$  concerning the dependency on both magnetic field and strain, is given by the strain scaling law [42]:

$$\frac{I_c}{I_{c0}} = \left[ \frac{B_{c2}^*(\epsilon)}{B_{c20}^*} \right]^n \left[ \frac{1 - B_{c2}^*(\epsilon)}{1 - B_{c2}^*(\epsilon_{01})} \right]^q$$

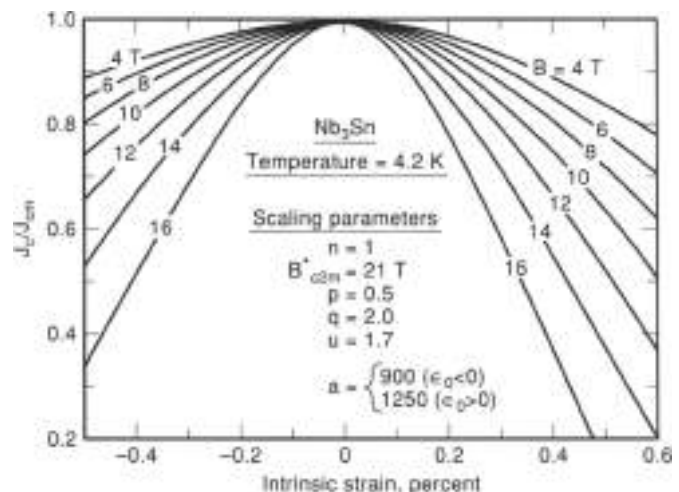


Figure 17. Relative critical current density  $J_c / J_{c0}$  as a function of intrinsic strain  $\epsilon_0$  for different magnetic field [42].

**Table 5. Scaling Parameters for the Use with the Strain Scaling Law (42)**

| Material           | $n$ | $p$ | $q$ | $B_{c20}$ |
|--------------------|-----|-----|-----|-----------|
| Nb <sub>3</sub> Sn | 1   | 0.5 | 2   | 21        |
| V <sub>3</sub> Ge  | 1.3 | 0.4 | 1.0 | 21        |

In this formula  $n$ ,  $p$ , and  $q$  are scaling parameters, which can be found together with the values of  $B_{c20}$  for Nb<sub>3</sub>Sn and V<sub>3</sub>Ge in Table 5.

The mechanical behavior of V<sub>3</sub>Ge depends mainly on the volume portion of V, while an increase of the Ge concentration reduces the tolerance with respect to mechanical loads. Conductors fabricated by the in situ technology show considerably higher mechanical values than filamentary bronze process conductors (43). There is not a clear  $I_c$  loss, but degradation is also smaller or completely recovered, respectively, after the load has been released.

Additions of third elements influence the  $I_c$ ,  $B_{c2}$ , and  $T_c$  values. Effects of strain can be seen as a function of  $B$  or  $B_{c2}$ . Due to the increase of  $B_{c2}$  by addition of Ti, Hf, or Ta to the matrix or to the core material, the effect of strain on  $B_{c2}$  is reduced. Further influence is given by the growth rate of the layer and, therefore, the remaining unreacted part of the core. Wires of Nb<sub>3</sub>Al are less strain-sensitive; even with an intrinsic strain of 0.6%  $I_c$  is reduced only by approximately 10%, as shown in Fig. 18.

The martensitic phase transition temperature  $T_c$  increases with the compressive strain, showing an influence of the cubic-to-tetragonal distortion of the lattice and the degradation of  $T_c$  and  $J_c$  (44). At the strain  $\epsilon_0$ , with  $J_c$  having its maximum, the Nb<sub>3</sub>Sn phase becomes cubic again. The effect of transverse compression on  $I_c$  is similar to that of axial strain. An increase

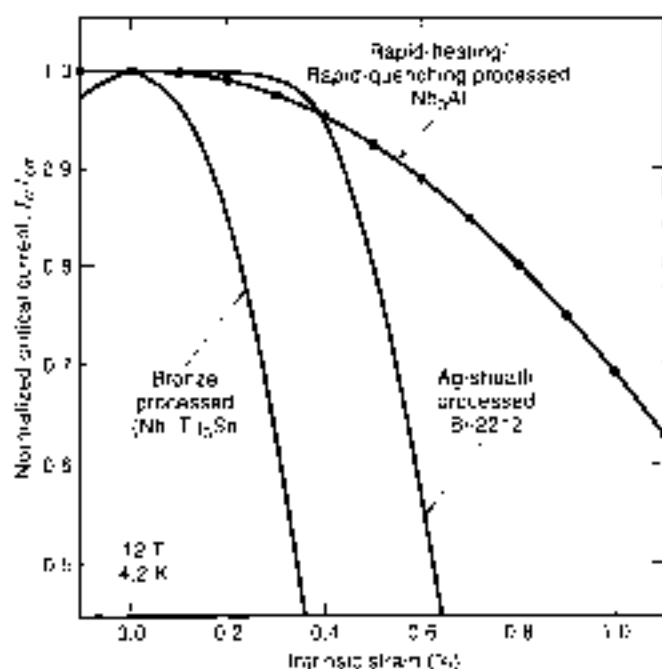
of transverse strain  $\sigma_1$  is leading to a small  $I_c$  enhancement. At higher  $\sigma_1$ , the  $I_c$  encounters a strong reduction. The sensitivity of  $I_c$  on transverse strain is higher than for axial strain and the irreversible behavior for  $\sigma_1$  starts at a level which is about 25% of the comparable axial strain (45). Transverse compression may occur in large magnet assemblies like Tokamaks, with each magnet having close and strong neighboring magnets. Because of the high currents necessary in such magnets, conductors are likely to be cabled and enveloped in a stainless-steel conduit. A mixture of compressive radial stresses and transverse pressures is obtained. Especially braiding procedures of the wires are leading to many crossover points with high stress concentrations. Tetragonal distortion is also caused by compression of the unit cell by radial stresses (46). Axial stresses are applied more frequently to conductors, but in view of the higher sensitivity to  $\sigma_1$ , the transverse compression has to be carefully taken into account for magnet engineering as well.

### HEAT-TREATMENT PRINCIPLES AND CONDITIONS

The enhancement of the critical properties of practical superconductors depends on improvements in the composite. It is further necessary to use production methods and diffusion treatments that are optimized toward the required features of magnets with high magnetic fields (>20 T) and high field homogeneity. Multiple, sometimes interacting measures are necessary to increase and to stabilize the values of  $J_c$  and  $B_{c2}$ . By far, not all effects in the many different A15 members are understood. Most of the information, gained by research work, is available for bronze conductors, especially for those of the Nb<sub>3</sub>Sn system, including internal-Sn and jelly roll conductors.

The speed of the reaction of the material in the course of the heat treatment, during which the A15 phase is formed, is correlated to the quotients of the atomic radii. The characteristic diffusion speed responsible for the ordering of the A ions to form the characteristic chains, depends on the radii of the B atoms. The smaller these radii are, the faster the diffusion process may occur (40). However, stoichiometric systems call for limits in the heat treatment. To have high concentrations of B ions, the solubility and workability of the components have to be shifted toward their limits. The thickness of the A15 layers formed per unit of time can be calculated by Fick's diffusion law, applied at the interface of two diffusion layers. The theoretical prediction for the amount of A15 phase that is formed, depending of the time  $t$ , is a proportionality of  $t^{1/2}$ . Due to influences such as grain growth or decreasing B ion concentration during the transition process, the kinetic parameters are reduced in relation to the theoretical prediction. Measurements of the speed of layer formation in Nb<sub>3</sub>Sn have shown proportionalities to powers of the time  $t$  in the range from 0.30 to 0.35 (32). In case where ternary or quaternary alloys are utilized, the mechanism of the diffusion process is similar, but the formation speed is influenced.

The pinning of fluxoids penetrating into type-II superconductors needs structural imperfections. Such imperfections may be grain boundaries, lattice disturbances, grain morphology, impurities, or combinations thereof. As the coherence length  $\xi$  in Nb<sub>3</sub>Sn is only about 8.5 nm, it is difficult to get a complete picture of the interactions which are necessary, in order to trap a fluxoid in the superconducting layers. In general, the increase of  $I_c$  depends on the heat-



**Figure 18.** Critical current  $I_c$  versus intrinsic strain  $\epsilon_0$  for bronze processed Nb<sub>3</sub>Sn wire, and a new Nb<sub>3</sub>Al wire processed by rapid heating/rapid quenching. The strain sensitivity of the Nb<sub>3</sub>Al compared with the (Nb, Ti)<sub>3</sub>Sn (30).

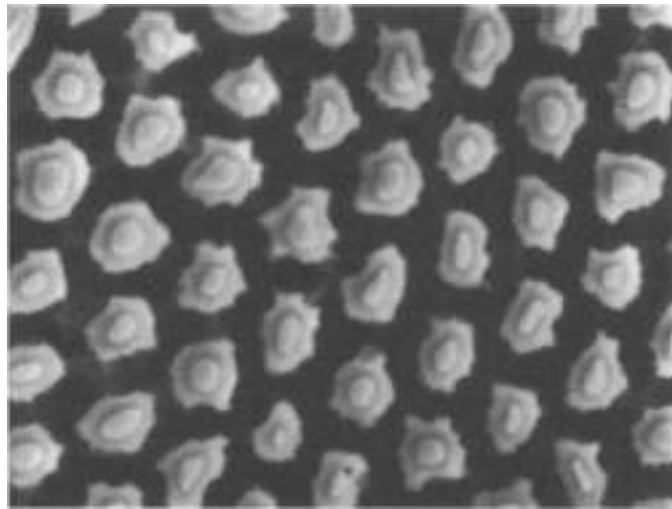


Figure 19. Layers (thickness = 1  $\mu\text{m}$ ) of Nb-Ti/Sn coated the adjacent Nb-Ti matrix of bronze conductor to improve its microstructure.

treatment temperature and time. Additional time and/or a higher temperature lead to a larger layer thickness. In Fig. 19, filaments with an A15 layer and an unreacted core of Nb in a bronze process conductor have been prepared so as to visualize the layer thickness of about 1  $\mu\text{m}$ . Nevertheless,  $J_c$  in the layer may be decreased by grain growth, because in the intermediate field range, the maximum pinning force is related to the average grain size and the grain size distribution in the A15 layer. Thus, the critical current  $I_c$  is inversely proportional to the grain size and, therefore, proportional to the number of grain boundaries per unit volume which act as pinning centers. Grain growth reduces the specific grain boundary area and diminishes the amount of fast diffusion paths. Diffusion at higher temperatures leads to faster layer formation and finer grain, but by depleting the bronze of B ions, Kirkendall voids are formed. This leads to a reduction in the diffusion rate. If those voids are located at the bronze-to-layer interface,  $I_c$  degradation may occur due to their influence on the strain behavior. Therefore, the gradient of the concentration of Sn over the cross section has to be taken into account for all heat treatment models. It is also important to have stoichiometric conditions, that is, a Sn concentration that is sufficiently high. For jelly roll and internal-Sn composites this can be achieved to a good degree. The so-called bronze route is limited by solubility and workability reasons to about 15 wt.% Sn. It is further necessary that the bronze be uniform, especially that the variation of the Sn concentration over the cross section is small. Filament sizes and filament spacings have to be watched, to leave sufficiently wide diffusion paths and to avoid bridging of filaments, in order not to end up with large effective filament diameters  $d$ .

### TERNARY ELEMENTS

The increase of  $J_c$  in the A15s is dominated at the intermediate field range by flux pinning at the grain boundaries. The pinning force density  $F_p$  is equal to the product of  $J_c$  and the corresponding magnetic flux density  $B F_p = J_c \cdot B$ . According to Kramer's law, the pinning force shows saturation in the

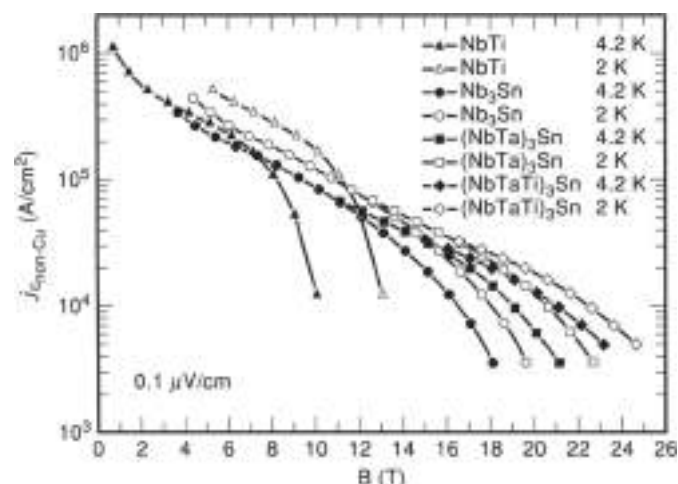
high field region  $B_{c2}$ . Increasing  $B_{c2}$  leads to an increasing  $J_c$  within the A15 layer. The value of  $B_{c2}$  is dominated by the normal state resistivity  $\rho_0$  and the critical temperature  $T_c$ . As it is not easy to increase  $T_c$  remarkably, the main means for varying  $B_{c2}$  is given by the normal state resistivity  $\rho_0$ , measured just above  $T_c$ , or by the resistivity ratio. Such an increase in  $\rho_0$  results in an increase of the Ginzburg-Landau-parameter  $\kappa = \lambda/\xi$ , where  $\lambda$  is the penetration depth (47). Because of the proportionality of the upper critical field  $B_{c2}$  to the Ginzburg-Landau parameter  $\kappa$ , the former is raised, too. Raising  $\rho_0$  is possible by impurities, irregularities in the chemical composition, causing microstructural defects.  $B_{c2}$  is not depending on the grain size, but its upper limit is determined by the susceptibility according to Pauli's paramagnetic effect (43). For that reason, the flux-pinning force and the grain boundaries are not relevant if the magnetic field  $B$  is close to  $B_{c2}$ .

Besides the more principal aspects, there are different other reasons which are influencing the performance of practical A15 superconductors. Disturbances in the microstructure are originating from the production process, chemical nonhomogeneities, or variations of filament diameters over the length (sausageing). Nonuniformity of A15 layers due to nonuniform distribution and supply of B ions is strongly influenced by the conductor design. These macroscopic effects are also observed for designs leading to irregular working and deformation conditions due to the combination of materials with quite different ductility, like Nb, Cu, and Sn. Microcracks occur in the layer itself, caused by thermal or handling defects. Bronze matrix conductors need, because of work hardening, intermediate heat treatments to preserve or to restore the ductility. Pre-reaction to a substantial degree may be encountered. It leads to heterogeneous deforming conditions, reduction of the Sn supply for the final diffusion treatment, and mechanical defects in the conductors. Therefore, intermediate heat-treatment temperatures must be chosen carefully and should not exceed 500 C. Additionally, time has to be restricted.

To achieve better properties of A15 conductors, doping with defined impurities like Zn, Mg, Fe, and Nb, and also alloying with higher contents of Ti, Ta, or Ga has been performed. The stoichiometry in a ternary or a quaternary compound is a rather demanding field. The variety of metallurgical treatments like alloying, in combination with numbers of different heat treatments, result in remarkable effects. Some of the additives, for example, in the bronze matrix, are influencing the diffusion process like speed and grain refinement, which enhances the performance of conductors in the intermediate field region. Others are more effective by increasing  $J_c$  at high fields due to an increase of  $B_{c2}$ . There is not a complete correlation given between composites and increase of  $J_c$  for the entire field range. This is not surprising at all, since additives may be deposited at the grain boundaries. Therefore, crossover of  $J_c$  versus  $B$  for doped or undoped conductors or for conductors doped with different additives is observed, as shown in Fig. 20.

The different atomic radii of the additives in relation to the base alloy is also of influence. Ti with a smaller atomic radius is incorporated more completely into the Nb lattice than zirconium (Zr) or hafnium (Hf) with larger atomic radii. The embedding of the Ti, originating from the matrix Cu-Sn into Nb-Sn, is at a larger degree than for the Ti alloyed with the Nb core. Elements like Ti, Zr, and Hf for alloying Nb result in A15 layers with fine grains. Mg in bronze, like Cu-Sn or Cu-Ga, is

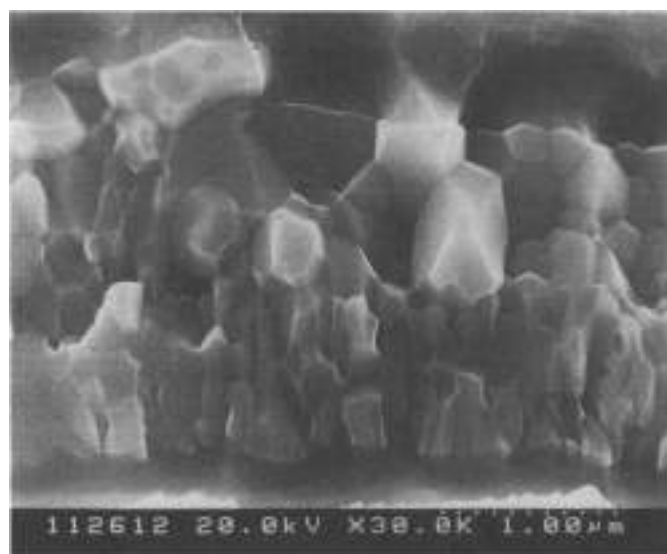




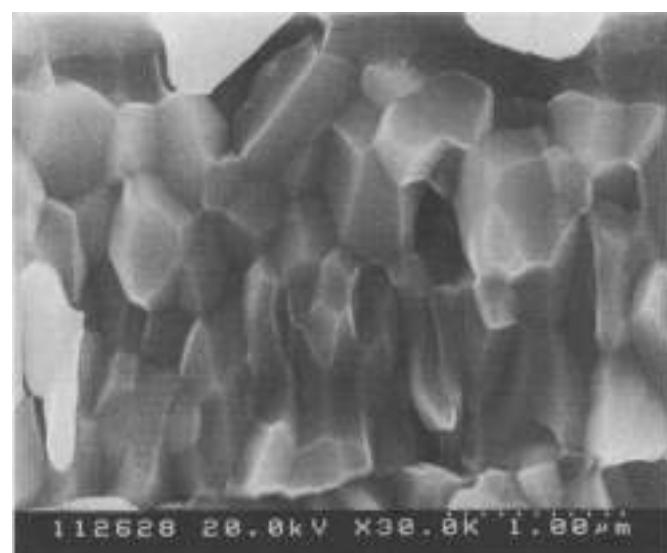
**Figure 20.** Non-Cu critical current density  $J_c$  versus magnetic field for  $\text{Nb}_3\text{Sn}$  (undoped, doped with Ta, or Ta and Ti, respectively) at temperatures of 4.2 K and 2 K (measurement: Vacuum-Schmelze)

increasing the formation rate of the A15 phase, which causes grain refinement. The low  $T_c$  of  $\text{V}_3\text{Ge}$  can be raised from 6 K to about 10 K by adding B at  $\text{V}$  (Al 143). Ti in Nb cores (about 2 at.%) is increasing the layer thickness. The improvement of  $J_c$  in  $\text{Nb}_3\text{Sn}$  by alloying Hf to the core and additional Ga to the matrix is remarkable and leads also to an increase of the irreversible strain  $\epsilon$  (43), but especially alloying with Ga is difficult and impractical.  $\text{V}_3\text{Ge}$ , which has already superior high field values, improves further above 20 T, when adding Ga to the V core and Mg to the Cu-Ga matrix. As Ti speeds up the diffusion rate of Sn in Nb, a doping with 0.8 wt.-% Ti, in combination with high-Sn-bronze (15 wt.-%) shows improved values of non-Cu  $J_c$  (33), see Fig. 10. By introducing Ge into the matrix of a  $\text{Nb}_3\text{Sn}$  composite, the thickness of the A15 layer is reduced significantly. This is most probably related to the formation of an additional phase with the Ge ( $\text{Nb}_4\text{Ge}_2$ ) at the interface between core and matrix (22). The grain size, however, is smaller, and so  $J_c$  in the layer is enhanced.

The heat-treatment time and temperature is in interaction with additional elements, or combinations thereof, responsible for the formation of the A15 layer. For wind-and-react technology, it is indispensable to limit the temperature, due to the insulation materials available. Practical glass-broid insulations for temperatures of up to 700 C and 500 C, respectively, are at hand. The temperature-time combination is further determined by the application of conductors. Additions of Ta, for example, Nb 7.5 wt.-% Ta, reduces the formation rate of the A15 phase (46), but increases  $J_c$  due to a longer heat treatment. In the case of intermediate magnetic fields, filament diameters should be smaller, leading, even with Ta doping, to a relatively short heat-treatment time, to realize, in this range of magnetic field, the required fine grain. Especially for conductors which are to be used at higher fields, the filament diameters are increased, as well as heat-treatment time and temperature. Often, more than one different cycle is executed, to improve the formation of the A15 layer, as not too much attention should be paid to grain growth, in view of the high field application. It is furthermore understood that, for each grain size, pinning interaction may be different, and flux line-lattice spacing is reason for different numbers of flux lines in



(a)

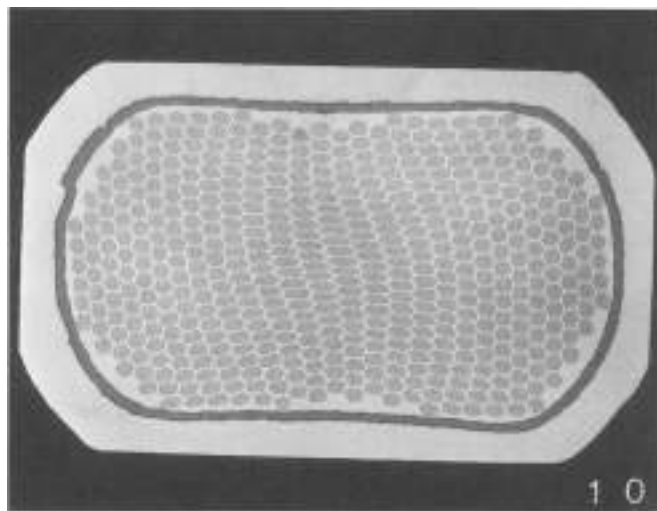


(b)

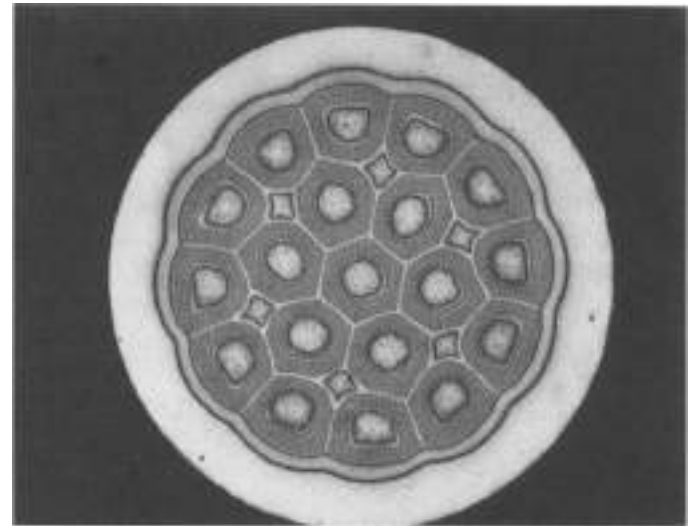
**Figure 21.** Microscopic photographs of reacted layers in fractured cross sections of samples heat treated at (a) 700 C for 100 h, and (b) 750 C for 150 h (courtesy of Kobe Steel).

each grain (49). The variability of the grain sizes for different heat treatments, but also within one sample, becomes evident in looking at the microscopic photographs in Fig. 21(a) and Fig. 21(b).

The necessity to fulfill the requirements of high critical currents  $I_c$  at high magnetic fields  $B_c$  leads automatically to large conductor cross-sections with a high number of filaments. Therefore, more than 100,000 filaments and cross section areas of more than 6 mm<sup>2</sup> are unavoidable. Figure 22 is an example of such an externally stabilized  $\text{Nb}_3\text{Sn}$  conductor. For several years, the demand for high magnetic fields and magnets with larger bores has been the driving force in the development of A15 superconductors. While the first successful magnets were built from tape conductors (10), round or rectangular wires became the more favorable solutions.



**Figure 22.** Externally stabilized  $\text{Nb}_3\text{Sn}$  conductor with a Ta diffusion barrier and outer Cu stabilization for high-resolution NMR magnets (courtesy of Vacuumschmelze).



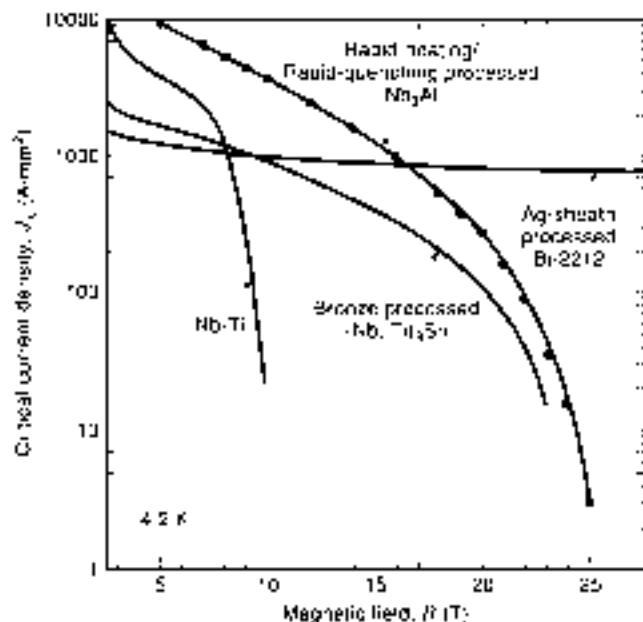
**Figure 24.** Cross section of an internal tin conductor for use in high-field dipole magnets (courtesy IGC).

For magnetic field strengths of above 9 T in the center of the magnet, materials with high  $B_{c2}$ , like conductors with A15 structure, are necessary. The study of polymers or macromolecules demands NMR systems of up to 1 GHz proton-resonance frequency (Larmor frequency); this corresponds to a magnetic field of 23.5 T. Thus all NMR systems working at frequencies of more than 400 MHz (corresponding to a magnetic field of 9.4 T) need other conductor material than NbTi. At a temperature of 2 K and a field of up to 21 T,  $\text{Nb}_3\text{Sn}$  still shows reasonable  $J_c$  values. Other composites, like the rapid-quenched  $\text{Nb}_3\text{Al}$ , may perform even better according to Fig. 23. Besides the NMR applications, which are dominated by bronze

matrix conductors, diffusion technology is substantial for conductors that are subject to a high magnetic field. For big machines like ITER, more than 1,000 tons of A15 conductors will be necessary.

Laboratory scale solenoids with magnetic fields of up to 20 T or high field split coils take advantages of the improved current densities  $J_c$  of, for example, internal-tin, modified jelly roll, or tube-processed wires. Special dipole and quadrupole magnets for accelerators have to reach magnetic fields of more than 11 T, which is only achievable with A15 conductors like the internal tin type shown in Fig. 24. Hybrid magnet systems, made of a combination of superconductive and resistive magnets, are capable of reaching fields of up to 45 T or 50 T. Not only the high critical fields  $B_{c2}$ , but also the superior values of the critical temperature  $T_c$ , are further advantages of these materials, which can even be used, with refrigerator cooling, for magnetic fields of 5 T to 10 T. Tapes of  $\text{Nb}_3\text{Sn}$  with a Nb-1% Zr-core are used to build split-coils acting as open MRI systems, operating at 4.2 K (16).

The results showing the improvements of the properties of A15 superconductors have been obtained, to a considerable degree, on an empirical basis. To force this technology forward, even more empirical work will be necessary. Fundamental tasks, like improvement of the composite, layer homogeneity, and grain morphology, will be inevitable for controlling the microstructural influences in future designs. Introducing artificial pinning centers (APC) and their possible variations and controlled grain refinement will be another route to be followed.



**Figure 23.** Non-Cu critical current density  $J_c$  versus magnetic field  $B$  for a bronze-processed  $\text{Nb}_3\text{Sn}$  and a  $\text{Nb}_3\text{Al}$  rapid-heating/rapid-quenching processed wire at a temperature of 4.2 K (20).

## BIBLIOGRAPHY

1. H. Hartmann et al., *Z. Anorg. Chem.* **198**, 116, 1931.
2. M. Weger and E. B. Goldberger, Some lattice and electronic properties of the  $\beta$ -tungstens, *Solid State Phys.* **28**.
3. W. Ruckel, *Supraleitung*, Weinheim, Germany: VCH, 1984.
4. J. Müller, A15 type superconductors, *Rep. Prog. Phys.* **43**: 1980.
5. G. F. Hardy and J. D. Hulm, *Phys. Rev.* **87**, 884, 1951.
6. B. T. Matthias et al., *Phys. Rev.* **86**, 1435, 1951.

7. W. Klose, *Sommerschule für Supraleitung*, Poggendorf, Germany: DPG, 1970, p. 14.
8. O. Henkel et al., *Supraleitende Werkstoffe*, Leipzig, Germany: VEB Verlag für Grundstoffindustrie, 1982.
9. J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Theory of superconductivity, *Phys. Rev.*, **108**: 1175, 1957.
10. B. W. Batterman and C. S. Barrett, *Phys. Rev.*, **13**: 390, 1964.
11. H. J. Williams and R. C. Sherwood, *Bull. Amer. Phys. Soc.*, **5**: 430, 1960.
12. British patent, GB.No 1203292, 1968.
13. W. D. Markiewicz et al., A 17.5 T superconducting core-cable Nb<sub>3</sub>Sn and VCEa magnet system, *IEEE Trans. Magn.*, **MAG-13**: 35, 1977.
14. K. Tachikawa and Y. Tanaka, *Japan. J. Appl. Phys.*, **5**: 804, 1966.
15. E. W. Callings, Processing of Nb<sub>3</sub>Al superconductors, Rep., Columbus, OH: Ohio State University, 1997.
16. C. G. King et al., Flux jump stability in Nb<sub>3</sub>Sn tape, *IEEE Trans. Appl. Supercond.*, **7**: 1524-1528, 1997.
17. E. W. Callings, Recent advances in multifilamentary Nb<sub>3</sub>Al strand processing, Rep., Columbus, OH: Ohio State University, 1997.
18. J. P. Kunzler et al., *Phys. Rev. Lett.*, **6**: 89, 1961.
19. M. R. Pickett et al., *Filamentary A15 Superconductors*, New York: Plenum, 1980, p. 331.
20. BMPT, *Entwicklung von Hochfeldsupraleitern*, BMPT-FB 1, 1976.
21. D. Lybellestier et al., Rutherford Lab. Report 74-135, *IEEE Trans. Magn.*, **MAG-11**: 247, 1975.
22. T. Pyon and E. Gregory, Some effects of matrix additions to internal tin processed multifilamentary Nb<sub>3</sub>Sn superconductors, *IEEE Trans. Appl. Supercond.*, **5**: 1760-1763, 1995.
23. M. Hansen, *Constitution of Binary Alloys*, New York: McGraw-Hill, 1958, p. 634.
24. E. Gregory et al., Development of Nb<sub>3</sub>Sn wires made by the internal tin process, *CEC/ICMC*, **AP 6**, 1997.
25. Y. Ikeda et al., Development of Nb<sub>3</sub>Sn superconducting wire using an in-situ processed large ingot, in *Advances in Cryogenic Engineering*, vol. 36a, New York: Plenum, 1990.
26. H. Fuji et al., Development of react and wind coils using in-situ Nb<sub>3</sub>Sn wires for ac applications, *MT15*, 1997.
27. S. Ceszara et al., *IEEE Trans. Magn.*, **MAG-15**: 639, 1979.
28. W. K. McDonald, *Expanded metal containing wires and filaments*, US patent 4414428, 1983.
29. R. C. Sharma, Multifilamentary VCEa wires and tapes with composite covers, *Cryogenics*, **25**: 381, 1985.
30. J. L. Jorda and R. Flückiger, Département de Physique de la Matière Condensée, *Genève. J. Low-Temperature Metals*, **75**: 227, 1990.
31. K. Inoue et al., *New Nb<sub>3</sub>Al Multifilamentary Conductor and Its Application to High-Field Superconducting Magnets*, Tsukuba-SEI, Ibaraki 305, Japan: RICEP, 1997.
32. K. Tachikawa, *Proc. ICSC*, **3**, 1970.
33. H. Hellmann, *Fabrication technology of superconducting material. Superconductor Material Science*, New York: Plenum, 1981.
34. T. Miyazaki et al., Improvement of critical current density in the bronze processed Nb<sub>3</sub>Sn superconductor, Paper CPB-7, CEC/ICMC, 1997.
35. H. Krauth et al., *Int. Workshop on High-Magn. Fields*, 1996.
36. Y. Miyazaki et al., Development of bronze processed Nb<sub>3</sub>Sn superconductors for 1 GHz NMR magnets, Paper CPB-6, CEC/ICMC, 1997.
37. M. Thoner et al., Aluminum stabilized Nb<sub>3</sub>Sn superconductors, *Adv. Cryog. Eng.*, **34**: 507, 1987.
38. E. Wüning et al., Evaluation of Nb<sub>3</sub>Sn superconductors for use in a 20.5 T NMR magnet, *IEEE Trans. Magn.*, **30**: 2344-2347, 1994.
39. H. C. Roggiers et al., Third round of the ITER strand benchmark test, EUCLAS conference applied superconductivity, no. 158, IOP conference series, 1997.
40. E. Buehler and H. J. Levingsstein, *J. Appl. Phys.*, **36**: 3856, 1965.
41. G. Rupp, The importance of being prestressed, in M. Suenaga and A. E. Clark, eds., *Filamentary A15 Superconductors*, New York: Plenum, 1980.
42. E. J. Kramer, *J. Appl. Phys.*, **44**: 1360, 1973.
43. J. W. Ekin, Stress-strain effects on critical current, *Cryogenics*, **35**: 825-828, 1995.
44. K. Tachikawa, Recent developments in filamentary compound superconductors, *Adv. Cryog. Eng.*, **28**: 1982.
45. R. W. Haard et al., The effect of strain on the martensitic phase transition in superconducting Nb<sub>3</sub>Sn, *IEEE Trans. Magn.*, **MAG-17**, 1981.
46. W. Sprickling et al., Effect of transverse compression on  $I_c$  of Nb<sub>3</sub>Sn multifilamentary wire, *Adv. Cryog. Eng.*, **34**: 569, 1988.
47. T. Lohmann and A. B. Seward, *Phys. Lett.*, **56A**: 355, 1976.
48. V. L. Ginzburg and L. L. Landau, *Exp. Test. Exp.*, **20**: 1064, 1950.
49. W. Sprickling, F. Weiss, and R. Flückiger, Effect of filament diameter and spacing on  $I_c$  of Nb<sub>3</sub>Sn wires in the intermediate field range and at high fields, *IEEE Trans. Magn.*, **23**: 1189-1191, 1987.
50. D. Rodriguez, Jr. et al., Flux pinning mechanisms in superconducting A15 materials and the optimization of their transport properties, to be published in *Adv. Cryog. Eng.*, **44**, 1998.

HEINHARD DIETRICH  
 Vauvasschmelze GmbH

## CRYOGENIC STABILIZATION

Kamerlingh Onnes (1) discovered superconductivity in 1911 when he was measuring the electrical conductivity of mercury as a function of temperature down to the temperature of liquid helium. He found that at 4.15 K the electrical resistivity ( $\rho$ ,  $\Omega$  m) became too small to measure. Although it is not possible to prove experimentally that a quantity is exactly zero, experiments to date have been able to show that the resistivity of a metal in the superconducting state is less than  $10^{-22}$   $\Omega$  m (compared to the resistivity of pure copper at low temperature,  $10^{-8}$   $\Omega$  cm) (2). Before long, a number of other elements were also found to exhibit this same phenomenon at similar low temperatures. A list of a few of these materials, called superconductors, with their transition (or critical) temperatures, is given in Table 1 (2-4).

Even at the time of its discovery, Kamerlingh Onnes realized that the phenomenon of superconductivity could have important technological uses. However, it was soon discovered that these early superconductors (the so-called type I superconductors) remained in the superconducting state only if they were not carrying a substantial electric current ( $I$ , A) and if they were not in the presence of any substantial magnetic field ( $H$ , A/m or T; see Table 1, footnote b). As the external magnetic field is intensified, or as the electrical current within the

Table 1. Properties of Some Superconductors\*

| Material  | Critical Temperature (K) | Critical Magnetic Field (T) <sup>b</sup> | Critical Current Density (A/cm <sup>2</sup> ) |
|---|--------------------------|--|---|
| Type I  |                          |  |   |
| Hg  | 4.15                     | $H_c = 0.04$                             |   |
| Pb  | 7.19                     | 0.08                                     |   |
| Sn  | 5.72                     | 0.03                                     |   |
| Type II   |                          |  |   |
| NbTi  | 9.1                      | $H_{c2} = 12.2$                          | $5 \times 10^8$                               |
| Nb <sub>3</sub> Sn  | 18.1                     | 25                                       | $5 \times 10^8$                               |
| Nb <sub>3</sub> Ge  | 23.2                     | 38                                       | $5 \times 10^8$                               |
| HTSC <sup>c</sup>   |                          |  |   |
| Bi <sub>2</sub> Sr <sub>2</sub> CaCu <sub>2</sub> O <sub>8</sub> (2212BSCCO)                | 90                       | $H_{c2} = > 1$                           |   |
| Bi <sub>2</sub> Sr <sub>2</sub> Ca <sub>2</sub> Cu <sub>3</sub> O <sub>10</sub> (2223BSCCO) | 110                      | 200                                      |   |
| YBa <sub>2</sub> Cu <sub>3</sub> O <sub>7</sub> (YBCO)                                      | 92                       | 30                                       |   |

\* Data from Refs. 2-4.

<sup>b</sup> It has become customary to use the units (T), equivalently a unit of magnetic induction, to express the value of the magnetic field.<sup>c</sup> At 77 K.

superconductor is increased, its normal-mode electrical resistivity is restored at a critical value of either current or field. Hence, the superconductors were inherently unstable under certain conditions. As Silsbee suggested in 1916, these are not two unrelated phenomena, because the magnetic field generated by passing a current through the wire destroys the superconductivity at the same value as does an externally applied field (5). The temperature at which superconductivity appears in zero applied magnetic field and carrying no current is called the critical temperature ( $T_c$ , K). However, the practical transition temperature (i.e., the maximum temperature at which a superconductor exhibits superconductivity) is a function of both external magnetic field and the current within the wire, as shown in Fig. 1. It can be seen that practical operation of a superconductor must be made within the parameter space under the surface  $F = f(I, H, T)$  shown, and thus below certain limits on temperature, current, and magnetic field.

Both the critical currents and the critical fields of the early superconductors were small. Consequently, although the

discovery of superconductivity was made early in twentieth century, its practical use for producing strong magnetic fields was not realized until much later. An efficient design of superconducting magnets for large physics experiments and energy devices has had to wait until near the end of the twentieth century (6).

By the 1950s, experimentation with intermetallic compounds and alloys had led to the discovery of materials that greatly relieved the above limitations. These materials remain in the superconducting state to somewhat higher temperatures (see Table 1) in higher magnetic fields, and also are able to transport larger electrical currents. They were given the name type II superconductors. A fundamental difference in the behavior of these two types of superconductors is in the way in which a magnetic field enters the material, as illustrated in Fig. 2(a) and 2(b). In type I materials, magnetic field is excluded from the superconductor [Fig. 2(a)]. However, when any one of the limits of critical temperature ( $T_c$ ), magnetic field ( $H_c$ ), and current ( $I_c$ ) is exceeded, there is an abrupt and total entry of any external field into the material as the material loses its superconductivity. In type II superconductors, above a threshold field  $H_{c1}$ , the field begins to enter in discrete units of field called fluxoids, and superconductivity is totally destroyed only when the field  $H_{c2}$  has completely entered the superconductor (see Fig. 2(b) and 2(c)).

The important characteristic of type II superconductors (alloys such as NbTi or intermetallic compounds like Nb<sub>3</sub>Sn) is their capacity to sustain high transport currents (7), which makes them suitable for use in high-current devices. However, for large enough magnetic field (and/or current) these superconductors eventually pass into the normal state as well. Consequently, these superconductors can also become unstable. For  $H < H_{c1}$ , a type II superconductor is in the superconducting state; for  $H_{c1} < H < H_{c2}$ , the superconductor is in a mixed state (the magnetic field penetrates into the regions existing in the normal state, but bulk superconductivity is not extinguished), and for  $H > H_{c2}$ , the superconductor is in the normal, resistive state (Fig. 2(c)). Type I superconductors have a critical field,  $H_c$ , below which there is no field within the material and it is superconducting (Fig. 2(a)). In contrast, type II superconductors

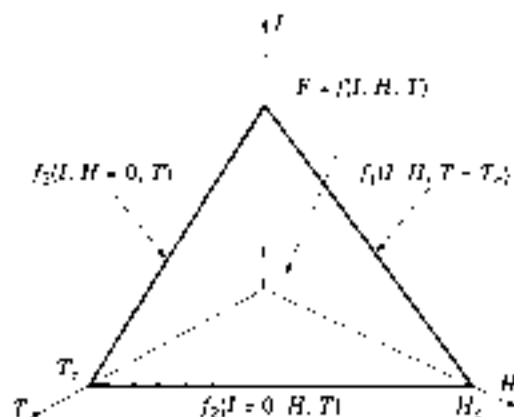


Figure 1. Critical linearized  $F(I, H, T)$  surface of a superconductor. It should be noted that in a typical magnet-grade superconductor the critical current  $I_c$ , magnetic field  $H_c$ , and temperature  $T_c$  are not linearly related. However, the classical theory of crystallinity assumes linear relationships.

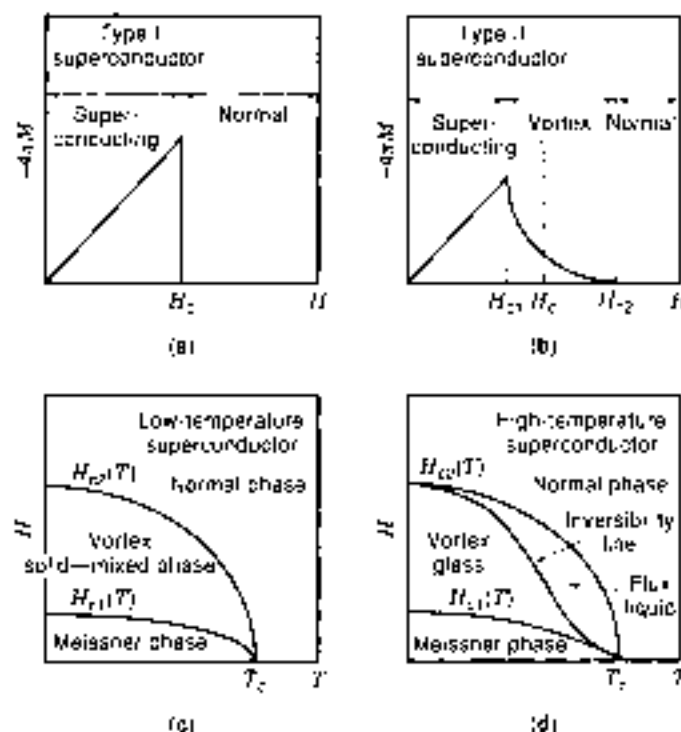


Figure 2. Superconductor types. (a) Magnetization (magnetic moment per unit volume,  $M$ ;  $4\pi M$  is also used) versus magnetic field for a type I superconductor. The magnetization increases abruptly from the Meissner value to zero. (b) Magnetization versus magnetic field for a type II superconductor. The magnetization decreases gradually from the Meissner value, hence providing three distinct regions: (1) the Meissner state (magnetic flux completely excluded), (2) the vortex state, and (3) the normal state. In the vortex state magnetic flux penetrates in the form of vortices but the material is still superconducting. The dotted line represents the comparison with the case for a type I superconductor. (c) Magnetic field versus temperature for a low-temperature superconductor. (d) Magnetic field versus temperature for a high-temperature superconductor.

have two critical fields: the lower critical field  $H_{c1}$ , at which the magnetic field begins to move into the superconductor, and the upper critical field  $H_{c2}$ , at which the penetration is complete and superconductivity is destroyed; see Fig. 2(b) and (c).

An important physical property of the intermetallic compound type II superconductors is that they are brittle and their fabrication into useful shapes may require that the superconducting compound be produced by heat treatment after the wire has been formed in its final shape (8). In the case of  $Nb_3Sn$ , the wire is drawn to its final dimension before heat treatment causes the reaction between the Nb and the Sn to form the compound.

In the early 1980s new compounds were discovered that were able to retain superconductivity up to very much higher temperatures. The first of these compounds was  $La_{2-x}U_xCuO_4$  with a critical temperature of 35 K (9). Since that time, other compounds with higher and higher operating temperatures have been developed. Now, practical superconductors that retain superconducting properties to temperatures near 135 K are available (4,9). These *high-temperature superconductors* (HTSCs) are different from the conventional superconductors in that they are complicated oxide compounds. In addition, they

are granular and exhibit their superconductivity only along certain planes within the crystal. Thus, in an application requiring a significant length, it is necessary to fabricate the finished superconductor with the grains aligned along the superconducting plane with no more than about 7° misalignment (9). A greater angle between the grains results in weak links between them, which drastically limits their ability to transmit current. Methods of addressing the alignment problem have been developed.

One type of HTSC consists of the oxides of barium, strontium, calcium, and copper, commonly referred to as BSCCO, one example being  $Ba_2Sr_2Ca_2Cu_3O_{10}$ , or 2223BSCCO. For the BSCCO compounds the oxide-powder-in-tube (OPIT) process (4) has succeeded in reducing the weak link problem to a manageable level by producing continuous conductors in lengths up to one kilometer capable of carrying acceptably high current densities. However, in a plot of critical magnetic field versus temperature, only a part of the area under the curve showing the phase space for superconductivity is available for the transport of current. Such a plot is shown in Fig. 2(d), which includes an additional line called the line of irreversibility, which varies considerably from one compound to another. This line shows a practical limit of current-carrying capability in that above it the flux is no longer pinned and the current-carrying capability vanishes, even though operation may still be within the superconducting envelope. The line of irreversibility is quite low for the highly anisotropic BSCCO: operation in magnetic fields above 1 T is only possible at temperatures below 40 K. Another HTSC is  $YBa_2Cu_3O_7$  or YBCO. The more isotropic YBCO compound has a much higher line of irreversibility, making possible operation at useful current densities in magnetic fields up to 5 T at 77 K (liquid nitrogen temperature). The OPIT process does not work to reduce weak links in YBCO, and useful YBCO conductors have only been produced by deposition onto specially textured substrates (10).

Like the intermetallic compounds of the type II superconductors, the HTSCs are brittle and difficult to fabricate. A major part of the effort to produce useful HTSCs has been toward their fabrication into useful shapes (9).

#### EXPERIENCE FROM EARLY MAGNET APPLICATIONS

The discovery of the type II superconductors allowed the superconducting state to be maintained in the presence of higher currents and at more elevated magnetic fields (compared to the previous situation with type I superconductors). Applications of superconductivity have been made in instrumentation devices, and a considerable amount of investigation into superconducting electric power transmission lines has been done (11, 12), but no such power lines have been built to date. The greatest application of superconductivity on a large scale has been in the building of magnets with very high field capabilities (13).

The achievement of the higher operating parameters of the type II superconductors was not in itself sufficient for the construction of high-field magnets with maximum performance. All the superconducting magnets designed until the early sixties of the twentieth century, without exception, suffered from so-called degradation, i.e., from the loss of the superconducting state before reaching the full design field. This phenomenon has been attributed to the fact that internal and/or external disturbances trigger an irreversible instability causing the

appearance of normal zones followed by a so-called quench (the destruction of the magnetic field, with the stored energy being converted to heat). Degradation can be attributed to a number of sources, which may be mechanical, magnetic, or thermal in nature. Any source of heat in the system can cause local temperature increases in the superconductor to the point where the appearance of a normal resistance region is imminent. In early attempts to build large magnets it was discovered that because of effects like these, the performance of superconducting wires fell far short of the short-sample results, a phenomenon called coil degradation (14). A magnet would quench after reaching only a fraction of its design field. However, after each such quench, the next try resulted in a higher operational field. After repeated attempts, the field attained can be considerably higher than that attained at first, but not as high as the short-sample results. This phenomenon is called *training*.

At first it was thought that degradation was the result of weak spots in the long lengths of superconducting wire. However, this would not explain training, and production techniques have been shown to produce remarkably uniform properties in long lengths of superconducting wires (many kilometers, even in small magnets). It was concluded that degradation must be caused by one or more of a number of possible disturbances, such as a source of heat or the penetration or rearrangement of the magnetic field.

Such disturbances can lead to severe consequences, in which both functional and mechanical integrity of the device can be compromised. It can become unstable. An engineering application cannot tolerate such an instability. It has become obvious that a new technological solution is needed to overcome these difficulties.

#### CAUSES OF INSTABILITIES (DISTURBANCES)

There are several mechanisms that can cause the generation of heat within a superconductor carrying current in a magnetic field. Because of the very low heat capacity of most materials at the low operating temperature, the generation of heat can cause a sufficient rise in temperature to create a local zone in the superconductor that is above the critical temperature and therefore has normal resistance. The normal zone then becomes an additional source of heat because of Joule heating ( $I^2R$ ,  $R$  being the electrical resistance,  $\Omega$ ). The consequent increase in temperature can exacerbate one or more of the processes creating the heat. This positive feedback (without a stabilizing influence) can lead to a runaway situation. Regardless of the cause of instability, an uncontrolled growth of the normal zone will promote further heat generation, and in the worst case cause a quench, and all of the energy in the entire assembly, which may be hundreds of megajoules, can be released as heat, generating a pressure increase in the cooling medium and overheating the entire assembly, possibly even catastrophically. The prevention of such an occurrence is a task of stabilizing the superconductor assembly.

Disturbances to superconducting systems can be abrupt and cause local heating, so that a maximum local effect results. Other disturbances are prolonged in nature and tend to heat larger portions of the system, so that their effect is distributed more widely. Causes of disturbances have been classified as magnetic, mechanical, and thermal, as indicated in Fig. 3.

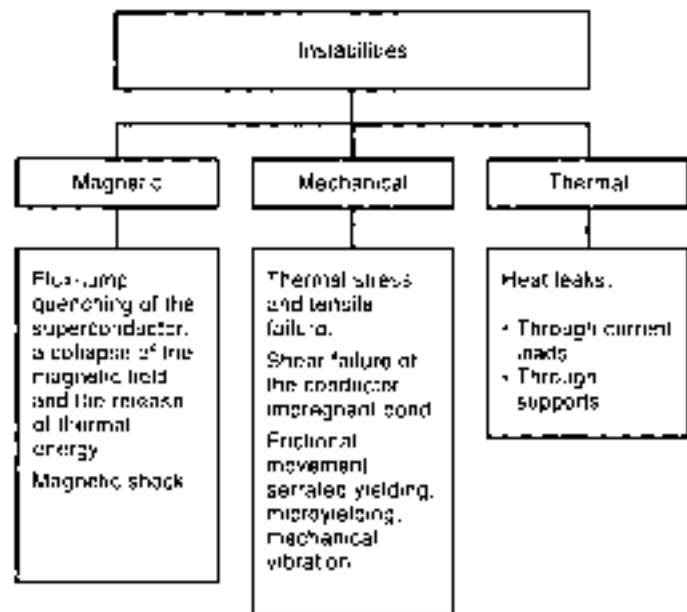


Figure 3. Instability sources. The instability disturbances are either magnetic (notably a flux jump), mechanical (mostly responses to thermal stresses), and thermal (such as heat leaks).

The forces generated by the magnetic field tend to stretch the magnet structure radially while causing it to contract axially (analogous to the reaction to a pressure increase of a cylindrical pressure vessel with ends closed by sliding pistons). This straining of the magnet system can cause sliding friction, or even the sudden rupture of components such as epoxy used for potting. Any sudden motion like this will generate heat as the motion is stopped. Once the external magnetic field reaches  $H_c$ , magnetic field penetration begins and the motion of the fluxoids also generates heat. Because the magnet system is operated well below ambient temperature, there is a constant input of heat, which may heat some portions of the system unequally. Any one of these disturbances can result in a normal zone within the superconductor.

#### Flux Jumps

The most serious disturbance is caused by flux jumps. In the mixed state, magnetic flux is present within the superconductor. In the presence of a current passing through the superconductor the fluxoids are acted upon by the Lorentz force, which is perpendicular to the surface of the wire and is proportional to the product of the current density  $J$  ( $A/m^2$ ) and the magnetic induction  $B$  ( $T$ )—in vector form,  $J \times B$ . The fluxoids are pinned to the crystal lattice by imperfections or impurities, so that motion does not occur unless a disturbance creates a location where the Lorentz force exceeds the pinning force.

Flux-jump instability is a characteristic of the superconductor itself, not the magnet as a whole. According to Wada et al. (15), fluxoids can be defined as quantized magnetic flux lines distributed within a type II superconductor. Degradation caused by a flux jump is related to a sudden motion of the fluxoid within the superconductor (16). The collective, discontinuous motion of fluxoids in such a superconductor can be caused by mechanical, thermal, magnetic, or electrical disturbances. Breakaway of the moving fluxoid vortices is

inherently associated with Joule heating. The increase of the local temperature may be sufficiently large to cause an avalanche of a fluxoid movement, thus creating additional Joule heating and a subsequent increase in temperature. This redistribution of the magnetic field in the superconductor is called a flux jump.

#### Mechanical Disturbances

Mechanical influences may be responsible for a disturbance in a superconductor that can lead to instability (17). For example, during the cooldown period before the startup of operation, as well as during the subsequent period of charging, stresses of various origins are always present in superconducting windings. These may be the thermoelastic stresses caused by the difference in thermal contraction of composite components such as superconductors, normal conducting material (e.g., copper), and epoxy impregnate. Similar stresses can arise from thermal gradients within the system, and stresses may arise during the manufacture of the composite conductor (drawing) and magnetic coils (winding and handling). Stresses are also caused by magnetic forces during charging. Small conductor movements activated during field cycling may cause local heating (18).

These stresses may lead to yielding or even tensile failure of a component in the magnet system such as the pulling resin, or to shear failure of the bond between an epoxy impregnate and the conductor. If thermal or magnetic stresses become great enough to cause some of the magnet material to yield, the yielding can take place in a series of discontinuous jumps (serrated yielding), which, because of their rapidity, cause rapid heating. Heat is produced as the motion is arrested, because of the low temperature, many materials are subject to cold embrittlement. Brittle fracture can be more serious than serrated yielding because the energy is all deposited locally.

The friction between the metallic and insulated surfaces can cause local heating. If this stress results in a stick-slip process, the effect will be similar to the brittle fracture mentioned above.

Any of these mechanisms can constitute a disturbance energy that may be strong enough to initiate a quench. Studies of thermomagneto-mechanical instabilities are numerous (19-22); however, the problem is open for further studies.

#### Distributed Heat Sources

Some of the above disturbances can result in slower, more distributed heat sources. However, there are additional sources of heat that are distributed widely through the magnet system. Although these heat inputs may not directly cause local heating within the magnet system, the heat transfer to the coolant must be sufficient to absorb the corresponding energy input.

**Heat Leaks.** Because essentially all superconductors are operated at temperatures in the cryogenic range, an ever present source of heat is the heat that enters the system through the thermal insulation, piping, and the support system. No insulation is perfect, so that this heat source is inevitable. Where, and how much, heat enters is determined by the design of the cryogenic system. The heat leak through the thermal insulation is usually distributed uniformly. However, the heat admitted through current leads, instrumentation, and piping may be more localized.

**Hysteretic Losses.** In some cases superconductors must carry an alternating or pulsed current. In these cases, a continuous

energy loss occurs and manifests itself as a heat source because of the Lorentz force on the uncoupled flux lines that move in and out of the superconductor as a result of the varying current. These losses are called hysteretic losses (3). They are influenced by the roughness of the surface of the superconductor. The smoother the surface, the lower the losses. Even more important is the twist pitch of the superconducting filaments and the resistivity of the matrix material (discussed below).

#### METHODS OF STABILIZATION

Because there are inevitable inputs of heat to the system as well as disturbances that can affect the temperature of the superconductor, it is necessary to provide methods to maintain the system at its operating temperature to prevent the catastrophic loss of the superconductivity. In the case of a superconducting magnet, the key problem of stability is that of sustaining the successful operation of the magnet system without loss of the magnetic field and without damage to the system. The problem is not only a possible appearance of normal zones within the superconductor, but the need to sustain conditions for the reestablishment of the superconducting state after the appearance of instability. Furthermore, the continuous operation of the device during the transition period must be assured.

*Flux-jump stabilization* suppresses an initiation of a cycle of disturbances that may cause the transition into the normal state. Thus, it is directed toward the prevention of an instability. The role of *cryostabilization*, in contrast, is to restore the superconducting mode of operation, once a disturbance has already initiated the existence of a normal zone.

There are several ways to approach the problem of stabilizing a superconducting system. First, the design of the superconductor wire can be used to minimize, or possibly even prevent, the damaging effect of a flux jump. The energy released during the passage of the flux lines through the conductor is proportional to the distance traveled across the wire. Thus the conductor is usually fabricated to consist of many very fine filaments (10  $\mu\text{m}$  to 100  $\mu\text{m}$  in diameter). These filaments are distributed within a matrix of a metal, which, while not a superconductor, does provide a highly conductive path for the current in case the superconductor can no longer handle the entire current. Typically the matrix metal is made of copper. However, for HTSCs silver is also used. One of the ac loss mechanisms arises from a coupling of the currents in adjacent superconductor filaments, and this loss can be minimized by twisting the superconductor composite so that the filaments are transposed along the length of the wire (3). An additional measure that can be directed toward a suppression of the ac losses is adding a more resistive metal, such as Cu-Ni, around the outer surface of the superconductor, thus decoupling the superconductor filaments and reducing the eddy currents flowing in the normal metal matrix (3).

Second, heat generated within the magnet system must be removed to maintain an acceptable operating temperature. This cooling can be provided by pool boiling in a suitable cryogenic fluid, or by the forced flow of the cryogen, either as a liquid, or as a fluid at supercritical pressure to avoid the problems of two-phase flow. In some cases where very low-temperature operation is required, use has been made of superfluid helium to obtain its excellent heat transfer capabilities. In this case an operating temperature in the vicinity of 2 K must be used.

The stabilization of the HTSCs utilizes the same methods as those for type II superconductors. In one respect, the task is eased by the fact that the specific heats of the magnet system become considerably higher as the temperature is raised. Thus the temperature increase from a given amount of heat is less than it would be at lower temperature. Another factor in favor of stabilization comes into play when the system is operated at the higher end of its temperature range. In this case, the current cannot be as high, which results in lower Lorentz forces, and thus the likelihood of flux jumps is diminished.

### Cryogenic Stabilization

According to Reid et al. (23), cryogenic stabilization is achieved if, after the release of a certain amount of Joule heat and a local rise in temperature (caused by either internal or external perturbations), efficient cooling is provided to remove that thermal energy more rapidly than it is generated. This goal assumes balancing of the following energy flows: (1) the energy brought to an intrinsically stable superconductor in the form of a thermomechanical disturbance or thermal energy from any source, including that generated by the current that is redirected into the stabilizing, resistive matrix, and (2) the energy removed from a superconductor element by convective cooling and by conduction. These processes can be highly transient in nature; indeed, their duration is usually very short (on the order of  $10^{-2}$  to  $10^{-3}$  s). Consequently, the use of a steady-state energy balance should lead to a very conservative stability criterion.

Cryostabilization can be full or limited. Full cryostabilization means stable operation after the entire conductor has been driven normal by a large disturbance (23). Design based on full cryostabilization is as a rule the most conservative and involves large conductors. Limited cryostabilization refers to recovery from a disturbance of limited size. By using the *stability criteria*, one can decide whether the superconductor is going to be stable (in the sense of either full or limited stabilization) or unstable. For example, full cryostability can be defined either using the so-called Stekly criterion or the Maddox-James-Norris model (see below).

All the early-developed stability criteria discussed in the following sections have been based on quasi-steady-state balances. The theory behind these static criteria will be (somewhat arbitrarily) called the classical theory of cryostabilization. The formulation of a transient problem and some of the issues involved will be given subsequently.

**The Stekly Criterion.** Stekly and his collaborators (24-26) were the first to formulate a method to prevent a catastrophic quench in a magnet. They discovered that reliable stabilization can be achieved by the simultaneous application of two measures: (1) an alternate path for the current through an adjacent material (such as copper) with high electrical conductivity (although nonsuperconducting) on the appearance of normal zones in the superconductor, and (2) very efficient cooling, such as in liquid helium. The so-called current sharing between a superconductor that abruptly loses its superconducting capability (becoming highly resistive) and the normal conductor (the matrix in which the superconductor is embedded, having a smaller resistivity than the adjacent superconductor in its normal state) secures a continuation of the magnet operation. By the intense cooling, the conductor can then be cooled back below the critical temperature for a given magnetic field and

electrical current. This, in turn, leads to a reestablishment of the superconducting state in the coil. Therefore, heat transfer to a cooling stream in addition to conduction through the conductor and the assembly will remove the generated heat and provide the conditions for the superconductor to reassume the total current flow, along with a disappearance of the current flow through the normal conductor, but without losing the magnetic field. Consequently, the coil becomes stable, and no quench will result. This approach constitutes *cryogenic stabilization*.

**The Concept of Stability and the Energy Balance.** The key nomenclature needed to define exactly the concept of stability is the energy balance of a conductor and/or magnetic device (a coil), which may be described in general as follows:

$$\begin{aligned} \text{Time rate of change of thermal energy of the conductor} \\ &= \text{conduction heat transfer rate} \\ &+ \text{rate of thermal energy generation} \\ &- \text{convection heat transfer rate} \end{aligned}$$

This balance assumes a composite conductor (material surrounded by a coolant and subject to transients as well as internal and external instabilities). The goal of a design is to keep the operating point of the conductor within the limits imposed by the critical surface ( $T_c, I_c, H_c$ ), but with an additional requirement formulated as follows: If the disturbance upsets this operation, the restoration of the superconducting state is still possible.

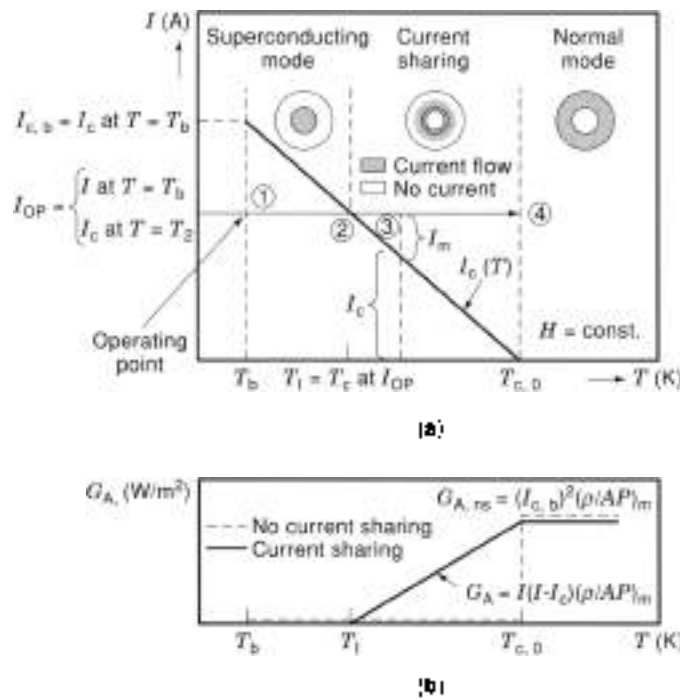
Let us first formulate the classical theory of cryostability, the concept originally introduced by Stekly and collaborators (25,26). We will not present this approach either in its entirety or in a chronological perspective. Rather, we will discuss the main points, which depend upon an energy balance that is astonishingly simple compared to the complexity of the superconducting instability phenomena. This simplicity commends this criterion as the most conservative of several that have been developed over the years.

First, the physical background for the analysis should be emphasized. Three distinct physical situations can be distinguished in a conductor operating at a given magnetic field  $H$ , carrying constant current  $I$ , and operating at various temperatures  $T$ . These situations are indicated in Fig. 4(a) as a superconducting mode, a current-sharing mode, and a normal mode.

In a (completely) superconducting mode, the conductor's operating current is less than critical, for a less than critical magnetic field, at a less than critical temperature. In a current-sharing mode, current flows partly through the matrix because of the appearance in the superconductor of normal zones characterized by a resistance that is much higher than the resistance of the matrix. Finally, when the temperature increases above the critical temperature, the superconductor operates in the normal mode and the current is carried exclusively by the matrix.

In an adjacent diagram [Fig. 4(b)], the corresponding distribution of thermal energy generated by Joule heating is presented. In the superconducting mode, Joule heating is zero. In a current-sharing mode, thermal energy generation per unit of area of the conductor surface is equal to  $G_s = H_n(\rho_s AP)_n = (I - I_c)(\rho_s AP)_n$  (W/m<sup>2</sup>), where  $\rho_s$  ( $\Omega \cdot m$ ) is the electrical resistivity of the resistive part (matrix), and  $I_n = I - I_c$  (A) is the current through the matrix. The symbols  $I$  and  $I_c$  represent the operating and critical currents as defined in Fig. 4(a). The





**Figure 4.** In  $I$ - $T$ - $H$  characteristic of a conductor: a linearized relationship  $I(T, H \text{ fixed})$ ; see also Fig. 1. Increase in the local temperature at given operating current  $I$  forces a superconductor to change from superconducting mode (state 1) to a current-sharing mode (state 2), and ultimately into the normal mode (state 3). In a current-sharing mode a part of the current flows through the resistive nonsuperconducting matrix and a part through the partially resistive superconductor. In the normal mode current flows only through the matrix, being excluded from the resistive superconducting part by its much higher electrical resistivity. Thermal energy generation by Joule heating versus temperature for a given current. In the case of no current sharing, Joule heating starts abruptly at, say  $T_c$  (dotted line). In the case of current sharing, thermal energy generation increases linearly with  $T$  in the current-sharing zone. In the superconducting mode there is no Joule heating ( $G_A = 0$ ). In the normal zone the Joule heating has a constant value (for a given current).

quantities  $A_m$  (m<sup>2</sup>) and  $P_m$  (m) are the cross-sectional area of the matrix and the conductor perimeter, respectively (Note that the units for electrical resistivity, area, and length, often used in practice, are  $\Omega \text{ cm}$ ,  $\text{cm}^2$ , and  $\text{cm}$ , respectively. Then the generated thermal energy per unit of conductor area is in  $\text{W/cm}^2$ . In a normal mode, say for  $I = I_{c,b}$ , thermal energy generation has a constant value of  $I^2(\rho/AP)_m$ .

To start the analysis we introduce a series of far-reaching assumptions but still preserve the main features of the process.

- The conductor is a composite, that is, it consists of a superconducting core and a highly conducting (but nonsuperconducting) matrix.
- The heat transfer phenomena involved, including the eventual release of thermal energy caused by a disturbance, are quasisteady.
- The temperature of the conductor is uniform across its cross section, and there is no heat conduction through the conductor.
- The electrical resistance between the core and matrix of the conductor is negligible.

- The electrical resistivity of a normal zone in the superconductor is several orders of magnitude larger than the resistivity of the matrix (thus, if the conductor operates in a normal, resistive mode, the electrical current tends to flow through the matrix).
- The electrical and thermal properties of the conductor materials and kinetic properties of the processes involved are independent of temperature.

The general energy balance equation in this case should include only thermal energy generation and heat convection, that is,

$$\begin{aligned} &\text{rate of thermal energy generation} \\ &= \text{heat transfer rate due to convection} \end{aligned}$$

The concept of cryostability assumes that the rate of thermal energy generation caused by instability is, in a limit, equal to the rate of heat removal by either pool boiling or forced convection of liquid helium (usually supercritical, and in some cases superfluid). This assumes that current may flow partly through the superconductor and partly through the matrix (current sharing) or it may be completely rerouted to the matrix if the superconductor becomes resistive (normal mode of operation). If the energy generated by Joule heating (both in the superconductor and in the matrix, in the current sharing mode, or entirely through the matrix, in the normal mode) is more than compensated by the heat removed from the conductor, its temperature will return to below the critical temperature, and the conductor will stay stable. So the stability criterion can be expressed as follows:

$$\frac{\text{rate of thermal energy generation}}{\text{heat transfer rate due to convection}} = \sigma$$

$$\sigma \begin{cases} < 1 & \text{(stable)} \\ = 1 & \text{(energy balance)} \\ > 1 & \text{(unstable)} \end{cases}$$

where  $\sigma$  is the so-called Stekly stability parameter. The above defined cryostability criterion can be written in an explicit form taking into account the energy generation under the condition that all the current is flowing through the matrix at the onset of normal mode of conductor operation. In this case the density of heat transfer rate generated within the conductor and normalized to the unit of heat transfer area between the conductor and coolant is given by

$$G_A = \frac{I_c r V_1}{P} = I_c r I_{c,b} \left( \frac{\rho}{AP} \right)_m = I_c^2 r \left( \frac{\rho}{AP} \right)_m = \frac{I_{c,b}^2 r}{4} \left( \frac{\rho}{A_m} \right) \quad (1)$$

where  $V_1$  (in  $\text{V m}^{-1}$  or  $\text{V cm}^{-1}$ ) is the voltage drop per unit length of the composite conductor. The right-hand side in the last equality of Eq. (1) expresses the thermal energy generation in terms of the current density  $(J = I/A)$ , the conductor diameter  $D$ , and the so-called filling factor  $\lambda = A_{superconductor}/A_{total}$ . The heat transfer rate density  $Q_A$  ( $\text{W/m}^2$  or  $\text{W/cm}^2$ ) on the matrix

surface is

$$Q_A = h(T_c - T_b) \tag{2}$$

Therefore, the Stekly parameter is given as

$$\alpha = \frac{f_m I_c b}{k P A_m (T_c - T_b)} \tag{3}$$

$$\alpha \begin{cases} < 1 & \text{stable superconducting state} \\ = 1 & \text{normal-state equilibrium} \\ > 1 & \text{unstable state} \end{cases} \tag{3}$$

When a composite conductor operates under current sharing, the total current  $I$  is the sum of the currents flowing through the superconductor,  $I_s$ , and the matrix,  $I_m$ . In general,

$$I_m = I - I_s : \begin{cases} I_m = 0, \quad I = I_c & \text{(superconducting)} \\ 0 < I_m < I & \text{(current-sharing)} \\ I_m = I & \text{(resistive)} \end{cases} \tag{4}$$

Under the current-sharing conditions, the voltage drop per unit length of the composite conductor,  $V_L$ , is given as follows [27]:

$$V_L = I_m \frac{\rho_m}{A_m} = (I - I_s) \frac{\rho_m}{A_m} = \left[ I - I_c \left( 1 - \frac{T_c - T_b}{T_c - T_b} \right) \right] \frac{\rho_m}{A_m} \tag{5}$$

Here a linearity of the critical  $I$ - $T$  curve has been assumed.

$$\frac{V_L}{I_c \rho_m / A_m} = \frac{I}{I_c} - 1 + \frac{T_c - T_b}{T_c - T_b} \tag{6}$$

or

$$v = i - 1 + \theta \tag{7}$$

where  $v = V_L / (I_c \rho_m / A_m)$ ,  $i = I / I_c$ , and  $\theta = (T_c - T_b) / (T_c - T_b)$  are the reduced values of the voltage drop, current, and temperature, respectively. The modes of conductor operation can be presented formally using a current-sharing factor [27]:  $f$  defined as follows:

$$f = \frac{I_m}{I} = \frac{I - I_s}{I} = 1 - \frac{I_s}{I} = 1 - \frac{I_c}{I} \frac{I_c b}{I} \tag{8}$$

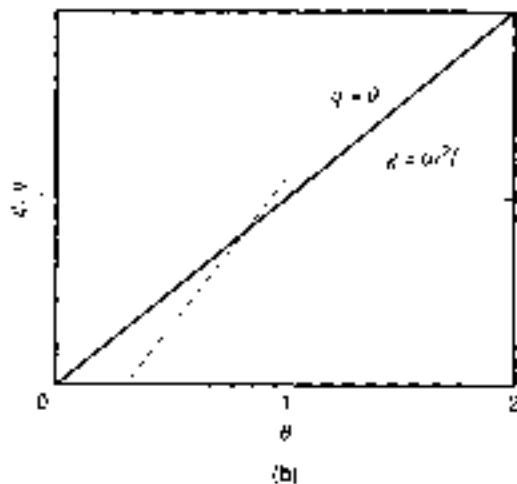
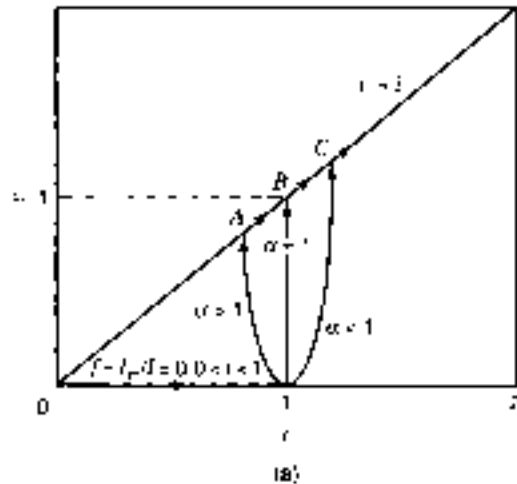
$$= 1 - (1 - \theta) \frac{1}{i} = \frac{i + \theta}{i}$$

and consequently,

$$f = \begin{cases} 1 & \text{(resistive)} \\ \frac{i + \theta}{i} & \text{(current-sharing)} \\ 0 & \text{(superconducting)} \end{cases} \tag{9}$$

From Eqs. (2) and (3), after rearrangement and introduction of the reduced voltage drop, current, and temperature, as well as the Stekly parameter, one can obtain

$$v = \frac{\alpha}{i} \tag{10}$$



**Figure 5.** Reduced characteristics. (a) Reduced voltage versus reduced current. If the conductor is fully superconducting (i.e.,  $f = I_m/I = 0$ ), then  $0 < i < 1$  (line 0-1). Current sharing takes ( $f = \theta = 1 - 1/i$ ), for linear  $I$  versus  $T$  correspond to the range of reduced current  $1 < i < 1 + \theta$ . Full crystallinity corresponds to the 0-1-C curve ( $\alpha > 1$ ) and sufficient cooling. The 0-1-A curve ( $\alpha < 1$ ) corresponds to insufficient cooling. For  $\alpha > 1$  the reduced voltage  $v$  is a single valued function of  $i$ . With an increase in  $i$  from 0 to 1 the superconductor quenches at  $i = 1$ . With decreasing current while in normal mode along the  $v = 1$  line the conductor switches back to the superconducting mode at  $i = 1/\alpha^2$ . (b) Reduced thermal energy generation  $\mu$  and reduced cooling rate  $\varphi$  versus reduced temperature  $\theta$  for a fixed reduced current. (for  $i = 1, \alpha^2 < \theta < 1$ ). Stable conditions for  $\alpha > 1$ .

By eliminating the reduced temperature from Eqs. (7) and (10) one can obtain the reduced voltage as a function of the reduced current in the matrix. This situation describes the variation of voltage with current. An interesting insight can be gained from the  $v$ - $i$  diagram using the relations given by Eqs. (7) and (10) and recognizing that the  $v = 1$  line in that diagram depicts the normal state (Fig. 5(a)). The voltage-current relationships are described in the legend of Fig. 5(a).

If the current is assumed to be fixed, an insight into a conductor modes of operation can be gained by using a diagram of the reduced thermal energy generation  $\mu = Q_A / (h(T_c - T_b))$ , and/or the reduced heat transfer rate Ref.  $\varphi = Q_A / (h(T_c - T_b))$ , versus reduced temperature (Fig. 5(b)), according to Gauster as reported in Ref. 27.

Finally, it should be pointed out that several definitions of the stability factor exist in the literature. In addition to the Stekly parameter  $\alpha$ , a stabilization parameter  $\xi = I_{oc}$  can be used (28):

$$\xi = \frac{\phi_2 I_{oc} A_m h_c T_c}{I_m J_c' h_c} \quad (11)$$

where also a parameter  $\phi_2$  has been introduced, equal to the fraction of matrix perimeter actually exposed to coolant. If the stabilization parameter as defined by Eq. (11) is to be used, note that  $\xi > 1$  implies a stable system and  $\xi < 1$  an unstable one. According to Ref. 17 the stabilization parameter  $\xi$  can be defined either at the critical temperature and the operating current, or at the critical temperature and zero current and magnetic field. In addition, a separate parameter can be defined for the current sharing case.

Prediction of the stability using the Stekly cryostability parameter has opened a new era in building large superconducting magnets. From Eq. (8) it is obvious that, for a given superconductor and matrix material (usually copper) as well as for the defined cooling conditions (to achieve greater stability, one should increase the ratio of matrix to superconductor in the cross section of the composite conductor. This criterion has led to very conservative design ( $\alpha$  can be smaller than 0.1, i.e.,  $\xi$  larger than 10), leading to very expensive solutions.

Let us consider an application of the Stekly criterion in an actual conductor design (for more details consult Ref. 27). For a typical critical current density of the order of magnitude of  $10^5$  A/cm<sup>2</sup> and a standard cooling heat transfer rate of  $2 \times 10^4$  W/cm<sup>2</sup> (a composite conductor with a low-temperature superconductor cooled by liquid helium), a change of the critical current between  $10^2$  A and  $10^3$  A leads to conservative conductor designs as follows. The ranges of the conductor diameters and the copper-to-superconductor ratios are between 1 mm and 21 mm and between 7 and 35, respectively. The design constraint assumes  $\alpha > 1$ . These rather large matrix-to-superconductor ratios can be reduced by further reducing the operating current (the standard operation is at  $I = 0.75 = 0.2I_c$ ). A more elaborate theory, based on the same approach, takes into account thermal resistance between the matrix and the superconductor and their actual sizes. Still, the stability criterion stays conservative,  $\alpha \geq 1$ . It should be noted that the transition from nucleate to film boiling of the coolant (liquid helium) may cause instability when  $\alpha > 1$  if the current sharing is not complete.

**Maddock-James-Norris Equal-Area Theorem.** The Stekly cryostability criterion has been developed taking into account removal of the generated thermal energy from a conductor only by convection, that is, the conduction mechanism has been neglected. This is a consequence of the assumption of the existence of a uniform temperature throughout the conductor. This assumption, however, can easily be violated, because the matrix material (usually copper) is thermally conductive. Hence, a question may be posed: How must the steady-state stability criterion be formulated if the conduction effect is to be taken into account? An elegant approach to the solution of this problem was given by Maddock et al. (29). We will summarize a one-dimensional formulation as introduced by Maddock (29), although a two-dimensional (or for that matter a three-dimensional) formulation can be readily devised.

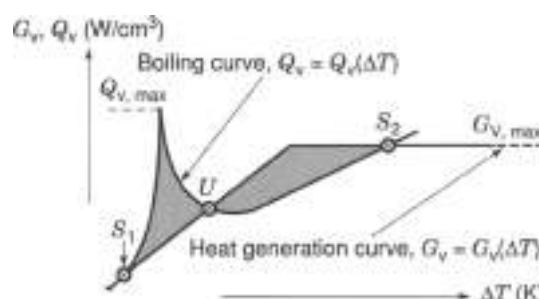


Figure 6. An equal-area theorem: boiling characteristic curve ( $Q_v$ ) and thermal energy generation curve ( $G_v$ ). The area between the  $Q_v$  and  $G_v$  curves on the segment between the points  $S_1$  and  $U$  represents an excess of heat transfer rate removed by boiling at the cold end. The area between the same curves on the same  $U$ - $S_2$  represents an excess of thermal energy generated at the hot end. Equality of the two energy rates (equality of the corresponding areas) corresponds to a balanced condition  $Q_v = G_v$  over the entire conductor.

Let us assume that a composite conductor carrying constant current  $I$  and with one-dimensional geometry (i.e., a long cylindrical rod with  $D/L \ll 1$ ,  $D$  being its outer diameter and  $L$  its length) is submerged in a liquid helium pool. Let a disturbance, caused by an instability, initiate the appearance of a normal zone in the superconducting part of the conductor, thus leading to current sharing. Thermal energy generation caused by Joule heating in the normal current-carrying parts of the conductor will cause an increase of the conductor temperature. Eventually, when all the current starts flowing through the matrix, any further increase in temperature will be accompanied by a constant thermal energy generation  $G$  (the current intensity is constant). Due to the conductor surface temperature increase, the boiling heat transfer mode of liquid helium may change from nucleate (at an initially small temperature difference) to film boiling (at more pronounced temperature differences). The boiling characteristic (i.e., the curve of heat transfer rate density versus temperature difference) will determine the rate of heat removal. In Fig. 6, both thermal energy generation and heat transfer rate are presented as functions of temperature difference between the conductor surface and the liquid helium pool (note that  $T_c$  remains unchanged; consequently the axis in Fig. 6 may be considered as the conductor surface temperature with  $T_c$  as a fixed parameter).

The area enclosed between the thermal energy generation curve and the boiling heat transfer curve over the segment  $S_1$ - $U$  is proportional to the excess thermal energy removed from the conductor. Similarly, the area enclosed over the segment  $U$ - $S_2$  between the thermal energy generation curve and the corresponding film boiling segment of the boiling characteristic curve is proportional to the excess of thermal energy generated within the conductor over the heat transfer rate density removed by boiling. Because of the temperature difference between the hot and cold ends, heat conduction will take place along the conductor leading to a removal of the excess thermal energy from the hot end by the excess of boiling heat transfer from the cold end. A stable equilibrium of this heat transfer process will be reached if the two excess thermal energy rates (represented by the shaded areas between the curves in Fig. 6) are exactly the same. This straightforward phenomenological description leads directly to the required stability criterion:

Equality of thermal energy rates, represented by equal areas in Fig. 6, corresponds to a limiting case of a cryogenic stability. If the cold-end thermal energy rate is larger than the hot-end thermal energy rate, the conductor temperature will ultimately return to its initial temperature, equal to the liquid bath temperature, thus leading to the disappearance of the normal zone and a restoration of the superconducting mode of the conductor. The limiting condition can be interpreted as an equality of the areas in Fig. 6 with, say, conductor temperature as abscissa. This conclusion constitutes the so-called equal area theorem introduced by Maddock et al. (29).

Formal mathematical proof of the above described theorem is straightforward. The general energy balance equation (with the uniform temperature assumption relaxed) reads as follows:

$$\begin{aligned} & \text{Input heat transfer rate by conduction} \\ & - \text{heat transfer rate by convection} \\ & - \text{rate of thermal energy generation} \end{aligned}$$

In an analytical form for a one-dimensional heat transfer problem this balance reads as follows:

$$\frac{d}{dx} [k(T) \frac{dT}{dx}] = \frac{q_0 l h}{A} (T - T_c) - \frac{\rho_0 I^2 l^2}{AA_m} \left( \frac{T - T_c}{T_{c0} - T_b} \right) \quad (12)$$

where  $A$  ( $m^2$ ) is the area of the conductor cross section, and  $k$  in  $(W/m K)$  is the thermal conductivity of the matrix. The bracketed term on the left-hand side and the two terms on the right hand side are all functions of temperature; thus, in a compact form,

$$\frac{dF(T)}{dx} = Q_c(T) - G_e(T) \quad (13)$$

where  $F$ ,  $Q_c$ , and  $G_e$  are the conductive heat flux, heat transfer rate per unit volume of the conductor, and thermal energy generation per unit volume of the conductor, all exclusive functions of temperature. After the formal integration of Eq. (13) one obtains

$$\int_{T_1}^{T_2} k(T) [Q_c(T) - G_e(T)] dT = \int_{T_1}^{T_2} F dT = 0 \quad (14)$$

where we assume that the integration boundaries ( $T_1$  and  $T_2$ ) are well apart so as to reach the zones in the conductor where the temperature gradients are equal to zero. Making the reasonable assumption that thermal conductivity  $k(T)$  depends linearly on temperature in the given temperature range ( $k(T) = cT$ ), a transformation of the dependent variable in Eq. (14),  $t = cT/2$ , will lead to

$$\int_{t_1}^{t_2} [Q_c(t) - G_e(t)] dt = 0 \quad (15)$$

A geometric interpretation of the Eq. (15) is straightforward. If one replaces the abscissa in Fig. 6 using the same transformations that led to Eq. (15), then Eq. (15) represents the equality of the areas presented in Fig. 6. If this condition is satisfied (i.e., the energy balance is preserved), any excess of thermal energy generated within the hot zone of a superconductor will be compensated by the excess of the heat transfer rate removed

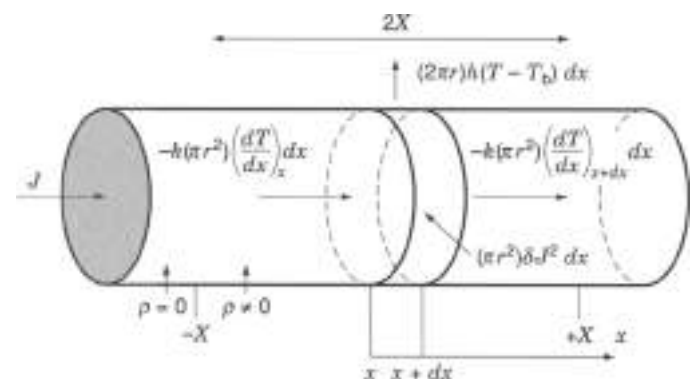
from the conductor in the cold zone thanks to the conduction between the two parts of the conductor.

**Wipf Minimum Propagation Zone.** The concept of the minimum propagation zone (MPZ) was introduced by Martucci and Wipf (30). It has since been developed into a comprehensive stabilization theory (17). The original approach was based on a simple energy balance of a superconducting material idealized to be infinite in space, but with a localized thermal energy generation source within a preexisting normal zone of finite size. This balance, in a generalized form, is as follows:

$$\begin{aligned} & \text{Net heat transfer rate due to heat conduction} \\ & - \text{rate of thermal energy generation by Joule heating} \end{aligned}$$

Hence, the approach assumed a balance between thermal energy generated by Joule heating (within the already established normal zone) and the rate of heat transfer from the normal zone by conduction through the superconductor. Therefore, the physical size of the normal zone that evolves into a quench has been assumed to depend on a tradeoff between Joule heating in the normal zone and the heat carried out of the normal zone by conduction. As a consequence, assuming the validity of the energy balance condition described above, the normal zone will neither grow nor collapse. The normal zone defined in such a way is called the MPZ (17).

The simplest *realistic* situation is the one that leads to a conservative stabilization limit, which requires the consideration of non-current-sharing conditions, and additional convective cooling. Let us assume that a linear superconductor, with an already developed normal zone of length  $2X$ , is cooled by (1) a coolant at temperature  $T_c$ , and (2) conduction in the axial directions (Fig. 7). In such a situation, the energy balance has an additional term [see Eq. (12)]. The conductor has a circular cross section (radius  $r$ ), with a constant electrical resistivity  $\rho$  (in the normal zone) and uniform thermal conductivity  $k$  (in both normal and superconducting parts). Under these conditions, the differential energy balance per unit length of the



**Figure 7.** A conductor with a formed normal zone ( $\rho \neq 0$ ) of total length  $2X$ . Outside the normal zone, the conductor is superconducting ( $\rho = 0$ ). The conductor is exposed to convective cooling, and to thermal energy generation within the normal zone. The heat conduction through the conductor removes heat from the normal zone.

conductor is as follows (17):

$$i\pi r^2 j_0 \frac{d^2 T}{dx^2} - (2\pi r) h_c (T - T_b) - \pi r^2 j_0 j_c^2 \quad (16)$$

with the following boundary conditions:

1. Superconducting zone (note that in this zone the second term on the right-hand side in Eq. (16) vanishes):

$$\begin{aligned} \text{at } x = -L, \quad T &= T_b \\ \text{at } x = +L, \quad T &= T_b \end{aligned} \quad (17)$$

2. Normal zone (with extra heating  $q_w$  at  $x = 0$ ):

$$\begin{aligned} \text{at } x = 0, \quad \frac{dT}{dx} &= -k \frac{dT}{dx} \\ \text{at } x = \pm X, \quad T &= T_b \end{aligned} \quad (18)$$

Note also that at the boundaries between the localized normal zone and both superconducting zones, the temperature gradients must be equal. Hence, Eqs. (16)–(18) (and also an additional equation representing the equality of the temperature gradients at the junctions between the zones) define the mathematical model of the temperature distribution in the conductor. The closed-form solution can be readily obtained (17) of particular interest is an analytical relationship between the heat flux caused by heating,  $q_w$ , [defined by Eq. (16)] and the length of the normal zone. This relationship can be obtained from the first of the two equations in Eq. (18) by differentiating the temperature distribution inside the normal zone, and by subsequently applying the result at the normal zone boundary. The final result is as follows:

$$\begin{aligned} \psi &= \frac{q_w}{i\pi r k / 2 h_c l^2 j_c^2} \\ &= 2 \left[ \xi \exp\left(\frac{X}{i\pi r k / 2 h_c l^2}\right) - \sinh\left(\frac{X}{i\pi r k / 2 h_c l^2}\right) \right] \end{aligned} \quad (19)$$

where  $\xi$  is a stability parameter defined in the same manner as the one introduced by Eq. (11). The relationship given by Eq. (19) is presented graphically in Fig. 5. The abscissa in Fig. 5 represents the dimensionless length of the normal zone,  $l = 2X / (i\pi r k / 2 h_c l^2)$ . Note that the length of the MPZ is equal to twice the normal-zone half length  $X$  for the critical magnitude defined by Eq. (19), that is, for  $q_w = 0$ :

$$\text{MPZ} = -i\pi r k / 2 h_c l^2 \ln(1 - 2\xi) \quad (20)$$

In Fig. 5, three distinct regions can be identified. The first region corresponds to the pairs of values of the dimensionless heating  $\psi$  and the dimensionless length of the normal zone,  $l$ , such that  $\xi < 0.5$ . The second region corresponds to the stability-parameter range  $0.5 < \xi < 1$ , and the third region to  $\xi > 1$ . In the first region, say for  $\xi = 0.49$ , the superconductor is in a least stable situation (15). In this region a MPZ exists for any  $0 < \xi < 0.5$ . At  $\xi = 0.5$  the MPZ is theoretically infinitely large; see Eq. (20). The second region ( $0.5 < \xi < 1$ ) is characterized by increasing stability. It should be noted that any  $\xi = \text{const}$  curve in this region has a minimum

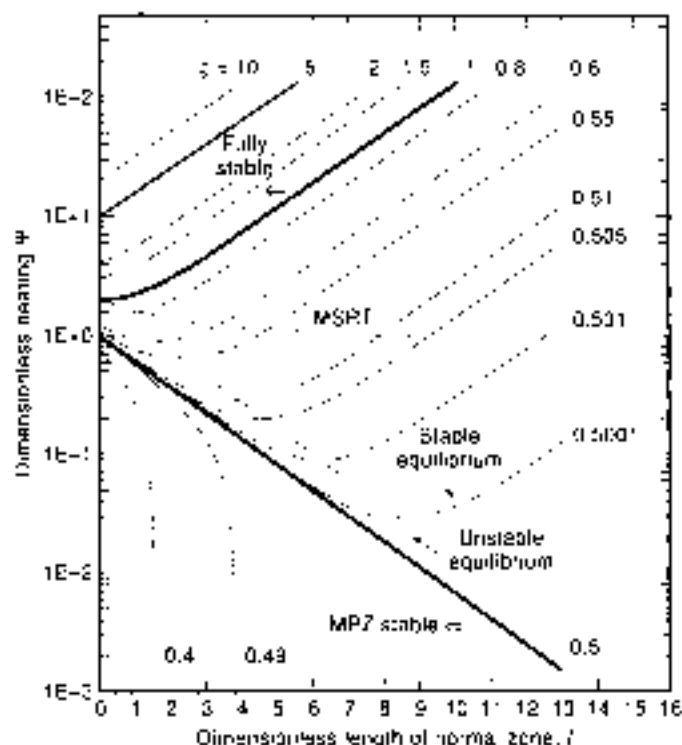
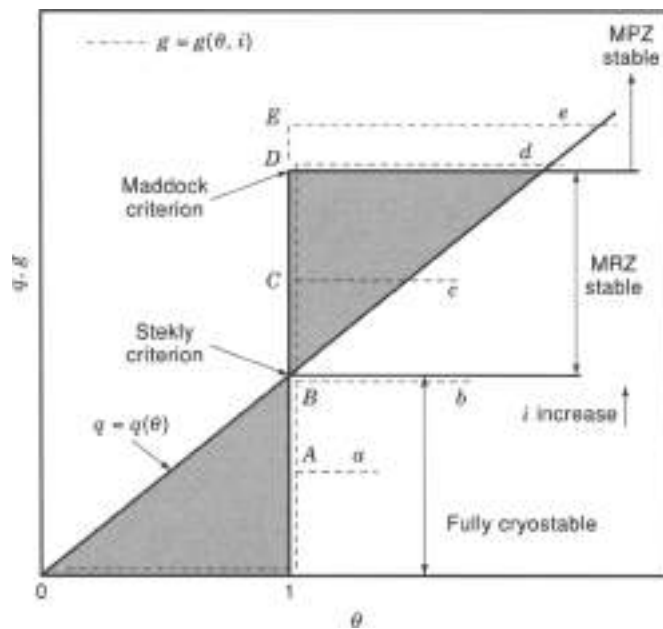


Figure 8. Dimensionless extra heating versus dimensionless length of the normal zone. No current sharing case. The region for  $\xi > 1$  corresponds to a fully stable conductor (crystalline region). The region for  $1 > \xi > 0.5$  represents the so-called minimum recovery zone (MRZ) of stability. The length of the MRZ is the length of a superconducting zone in equilibrium with neighboring normal zones of infinite length. A longer superconducting zone will spread into the normal zones; a shorter one will shrink. The region for  $\xi \leq 0.5$  corresponds to the so-called minimum propagation zone (MPZ) of stability. The MSRT line is the line of maximum size of recoverable transition. The locus of minima in the MRZ region provides a boundary between stable and unstable equilibrium (17).

Hence, the left-hand branches of these lines correspond to an unstable equilibrium (between the  $l = 0$  value and  $l$  or  $\psi_{\text{crit}}$ , for any  $\xi$ ). The locus of these minima defines the bifurcation line between the unstable and stable equilibrium. The *critical size of recoverable transition* (MSRT) line is defined by  $X / (i\pi r k / 2 h_c l^2) = \ln(2\xi - 1)$ . Finally, the third region ( $\xi > 1$ ) corresponds to fully stable (i.e., crystalline) conditions. In that region any created normal zone is in stable equilibrium. The size of the normal zone increases monotonically with increase of the heat release per unit conductor volume. Note that  $\xi \geq 1$  corresponds to the condition imposed by the conservative Stekly criterion.

The Wipf approach has been refined even further for periodic temperature distributions and for a current sharing situation (17,27,31).

It is instructive to provide a simplified representation of the main features of all three stabilization criteria discussed so far (Stekly, Maddock, and Wipf). A diagram of reduced cooling rate versus reduced temperature may be used (10); see Fig. 9. There the reduced cooling rate is assumed to be a linear function of reduced temperature. The non-current-sharing case will be considered. The following situations may be distinguished. For the reduced heat generation, the one described by curve a, the



**Figure 9.** Reduced cooling rate,  $q = Q_{cl}/(h(T_{cl} - T_b))$ , and reduced thermal generation,  $g = G_{cl}/(h(T_{cl} - T_b))$ , versus reduced temperature,  $\theta = (T_1 - T_b)/(T_{cl} - T_b)$ . Curves 0-1 A-a, 0-1-B-b, 0-1-C-c, 0-1-D-d, and 0-1-E-e, each denoted by a dotted line, correspond to various values of the reduced thermal generation  $g$ , each with different but fixed current. The reduced cooling curve is the same for all these cases. In a fully cryostable region (curve 0-1-A-a) the corresponding current causes a Joule heating lower than the cooling rate for any reached temperature. The stability conditions for the curve 0-1-B-b satisfy the Stekly criterion; see the curve for  $i = 1.0$  in Fig. 8. The stability conditions for the curve 0-1-D-d satisfy the Maddock criterion (the shaded areas both below and above the cooling curve are equal in each other); see the curve for  $i = 0.5$  in Fig. 8. This representation is valid in the absence of current sharing (compare with Fig. 9).

superconductor is in a fully stable mode (cryostable case). The limit of this type of behavior is at the condition described by the Stekly criterion (the reduced heat generation denoted by curve b). If current further increases (represented by an increase in reduced current  $i$  in Fig. 9), the superconductor reaches the region of MRZ stability, as described by Wipf (curve c). Further increase in current may eventually provide the condition that corresponds to the Maddock equal-area stability criterion (curve d). Note that for curve d, the areas between the curves representing the reduced cooling rate and reduced heat generation are equal. If current increases even more, the MPZ stability zone will be reached (curve e).

#### ADVANCED STABILITY AND THERMAL MANAGEMENT PROBLEMS FOR SUPERCONDUCTORS

The conventional theory of cryogenic stability of both low-temperature and high-temperature superconductors has not been able to address a number of issues related to the design of modern superconducting devices. Without trying to provide a comprehensive review (for the details consult the bibliography contained in Ref. 27), let us discuss briefly two characteristic topics.

As commented on in the introductory part of this article, instability phenomena are inherently transient. Consequently, the quench and recovery are also inherently transient. The existence of a normal zone, as introduced first by Wipf and others, should be considered in the light of the propagation velocity. This velocity is positive during quench and negative during recovery. Under the equilibrium conditions it is equal to zero. A good review of the early attempts to determine this velocity is provided in Ref. 27. A modern treatment of the related problems (such as the existence of the so-called traveling normal zones (TNZs), the propagation velocities in the uncooled superconductors, the influence of nonlinearities introduced by temperature-dependent material properties, and the issues related to thermal and hydrodynamic management of internally cooled superconductors) is reviewed in Ref. 31. It should be added that the stability conditions defined by the conventional theory based on the presence of a continuous disturbance does not reflect the proper conditions for stability against transient heat pulses. So transient stabilization must be considered (27). An important part of these studies is an appreciation of transient heat transfer phenomena. For example (31), for very short transient phenomena (of the order of magnitude of tenths of a millisecond), the heat transfer coefficient may be in the range of  $0.5 \text{ W/cm}^2\text{K}$  to  $1.5 \text{ W/cm}^2\text{K}$  (the propagation velocity was between  $5 \text{ m/s}$  and  $20 \text{ m/s}$ ). These transient heat transfer coefficients are larger by an order of magnitude than the film boiling heat transfer coefficient.

The introduction of HTSC materials brought an additional appreciation of the need to include in the analysis the temperature-dependent thermophysical properties and kinetic heat transfer characteristics over a much broader temperature range. A direct consequence of this fact in any thermal management problem for a superconductor (not necessarily related to cryostability) is a requirement to model the problem with its increasingly nonlinear character. Under these circumstances, the originally introduced assumptions must be revisited (32).

#### THERMOPHYSICAL AND HEAT DATA ON CRYOGENS

Although the HTSCs can exhibit some superconducting properties even at temperatures approaching ambient, superconductivity applications under practical conditions still depend upon cooling to cryogenic temperatures. This fact does not diminish the value of the HTSCs, because of the greater ease of operation and savings in refrigeration power that are effected by operating at temperatures nearer to that of liquid nitrogen than to that of liquid helium. The theoretical power input required to produce one unit of refrigeration at cryogenic temperatures is given as  $(T_1 - T_2)/T_2$ , where  $T_1$  is ambient temperature, or the temperature at which the heat must be rejected, and  $T_2$  is the refrigeration temperature, the temperature at which the heat is to be removed. For an ideal cryogenic refrigerator to remove heat at the temperature of liquid helium (4 K) and reject it at ambient temperature (300 K) would require 74 W of input power per watt of refrigeration. At liquid nitrogen temperature (77 K) this ratio is reduced to a little less than 3 W power input per watt of refrigeration. As Strachan (33) has shown, cryogenic refrigerators to date do not approach this theoretical limit very closely. The degree to which this limit is approached is not dependent upon the temperature level of refrigeration.

Table 2. Pertinent Thermophysical Properties of Cryogen<sup>a</sup>

| Cryogen (Liquid) | NBP <sup>b</sup> Temperature (K) | Triple Point Temperature (K) | Critical Temperature (K) | NBP Density (kg/m <sup>3</sup> ) | Critical Pressure (MPa) | NBP <sup>b</sup> Heat of Vap. (kJ/kg) |
|------------------|----------------------------------|------------------------------|--------------------------|----------------------------------|-------------------------|---------------------------------------|
| He               | 4.215                            | —                            | 5.25                     | 125                              | 0.23                    | 20.5                                  |
| H <sub>2</sub>   | 20.27                            | 13.80                        | 32.98                    | 71                               | 1.20                    | 415.1                                 |
| Ne               | 27.09                            | 24.8                         | 44.4                     | 1236                             | 2.37                    | 86.2                                  |
| N <sub>2</sub>   | 77.1                             | 63.1                         | 126.2                    | 809                              | 3.39                    | 199                                   |
| Ar               | 87.3                             | 83.8                         | 150.9                    | 1394                             | 4.86                    | 163                                   |
| O <sub>2</sub>   | 90.2                             | 54.4                         | 154.6                    | 1140                             | 5.11                    | 213                                   |

<sup>a</sup> Data from Ref. 44.

<sup>b</sup> Normal boiling point.

Helium does not have a true triple point and can only be solidified at a minimum pressure of 2.5 MPa.

<sup>c</sup> Properties given are those of parahydrogen.

but depends strongly upon the capacity of the unit. The larger the refrigerator, the better is the degree of approach, with the best (largest) units reaching only 35% to 40% of the limit (88).

Table 2 lists some of the thermophysical properties of cryogen (34) that might be used in the cooling of a superconducting system.

If cooling is to be done by pool boiling of a liquid, the range of temperature that is available is of interest. For a given cryogen, the boiling temperature can be fixed at any temperature between the triple point and the critical point by maintaining the corresponding system pressure. From Table 2 it can be seen that over the range from below 4.15 K to over 150 K, there are only two gaps; from 5.25 K to 13.8 K, and from 44.4 K to 54.4 K.

Also, the application of supercritical pressure can allow forced-flow cooling without encountering two-phase flow with its consequent complications of pressure and flow oscillations and excess pressure drop in the flow channels. However, only in the case of liquid helium can this be accomplished at a pressure as low as 0.23 MPa.

The main heat transfer phenomena involved with cooling of superconducting magnets are as follows: (1) forced convection of a single-phase coolant (say, gaseous or supercritical helium), (2) phase-change heat transfer (nucleate and film boiling) in both pool boiling and channel flow conditions, and (3) heat transfer in superfluid helium. The most important mode of heat transfer for cryogenic stabilization is boiling (both nucleate and film boiling). In Table 3 typical data for (1) the critical heat flux (i.e., the maximum heat flux for nucleate boiling at the given temperature difference) and (2) the Leidenfrost-point heat flux (i.e., the minimum heat flux for film boiling) are given. It should be noted that the available heat transfer data scatter widely and depend greatly on heat transfer surface orientation, geometry, and surface conditions. Transient heat flux data for nucleate

boiling of cryogen are greater by an order of magnitude than steady-state values.

## SAFETY

In working with any cryogenic fluid, safe operation requires that there be a satisfactory understanding of the hazards that can arise and also a strict compliance with safe operating principles for these fluids. These hazards can stem from the temperature of the fluid, or from its reaction with other materials that come in contact with it. The low temperatures can embrittle some structural materials, or can cause freezing of human tissue if personnel should contact the cold fluid or the exterior of a cold, uninsulated pipe. The low temperatures also cause a significant thermal contraction which, if not sufficiently compensated, can give rise to high stresses that can cause their own hazard.

Any of the cryogen that have boiling points below that of liquid oxygen can condense the atmospheric air, resulting in a condensate that is enriched in oxygen (as much as 50% oxygen) which is even more hazardous than liquid air. If this condensed air is allowed to fall on sensitive equipment, some materials can be sufficiently embrittled to crack, other equipment may no longer function properly, and, if the enriched liquid air should fall on a combustible material (such as asphalt), an explosion can result. In an improperly purged system, condensation of air, or its constituents such as water vapor or carbon dioxide, can result in the obstruction of pressure relief passages or the connections to instrumentation needed for the safe operation of the system.

The use of liquid hydrogen or liquid oxygen is less probable in the application of superconductivity. However, such an application is not impossible. Use of these fluids entails the additional hazard of combustion, or even explosion, if strict safety measures are not observed.

The most likely safety hazard to be encountered in applications of superconductivity is the excessive buildup of pressure within a cryogenic system. If a cryogenic fluid is totally confined, as in a pipe between two closed valves, the pipe must try to maintain the density of the liquid as the contained fluid becomes a gas that approaches ambient temperature. Because of the increase in the compressibility factor of the gas at high pressure, the pressure obtained is higher than what would be computed by applying the ideal gas law to the appropriate density ratios. Consequently, helium trapped in this fashion could reach a pressure of 103 MPa (15,000 psi), and

Table 3. Typical Boiling Heat Flux Data

| Cryogen (Liquid) | Critical Heat Flux <sup>a</sup> |                | Leidenfrost Point Heat Flux (W/cm <sup>2</sup> ) |
|------------------|---------------------------------|----------------|--|
|                  | Value (W/cm <sup>2</sup> )      | $\Delta T$ (K) |  |
| Helium           | 1                               | 1              | 0.5  |
| Hydrogen         | 10                              | 5              | 0.5  |
| Neon             | 15                              | 5              | 1  |
| Nitrogen         | 15                              | 13             | 1  |
| Oxygen           | 25                              | 33             | 2  |

<sup>a</sup> At atmospheric pressure. These numbers are only an approximate representation of a range of typical results. (2)

nitrogen could reach a pressure of 296 MPa (43,000 psi) if the container did not rupture beforehand (35). In the case of a quench of a superconducting magnet, there is a very rapid release of energy that will quickly enter the cryogenic coolant. If sufficient venting capacity is not provided, a rapid, hazardous buildup of pressure in the coolant passages or the container can result.

#### ADDITIONAL READING

Several references mentioned in the bibliography deserve additional attention as useful sources for further reading. The book of Wilson (14) is a classical text and must be read as an upper level introduction to the problems of design of superconducting magnets. A two-volume book by Collings (27) provides extensive insights into many highly technical aspects of metallurgy and physics of low-temperature superconductors, as well as a very comprehensive bibliography. The book by Dresner (31) is an advanced text that provides an insight into the modern treatment of the stability of superconducting devices. For technical calculations related to design of cryogenic devices, including cryostability aspects, the book by Iwasa (28) will be very useful. A number of useful sources can be found that deal with cryogenic and heat transfer aspects of the design of cryogenic devices. In addition to those mentioned in the bibliography, the following one may be consulted: S. W. van Sciver, *Helium Cryogenics*, New York: Plenum, 1986.

#### BIBLIOGRAPHY

1. H. K. Onnes, *Commun. Phys. Lab. Univ. Leiden*, Nos. 120k, 122b, 124c, 1911; see also K. Gavrilin, and Y. Goubarovs (eds.), *Heike Kamerlingh Onnes: Through Measurement to Knowledge, The Selected Papers*, Norwell, MA: Kluwer, 1990.
2. K. D. Timmerhaus and T. M. Flynn, *Cryogenic Process Engineering*, New York: Plenum, 1989.
3. R. V. Carlson, Applications of superconductivity, in: K. D. Williamson, Jr., and E. J. Edeskuty (eds.), *Liquid Cryogenics*, vol. II, Boca Raton, FL: CRC Press, 1983, chap. 4.
4. D. F. Larbalestier and M. P. Maley, Conductors from superconductors: Conventional low-temperature and new high-temperature superconducting conductors, *Mater. Res. Soc. Bull.*, **18**: 50-56 (1993).
5. K. Mendelssohn, Superconductivity and superconducting devices, in C. A. Bailey (ed.), *Advances in Cryogenics*, London: Plenum, 1971, chap. 10.
6. D. S. Beard et al., The IFA large coil task, *Fusion Eng. Des.*, **7**: 1-232, 1988.
7. V. A. Alflov et al., *Stabilization of Superconducting Magnetic Systems*, New York: Plenum, 1977.
8. E. J. Edeskuty, Cryogenics, in E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers*, 23th ed., New York: McGraw-Hill, 1990, chap. 19.2.
9. M. P. Maley, personal communication, 1998.
10. X. D. Wu et al., Properties of  $YBaCu_3O_7$  thick films on flexible buffered metallic substrates, *Appl. Phys. Lett.*, **67**: 2387-2389, 1995.
11. P. Chawdhari and E. J. Edeskuty, Bulk power transmission by superconducting dc cable, *Electr. Pow. Syst. Res.*, **1**: 41-49, 1977, 76.
12. E. B. Forsyth, The high voltage design of superconducting power transmission systems, *IEEE Trans. Insulation Mag.*, **6**: 7-16, 1990.

13. R. J. Thorne, Magnet program overview for the international thermonuclear test reactor, *IEEE Trans. Magn.*, **30**: 1595-1601, 1994.
14. M. N. Wilson, *Superconducting Magnets*, Oxford: Clarendon Press, 1987.
15. H. Wada et al., VD-1, Appendix—terminology for superconducting materials, *Cryogenics, VAMAS Suppl.*, **35**: S113-S126, 1995.
16. A. A. Abrikosov, On the magnetic properties of superconductors of the second group, *Sov. Phys. JETP*, **5**: 1174-1182, 1957, transl. of *Zh. Exp. Teor. Fiz.*, **32**: 1442-1452, 1957.
17. S. I. Wipl, *Stability and degradation of superconducting current-carrying devices*, Tech. Rep. LA-7275, Los Alamos Scientific Laboratory, Dec. 1978.
18. M. N. Wilson, Stabilization of superconductors for use in magnets, *IEEE Trans. Magn.*, **MAC-13**: 440-446, 1977.
19. O. Tsukamoto, T. Takao, and S. Homyo, Stability analysis of superconducting magnet: An approach to quantification of energy disturbances raised conductor motion, *Cryogenics*, **20**: 616-620, 1989.
20. R. G. Mints and A. L. Rukhmanov, *Instabilities in Superconductors* (in Russian), Moscow: Nauka, 1984.
21. Y. Iwasa et al., Experimental and theoretical investigation of mechanical disturbances in epoxy-impregnated superconducting coils, part 1: General introduction, *Cryogenics*, **25**: 304-306; part 2: Shear stress-induced epoxy fracture as the principal source of premature quenches and training—theoretical analysis, 307-316; part 3: Fracture-induced premature quenches, 317-322; part 4: Pre-quench cracks and fractional motion, 323-326, 1985.
22. V. I. Dubenko et al., Experimental study of thermomechanical instabilities in superconducting composites, part 1: Magnetic flux jumps induced by plastic deformation, *Cryogenics*, **29**: 10-15; part 2: Dynamics of normal zones, 16-21, 1989.
23. H. T. Beal et al., Definitions of terms for practical superconductors: 3. Fabrication, stabilization and transient losses, *Cryogenics*, **19**: 327-332, 1979.
24. A. R. Kantorowicz and Z. J. J. Steky, A new principle for the construction of stabilized superconducting coils, *Appl. Phys. Lett.*, **8**: 561-57, 1965.
25. Z. J. J. Steky and J. L. Znr, Stable superconducting coils, *IEEE Trans. Nucl. Sci.*, **12**: 367-372, 1965.
26. Z. J. J. Steky, B. Thome, and B. Strauss, Principles of stability in solid superconducting magnets, *J. Appl. Phys.*, **40**: 2238-2245, 1969.
27. E. W. Collings, *Applied Superconductivity: Metallurgy and Physics of Titanium Alloys*, New York: Plenum, 1986.
28. Y. Iwasa, *Case Studies in Superconducting Magnets: Design and Operational Issues*, New York: Plenum, 1991.
29. D. J. Madlock, G. H. Jones, and W. T. Norris, Superconductive composites: Heat transfer and steady state stabilization, *Cryogenics*, **9**: 261-275, 1969.
30. A. E. Martineb and S. I. Wipl, Investigation of cryogenic stability and solubility of operation of Nb<sub>3</sub>Sn coils in helium gas environment, in H. M. Long and W. F. Giauque (eds.), *Proc. 1972 Appl. Superconductivity Conf.*, IEEE Publ. 72CH0682 8 TABSC, 1972, pp. 341-349.
31. L. Dresner, *Stability of Superconductors*, New York: Plenum, 1995.
32. D. P. Sekulic, E. J. Edeskuty, and Z. G. Yao, Heat transfer through a high temperature superconducting current lead at cryogenic temperatures, *Int. J. Heat Mass Transfer*, **40** (16): 3917-3926, 1997.
33. T. R. Searbridge, *Cryogenic Refrigeration—An Updated Survey*, NBS Technical Note 135, Boulder, CO: US Department of Commerce, National Bureau of Standards, 1974.



34. R. L. Mills and F. J. Edeskuty, *Cryogenics and their properties*, in K. D. Williamson, Jr., and F. J. Edeskuty (eds.), *Liquid Cryogenics*, vol. D, Boca Raton, FL: CRC Press, 1983, chap. 1.
35. F. J. Edeskuty and W. P. Stewart, *Safety in the Handling of Cryogenic Fluids*, New York: Plenum, 1986.

D. F. ŠEKULIĆ  
University of Kentucky  
F. J. EDESKUTY  
Edeskuty Engineering

## FORCED FLOW CONDUCTOR MANUFACTURING

Electric windings that use low temperature superconductors need to be actively cooled. An easy and effective way to maintain the winding at operating temperature is to fill the space around the coil with liquid helium (pool-cooled magnets), allowing the liquid to penetrate between the turns. Alternatively, the superconductor may include a cooling path where the coolant is forced to circulate under pressure. This article is a review of the manufacturing methods for forced flow conductors, starting from the early prototypes (80 years ago) to the very large cable-in-conduit conductors of the fusion magnets. The design aspects of forced flow conductors are not deeply discussed here, although they have sometimes driven the effort toward advanced manufacturing technologies. Essential references on the design aspects of forced flow conductors can be found in the articles: STABILITY IN FORCED FLOW SUPERCONDUCTIVITY, HYSTeresis AND COUPLING LOSSES, SUPERCONDUCTING MAGNETS, QUENCH PROTECTION.

The manufacturing methods for forced flow superconductors sound like a brand new subject. Actually, it is old enough to have a history. The initial proposal dates back to 1965, when Kolmitz suggested using the circulation of supercritical helium for reliable cooldown and heat transfer in large superconducting magnets. The first milestone is the spark chamber detector built in 1970 at CERN (2), with 50 MJ stored energy and 24 tonnes of NbTi hollow conductor.

Besides the clear advantage in thermal behavior, a key argument in favor of forced flow conductors is electromechanical integrity. In contrast with pool-cooled windings, where the individual turns are insulated only by spacers to allow the coolant to wet the metallic surface, the turns of a winding made of forced flow conductor can be fully insulated and potted. Such a rigid structure offers superior mechanical performance under electromagnetic load and withstands very high operating voltages (e.g., in pulsed or quench mode), not limited by the dielectric strength of the coolant. Clearly, the choice of forced flow conductors becomes mandatory for magnets with very large stored energy, subjected to high operating loads and requiring high-voltage discharge in case of quench.

Comparing the latest forced flow conductors with the early ones, manufactured thirty years ago, the difference is impressive. On one side, the conductor design has made substantial progress, using the results of extensive R&D work (mostly in ac loss and stability) and sophisticated thermohydraulic codes. On the other, the manufacturing methods have evolved under the pressure of the design goals, and advanced technologies have become reliable and affordable.

The present review is focused on the manufacturing aspects. In general, only practical industrial manufacturing experience is quoted. Short prototype conductors are reported only if an original method was used. The review is organized in three main sections. The first describes the conductor generation where the helium flows in smooth pipes and the superconductor is attached outside the pipes, without direct contact with the coolant. The second generation is a transition group of conductors, where a flat cable is encased in a welded steel sheath. The third generation includes the rope-in-pipe, or cable-in-conduit, conductors, where the coolant is forced to flow through a bundle of strands with large void fraction, encased in a structural conduit.

Superconducting cables for power transmission lines are not in the scope of this review. Their manufacturing methods, dominated by the high-voltage insulation requirements, are similar to those for conventional high-voltage conductors for underground lines.

### CONDUCTORS WITH SMOOTH PIPES

In the late sixties, the use of a helium-tight conduit produced by extended welded seams was considered too risky for large-scale applications (particle detectors, fusion magnets), where a single leak in the winding would have either severely delayed or compromised the entire project. It is understandable that copper pipes were initially selected for forced flow conductors, because of their high thermal conductivity, small pressure drop, and reliable leak-tightness.

Four conductor subgroups are identified according to the assembly procedure for attaching the superconductor to the copper pipes:

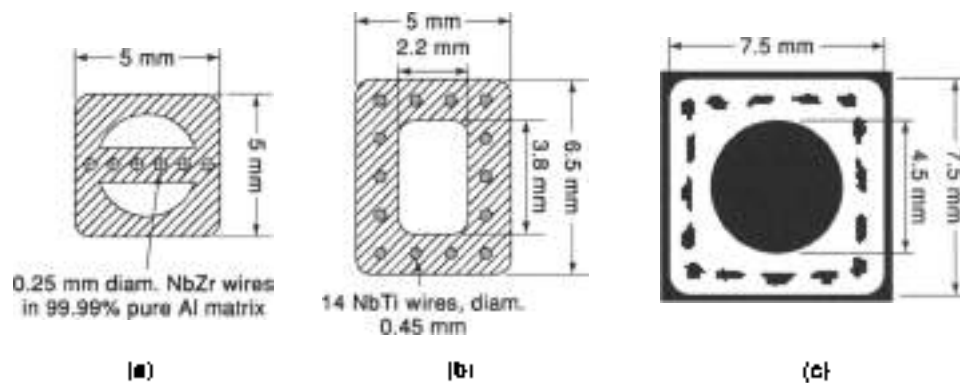
#### Superconducting Pipes

The first experiments with forced flow conductors look today like small masterpieces of technology. The idea was not to attach a superconducting wire to a pipe, but to create a hole in a superconducting composite.

A cross section of the first ever reported hollow superconductor is shown in Fig. 1(a). The extruded composite has a pure aluminum matrix with six NbZr filaments (0.25 mm thick) in the midplane, to allow bending radii down to 10 mm. The copper-plated NbZr filaments were supplied by Supercon (United States), and the hollow extrusion was done in 1967 at the Atelier électromécanique de Gascogne (France). Over 400 m of conductor have been produced and successfully wound into a solenoid at CERN (3).

The hollow conductor shown in Fig. 1(b) was commercially produced by Imperial Metal Industries (England) at the end of the sixties. The rectangular extruded pipe, 8.5 x 5 mm, is made of copper, with 14 NbTi filaments (0.45 mm diam) embedded in it. The production lengths ranged up to 200 m. At CERN, the conductor sections were electron beam welded (4) and wound into a solenoid of 32 double pancakes, for a total of 2500 m of conductor. At Saclay, France, a smaller, racetrack coil (5) was wound with 300 m of the same conductor.

At M.F.O. (Switzerland), later BNL and ABB, a larger square superconducting pipe was developed (6) by hollow extrusion of a copper billet with 16 NbTi filaments (0.55 mm diameter) (see Fig. 1(c)). Because of the large cross-section, the conductor unit length was limited to 100 m.



**Figure 1.** Three extruded hollow superconductors: (a) Al matrix with NbZr filaments (1967, courtesy of CERN), (b) Cu matrix with NbTi filaments (CMI, 1969), (c) round to square Cu matrix with thick NbTi filaments (MFO, later ABB, about 1968).

The manufacture of a hollow superconductor at the Electro-Technical Laboratory, Tsukuba, Japan (ETL) was reported in 1974 (7). The method is shown schematically in Fig. 2(a). Twelve square rods of superconducting composite (6 × 6 mm), each with 18 NbTi filaments in a copper matrix, were inserted in a rectangular copper tube and rolled with intermediate annealing to a flat strip, 1.8 mm thick. The strip was then welded in a conventional tube mill and drawn down to form a rectangular hollow conductor, 4 × 5 mm; see Fig. 2(b). Due to the rolling process, the 116 filaments, with equivalent diameter 0.25 mm, are flattened and nonhomogeneously spaced.

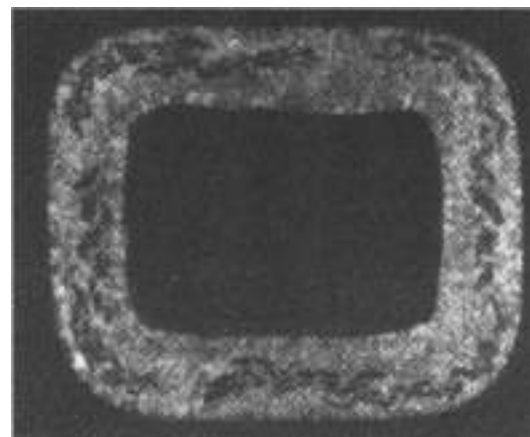
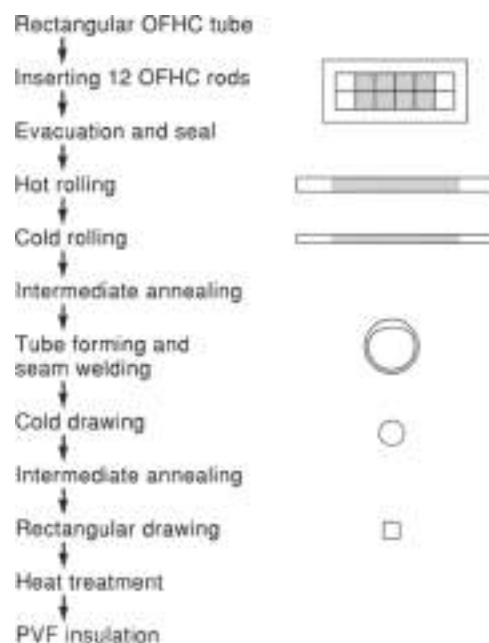
The major drawbacks of such "superconducting pipes" are the large size (leading to poor critical current density) of the superconducting filaments due to the limited reduction ratio, and their nontransposition with respect to the transverse field,

due to the lack of twisting. As a result, large flux jumps were observed in operation.

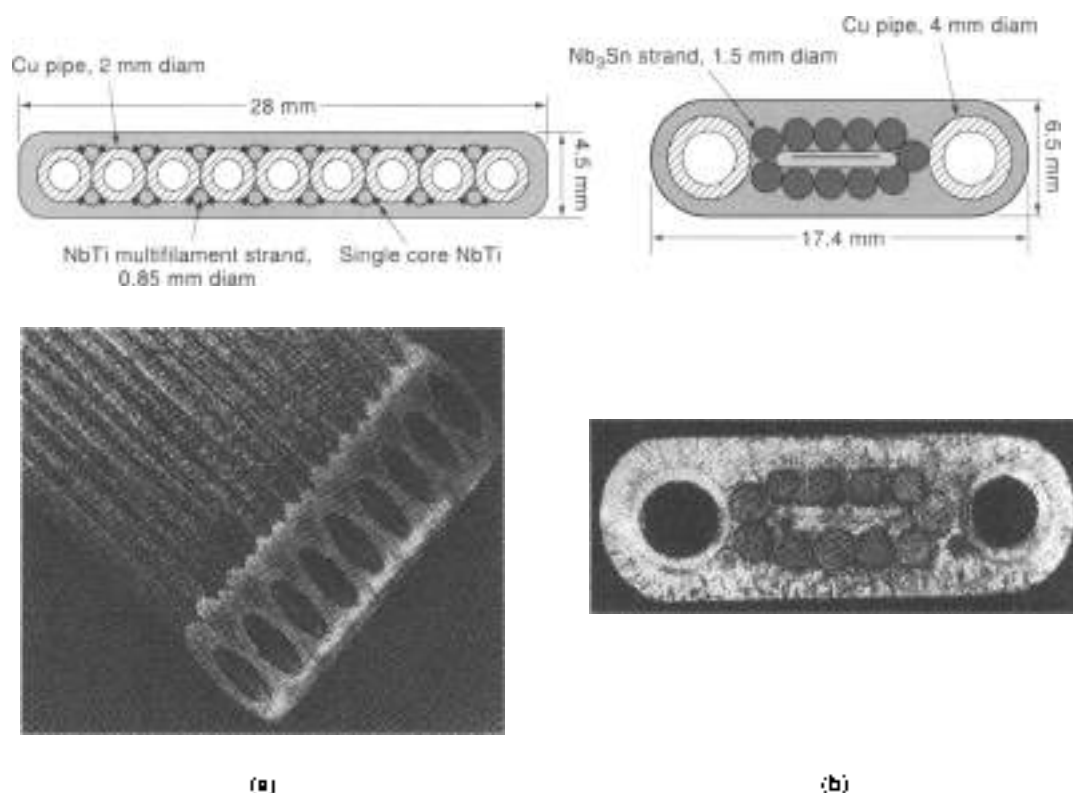
#### Conductors Assembled by Electroplating

An original method was developed at the Kurchatov Institute (Moscow) in the late sixties to bond together the superconducting strands into large composites by continuous electroplating in a  $\text{CuSO}_4$  electrolyte (8). The same method was later used to assemble the forced flow conductors for the toroidal coils of the Tokamak-7 (U7) and Tokamak-15 (T-15).

The T-7 conductor (9) is made from a linear array of four copper pipes, 2 mm inner diameter, with 16 multifilamentary NbTi strands sitting in the grooves between the pipes; see Fig. 3(a). Another 32 thin, single core NbTi wires are placed



**Figure 2.** A hollow superconductor (7), manufacture in 1974 by a tube mill method (courtesy of K. Aoyama, ETL, Japan).



**Figure 3.** Cross section of the conductors assembled by copper electroplating at Kurchatov Institute: (a) the NbTi-based T-7; (b) the react-and-wind Nb<sub>3</sub>Sn conductor for the T-15. (Courtesy of Kurchatov and Bochar Institutes, Moscow.)

in the grooves between the multifilamentary strands and the copper pipes to obtain an even surface. The pipe strands assembly is bonded by electroplating a 0.6 mm copper layer up to the final size of 28 × 4.5 mm. The strands are not transposed, and the conductor has suffered from severe flux jumps, triggering quenches during ramp up and down. The large number of parallel cooling channels, supplied by two-phase helium, results in a highly nonhomogeneous distribution of the mass flow rate. Over 10 km of conductor have been manufactured in units longer than 200 m.

In the conductor for the T-15 (10), a flat cable of 11 nonstabilized Nb<sub>3</sub>Sn strands (1.5 mm diam) around a bronze strip was bonded after heat treatment to two copper pipes (3 mm inner diameter) by an electroplated Cu layer, 1.2 mm thick, see Fig. 3(b). Some copper wires were included between the cable and the pipes to obtain a more regular envelope. To minimize the bending strain during winding, the heat-treated cable is close to the neutral bending axis of the conductor. In the winding of the T-15 coils, a wet insulation method had to be applied to smooth the uneven conductor surface. Over 100 km of conductor have been manufactured, in units of 200 m.

After the T-7 and T-15 manufacturing experience, the technique of assembly by electroplating was not applied any more. The low process speed, the large electric power requirement, and the poor dimensional tolerance overcame the advantage of a good, low-resistance bonding of the components.

#### Strands Soldered on a Central Copper Pipe

**The OMEGA Conductor.** The conductor for the OMEGA detector at CERN is a historical milestone in the development of

forced flow superconductors (11). The development started in 1968, with a prototype consisting of a round-in-square copper pipe with eight multifilamentary strips located in longitudinal slots. The 40 × 40 mm assembly was fed into an outer copper pipe and compacted by drawing to the final size of 16 × 15 mm (Fig. 4(a)). This first approach was abandoned because of the high cost, the limited manufacturing length (10 m), and the lack of transposition.

The final cross section is shown in Fig. 4(b). Manufactured in 1970 at BBC (Switzerland), the 18 × 18 mm conductor is a square copper pipe with residual resistance ratio (RRR) > 250, surrounded by two layers of wires (1.5 mm diam) cabled in opposite directions. In the first layer, 30 out of 36 wires are superconducting, with four Nb<sub>3</sub>Sn thick filaments (250 μm) in a low RRR copper matrix. The strands are initially not twisted, but, as they are cabled on the copper pipe, no back rotation of the cage strander is allowed, so that the filaments in the strand have the same twist pitch as the cable (200 mm). The second layer, with protective function, contains 40 copper wires and is cabled in the opposite direction, with pitch 180 mm.

The cabled conductor was driven into a SnAg<sub>5</sub> bath and pulled through a square die, with a die angle of 45°, cooled at constant temperature to solidify the solder; see Fig. 5. Twelve sections, about 1 km each, were produced for a total of 24 tonnes (11.4 km). The soldering process for each section took about 35 h, at a speed of 0.45 m/min. The online quality assurance included eddy currents and ultrasonics to check the quality of the solder impregnation.

**The OMEGA's Brothers.** The smallest brother of the OMEGA is a rectangular conductor developed in the early seventies

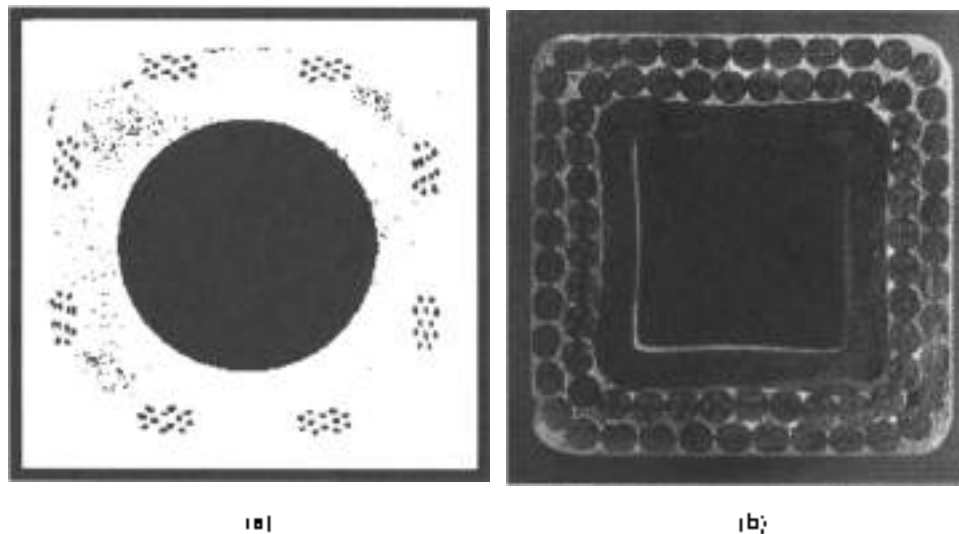


Figure 4. OMEGA conductor: (a), prototype (left); and (b) final soldered cable. Courtesy of ABB.

at ETL, Japan (7), consisting of a copper pipe with one layer of 52 wires (0.82 mm diam) cabled around and impregnated with SnAg solder. Half of the wires are multifilamentary NbTi strands, alternated to copper wires. In contrast with the OMEGA, the wires are cabled and soldered as a round conductor and later drawn down to the final rectangular size of  $5 \times 4$  mm; see Fig. 6(a).

In the mid seventies, a conductor of identical size to the OMEGA ( $15 \times 18$  mm) was manufactured at CERN for a large superconducting dipole (12). The copper pipe, with 10 mm bore, is round, and only one layer of NbTi multifilamentary strands and copper wires is cabled on it. The cable is encased between two copper profiles and bonded by SnAg soldering; see Fig. 6(b). The unit length is up to 270 m, and the overall production is more than 5 km.

In the late seventies, a hollow conductor [see Fig. 6(c)] was manufactured at VAC, Germany (13). A loose layer of 20 NbTi

multifilamentary strands and 10 copper wires were cabled and soldered in the same process around a square copper pipe. The solder is SnPb eutectic alloy, and the final size is  $8.2 \times 8.2$  mm. A first conductor batch had to be replaced, as it did not withstand the tight bending radius (over 4% bending strain). About 500 m of conductor was used for the test facility SAMPO at ENEA (formerly CERN), Italy.

A similar, bigger conductor [see Fig. 6(d)] was manufactured in 1981 at Europa Metall, Italy (14). The 33 NbTi strands (1.3 mm diam) were first cabled on a round copper pipe and soldered by SnPb alloy. The conductor was then shaped approximately to a rounded square,  $14.6 \times 14.6$  mm. About 9 km of conductor have been manufactured and wound into the inner module of the SULTAN test facility (Villigen, Switzerland).

Another little brother of the OMEGA [see Fig. 6(e)] has been manufactured at BBC in 1981, for a Sector Cyclotron, SuSe (15), in Munich, Germany. The rectangular copper pipe has a tiny hole,  $1 \times 2$  mm. Two kinds of NbTi strands, with 0.7 mm diameter and different Cu non-Cu ratios, were cabled around the copper pipe and soldered with SnAg alloy. The overall conductor size is  $4.5 \times 5.5$  mm. About 4.5 km of conductor has been wound into the winding of the sector cyclotron, with a minimum bending radius of 68 mm (3.3% bending strain).

**The Nb<sub>3</sub>Sn Hollow Conductor at ETL, Japan.** Apparently, the method of cabling and soldering the superconducting strands on a copper pipe does not apply to Nb<sub>3</sub>Sn conductors, soldering must be carried out after heat treatment, but a heat-treated strand cannot be cabled because of its brittleness. Agatauma et al. (16) reported in 1978 an ingenious method to prepare a Nb<sub>3</sub>Sn conductor soldered on a copper pipe, according to the steps in Fig. 7, called the 'react, wind, and solder' method.

Initially, two round, hollow braids (12 and 24 strands) are made from thin, Sn plated Nb bronze composites (0.1 mm diam). The hollow braids are flattened to a  $\approx 0.6$  mm thick ribbon, wound on a large-diameter holder, and heat treated to form the Nb<sub>3</sub>Sn composite. In a single process, two small and two large braids are drawn, together with a SnAg solder foil, into longitudinal slots on a rectangular copper pipe. The conductor, wrapped with a polyimide film and a reinforcing

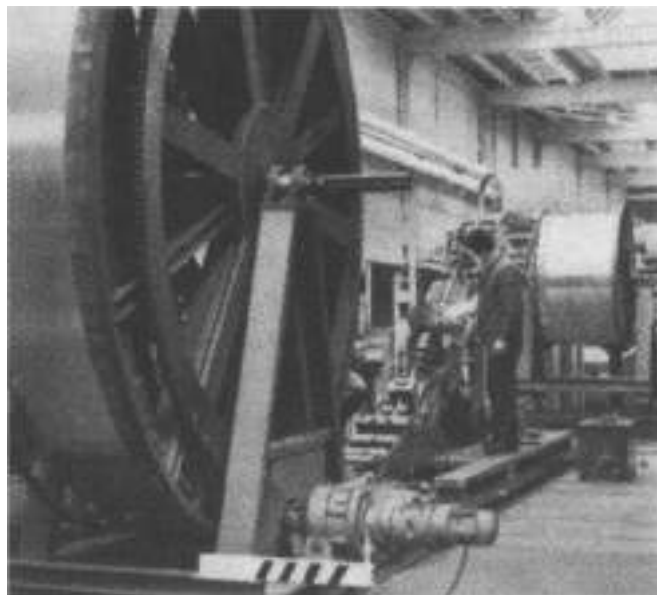


Figure 5. Soldering of the OMEGA conductor (courtesy of ABB).

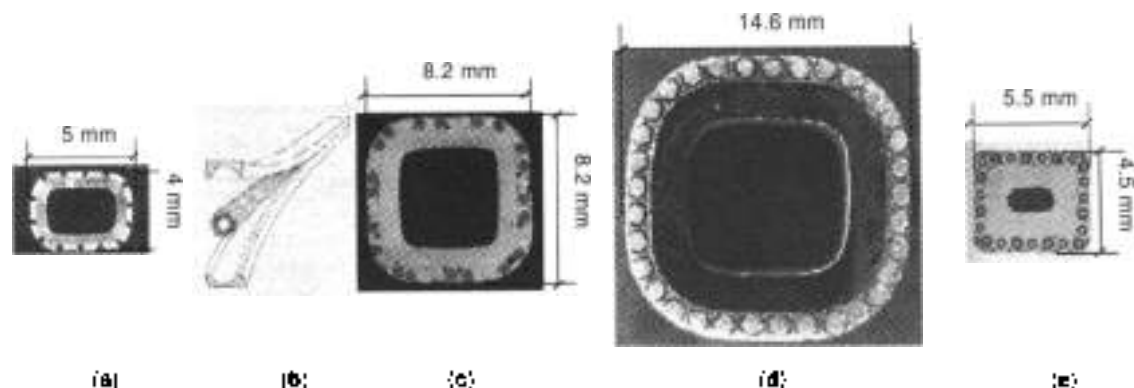


Figure 6. Soldered hollow conductors manufactured similarly to the OMEGA conductor. See text for details.

steel tape, is wound into a coil with 75 mm minimum bending radius. After winding, the coil is heated in vacuum to melt the solder and bond the braids to the copper pipe, see Fig. 7.

If the conductor were a solid monolith during the winding process, the bending strain at the Nb<sub>3</sub>Sn filaments would be in excess of 8%, which would be fatal for the conductor (to avoid irreversible damage in Nb<sub>3</sub>Sn, the bending strain should be below  $\pm 0.5\%$ ). However, as the braids have not yet been bonded to the copper pipe, the strands can settle in the slots and the actual bending strain is much smaller. A length of 160 m conductor was used for the test coil, which was successfully tested, demonstrating the viability of the process.

The strands in the braids are transposed, but not the four braids placed in the longitudinal slots. To obtain a high current

conductor, a very large number of thin strands must be braided. The conductor layout, with thin braids placed on the perimeter of the pipe, does not allow a large operating current density. Since the successful demonstration, no further application of the process has been reported.

**The Swiss ECT Conductor.** The most complex conductor belonging to the OMEGA family is the Swiss ECT conductor, manufactured in 1981–1983 at BNL for the Swiss contribution to the Large Coil Task at Oak Ridge. Initially, two forced flow conductor layouts were developed, one being a two-stage cable in a conduit and the other a multistage soldered cable around a copper pipe (17). The reasons to prefer the second option were mainly the cost and the slowness of the first. Today, this choice may sound surprising and the same reasons would rather lead

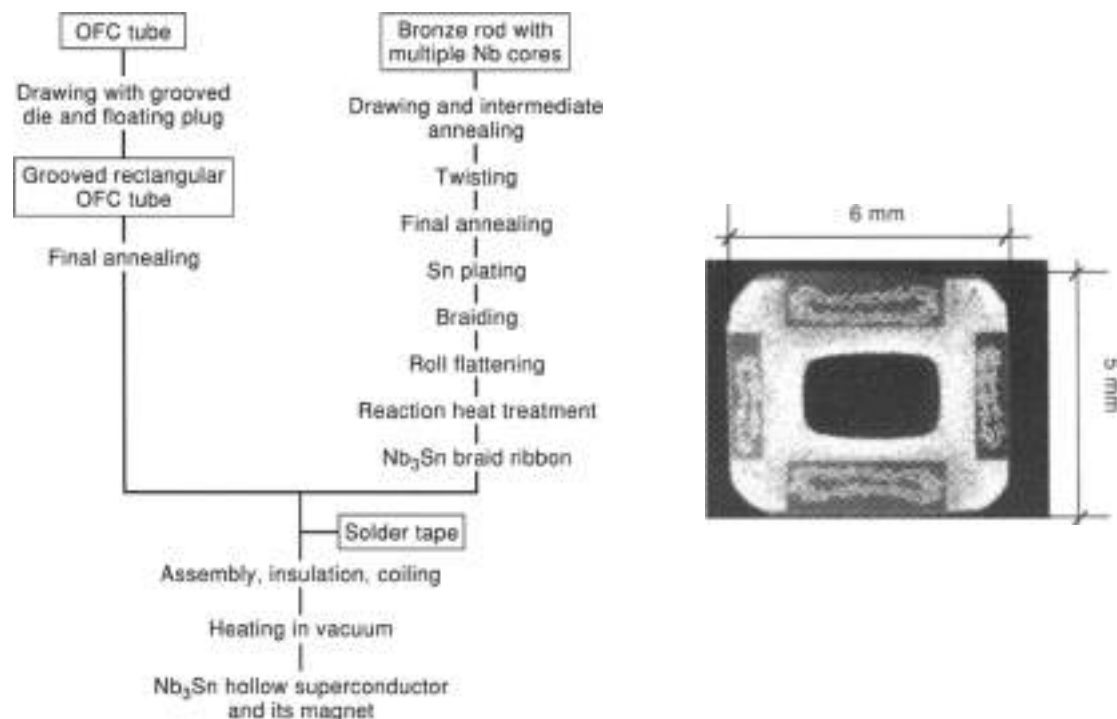


Figure 7. Nb<sub>3</sub>Sn hollow conductor, with four braids soldered in the slots of the Cu pipe (16) according to the react, wind, and solder process (courtesy of K. Agatsuna, ETL, Japan).

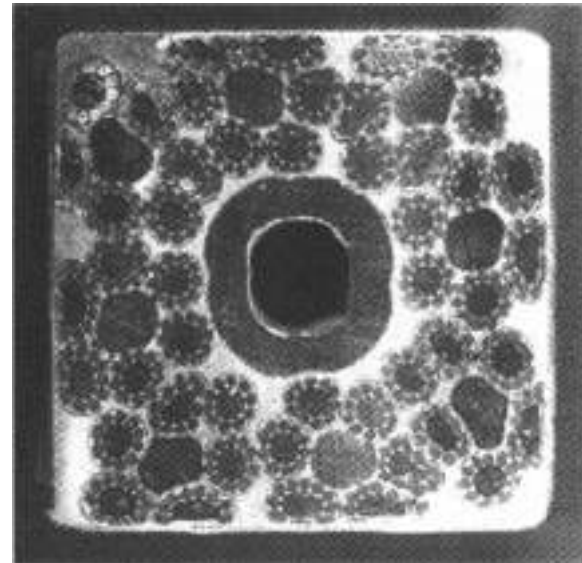
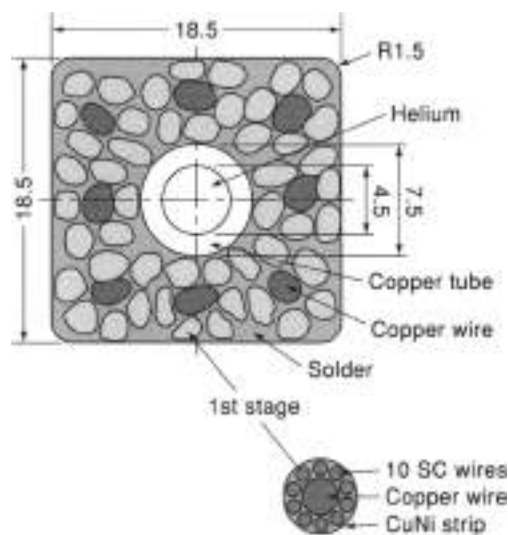


Figure 8. The Swiss LCT conductor. Dimensions in millimeters. (Courtesy of ABB.)

us to prefer the cable-in-conduit option. However, in 1981 BSC had developed a broad knowhow on soldered conductors, and almost no industrial experience was available on welding of long lengths of steel conduit.

The cable layout is summarized in Fig. 8. The first cable stage consists of 10 multifilamentary NbTi strands (0.46 mm diam) cabled around a copper wire and solder-impregnated. To limit the air losses, the first cable stage is tightly wrapped with a 0.05 mm thick CuNi strip and soldered again to bond the wrap to the cable. To maintain full transposition, the second cable stage is made of six subcables around a thick copper core (2.25 mm diam). Eight second cable stages are eventually cabled with the tight pitch of 150 mm around a round copper pipe (7.5  $\times$  4.5 mm), wrapped with CuNi foil. The cable is compacted and shaped by rolls to a square of 18.5  $\times$  18.5 mm. Despite the large load required to compact the rigid, solder-impregnated components down to a local void fraction of 30%, neither damage nor performance degradation was observed in the compacted cable.

For the final solder filling of the multistage cable, a special CdZnAg alloy was selected (18), with superior mechanical properties to those of PbSn and SnAg alloys. As the melting point of the solder is about 350 C, the process speed had to be maintained in the range of 2 m/min to avoid degradation of the NbTi critical current density. Due to the high speed, the die needed to be specially laid out. To prevent solder oxidation and for environmental reasons (Cd vapor), the process was carried out under nitrogen atmosphere. The overall length produced for the Swiss LCT conductor was  $\sim$  5.5 km, in units of 250 m.

The Swiss LCT conductor marked the end of a successful series of hollow conductors cabled and soldered around a copper pipe. As in the OMEGA and its brothers, a single-stage cable was used, so the achievable current density was low, because only one layer of strands could be placed on the outer perimeter of the pipe. The option to use a multistage cable around the pipe was visible, but turned out to have drawbacks: the solder cross section climbed above 25%, and complex, high-resistivity barriers had to be included to avoid large coupling-current losses.

**Flat Cables Soldered to Copper Pipes.** In Nb<sub>3</sub>Sn react-and-wind conductors, the superconducting strands must stay as close as possible to the neutral bending axis to avoid substantial strain-induced degradation during winding. This requirement, together with the wish to maintain the advantages of the copper pipes, drove in the eighties the development of a number of flat cables assembled by soldering to copper pipes. The ECN (The Netherlands) and the SIN (later CRPP, Switzerland) carried out the conductor development and manufacture as part of the construction of the SULTAN test facility.

The smallest of this conductor subgroup is a NbTi-based flat cable soldered to the broad side of a rectangular, half-hard copper pipe to build a square conductor, 8.4  $\times$  8.4 mm; see Fig. 9(a). The cable is made by 16 multifilamentary strands with 75 mm pitch and is not at the neutral bending axis, but this is not crucial for the ductile NbTi conductor: the bending strain in the winding is smaller than 1%. The soldering process was carried



Figure 9. Flat cables soldered to copper pipes: (a) NbTi ECN 9 T, 8.4  $\times$  8.4 mm; (b) Nb<sub>3</sub>Sn SIN 12 T, 21.5  $\times$  8.4; (c) Nb<sub>3</sub>Sn ECN 12 T, 26.2  $\times$  7.0 mm; (d) Nb<sub>3</sub>Sn SIN 9 T, 26.3  $\times$  9.6 mm. (Courtesy of the SULTAN group.)

out at MCA, Massachusetts, in 1982 (19). A wave soldering device was not sufficient to fill the interstices satisfactorily and had to be modified by adding a restraining die, where the solder becomes solid, as in the OMEGA process. The solder cross section is less than 5%. Ten sections have been produced, each 500 m long.

The  $Nb_3Sn$  react-and-wind conductor in Fig. 9(b) was manufactured at SIN (later CRPP) in 1985 (20). The flat cable is made of 2016 multifilamentary strands (0.125 mm diam) according to the external Sn method. The first cable stage (not fully transposed) is composed of 7 strands (1 + 6). The second cable stage (also nontransposed) is composed of two layers of strands (6 + 12) cabled around a copper core. The last stage is a 970 m long flat cable (0.3 × 13.8 mm) of 16 subcables around a bronze strip. The flat cable is heat-treated on an Inconel drum (7 layers) at a radius of 600 mm. This allows straightening the cable after heat treatment without irreversible damage of the  $Nb_3Sn$  filaments and bending it eventually to the minimum winding radius, ≈300 mm, with the tolerable bending strain of ±0.28% in  $Nb_3Sn$ .

The other components to be attached to the flat cable are a copper pipe, a copper profile as stabilizer, and two stainless steel strips as a structural reinforcement. These four parts have been assembled with three brazing foils and brazed at 630 °C for a total length of 970 m in a continuous process, at the rate of 0.3 m/min. The process could be stopped and restarted to replace the heating elements. The 200 m long sections of the copper pipe were preliminarily brazed to a full length, using a higher-melting alloy.

The brazed steel-copper composite and the heat-treated  $Nb_3Sn$  cable were assembled in a combined continuous milling-soldering process. A slot is milled in the copper strip, and the cable is carefully placed in the slot. Bonding of the cable to the copper is achieved by a wave soldering device: the PbSn alloy fills up to 80% of the void area in the cable. The process speed is 0.5 m/min. Stop and restart was necessary every 70 m to replace the cutting tool. As the thickness of the heat-treated cable varied at the layer transition, the thickness of the slot had to be periodically adjusted. The 21.5 × 8.4 mm conductor was produced in a single length of 870 m and wound into a solenoid. The coil performance confirmed that, besides the expected bending strain, no significant degradation occurred during the manufacturing process. The de-operated coil did not suffer from the lack of transposition in the cable.

For the  $Nb_3Sn$  conductor manufactured in 1986 at ECN (The Netherlands), see Fig. 9(c), some features have been simplified (21). The flat cable, 18.5 × 1.65 mm, is made of 36 multifilamentary strands (1 mm diam, powder-in-tube method, without central strip). The seven cable sections, about 140 m long, are wound as loose pancakes on steel supports and heat-treated. The other conductor components (two steel strips, two copper pipes, and two copper strips) are all assembled to the heat-treated cable in one run by a wave soldering device. The steel strips were preliminarily electroplated with solder to ease the bonding. The process speed was 1 m/min. Six out of seven conductor sections (total about 800 m) were wound as double pancakes with interpancake joints.

To complete the SULTAN test facility in the split-coil configuration, three additional conductor lengths were manufactured at SIN in 1988 and 1990; see Fig. 9(d), (22). The flat cable, 13.7 × 3.54 mm, is made of 98 multifilamentary strands (0.65 mm diam, bronze method), cabled in two stages, (1 + 6) ×

14, around a central bronze strip. After heat treatment, the components are assembled in a single run (See Fig. 10), as in the ECN conductor, by a wave soldering device, at the rate of 0.3 m/min. Two conductor lengths, 27.0 × 10.3 mm, each ≈ 800 m long, were wound into the two 9 T coils of the SULTAN facility. A third conductor section, 26.1 × 7.6 mm, 800 m long, with thinner steel strips, was used for the innermost, 12 T coils.

## SHEATHED FLAT CABLES

The first attempts to place the active superconducting elements inside a hydraulic conduit date back to the late seventies. Besides the group of bundle conductors (see below), a small number of conductors have been developed by encasing a more or less rigid flat cable into a welded conduit.

### The EU-LCT Conductor

Vacuumschmelze (Germany) manufactured the EU-LCT conductor in 1980–1981 (23) after three years of development (24). The design optimization effort (ac loss, stability, and structural aspects) led to a sophisticated conductor layout, requiring four continuous manufacturing processes to assemble the components at very tight tolerance. The single-stage, NbTi flat cable (stainless-steel-sheathed, 40 × 10 mm) is shown in Fig. 11.

Twenty-three NbTi rectangular composites, 2.35 × 3.1 mm, are cabled and soldered on a flat core, which has a central insulating foil to avoid large interstrand coupling-current losses. The core is prepared starting from a NiCr 0.5 mm thick strip, clad on one side with a 50 μm thick CuNi layer to provide a base for soldering. A 4 μm thick Sn coating is electroplated on the CuNi side of the strip. The strip is folded on a 50 μm thick Kapton foil to form the core of the cable. To improve the strength, the folded strip is stretched, adding 6% cold work. The tolerance on the final size of the strip is (33.5 ± 0.55) × (1.0 ± 0.03) mm. Two foils of SnPbIn solder, 50 μm thick, are then attached to the strip by fusing at both edges. The rate of the process is 0 m/min.

A cage strander with full back twist is used to wind the 23 composites edgewise on the flat strip, without adding torsion to the bending. A Ruebel head at the cabling point bends the composites accurately at the edge of the flat core, always maintaining the short side of the composite parallel to the flat core; see Fig. 11. Spacing wires cabled together with the superconducting composites keep the 0.8 mm space between the strands. In the same process, the strands are bonded to the core by melting the solder foils. This is achieved by pressing hot metal blocks on the cable, the strand temperature is 200 to 210 °C for about 10 s. The load is maintained till full solidification occurs. Afterwards, the NiCr spacing wires, coated with a refractory varnish, are removed. The transposition pitch is long (400 mm); however, the interstrand coupling-current loss is small, as the strands are electrically in contact only through the thin solder layer. The speed of the cabling-soldering process (See Fig. 12) is 1 m/min.

The 0.5 mm thick sheath is a folded strip of nitrogen-alloyed austenitic steel (316LN). To create additional cooling channels and to limit the heat load to the superconductor during the welding, six steel spacers, 0.6 mm thick, are placed between the cable and the sheath; see Fig. 11. The six spacers are preliminarily attached to the 98 × 0.8 mm strip in a continuous

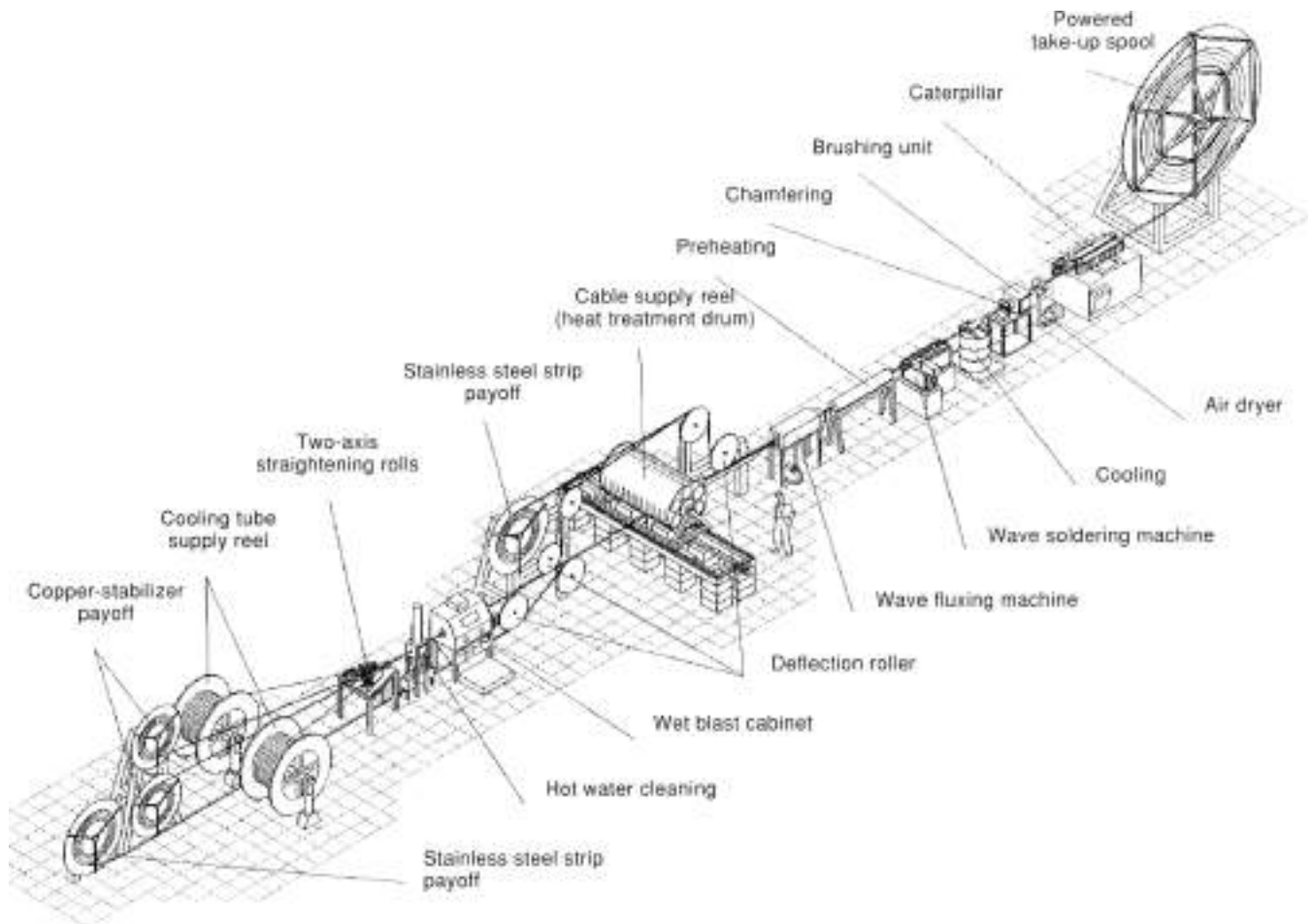


Figure 10. The milling-soldering line for the final assembly of the  $Nb_3Sn$  react-and-wind conductor for SULTAN (courtesy of G. Pasztor, CRPP).

process, by spot electron beam welding, with minimum heat input to the thin strips. To avoid excessive outgassing in the vacuum environment of the electron beam, the steel for the spacers is 316Ti (not nitrogen-alloyed).

The cable sheathing is the fourth and last continuous process for the conductor assembly; see Fig. 12. Six shaping rollers fold the ribbed strip into a U, where the cable is fitted, and another six mills fold the strip around the cable, leaving a welding gap  $\approx 0.1$  mm. A tungsten inert gas (TIG) pulsating torch carries out the closing weld. The strand temperature underneath the weld increases to 260 °C during 5 s; the temperature rise does not affect the  $NbTi$  critical current density. Due to the electrode replacement, the process needed to be interrupted every 50 m. The welding parameters at the restart points are adjusted to guarantee a reliably tight weld transition.

The final tolerance on the conductor size is  $\pm 0.1$  mm. The corner radius is 2 mm. The conductor has been produced in units of about 500 m, for a total length of 7 km.

### $Nb_3Sn$ Sheathed Flat Cables

The flat cable geometry is attractive for the  $Nb_3Sn$  react-and-wind method because the active, strain-sensitive superconductor can be placed, after the heat treatment, on the neutral bending axis, allowing the conductor to be wound into

a coil with minimum degradation. Several such developments were started in Europe and Japan in the eighties, but only few led to significant manufacturing applications.

The three prototype conductors, developed in the mid eighties for the initial phase of the Next European Torus (NET) project, were eventually manufactured only in  $\approx 10$  m sections (25–27). The rectangular conductors include a heat-treated  $Nb_3Sn$  flat cable, 4.0 to 6.6 mm thick. An insulating core in the center of the cable to reduce the coupling loss was shown not to be necessary (27), so the cable thickness could be reduced. The stabilizer, assembled either as a mixed matrix monolith or as a transposed cable, is soldered to both sides of the flat cable. The conduit is assembled around the soldered conductor by welding four steel strips. The cooling channels are placed underneath the full-penetration weld seams: a strip sliding process is not allowed, due to the required conduit thickness (2.5 to 3.2 mm). Laser beam welding was needed to reduce the heat load to the soft soldered cable and to maintain the tight assembly tolerance.

A Russian prototype react-and-wind conductor (28) has a layout similar to the NET prototypes, with four stainless steel strips welded to form the outer sheath. However, the helium channels are eight parallel copper pipes soldered to the central  $Nb_3Sn$  flat cable, and the weld seams do not need to be heli-



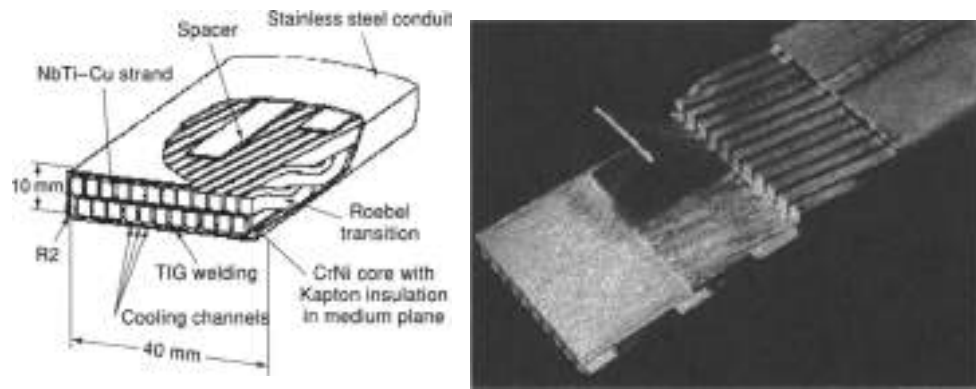


Figure 11. The EU-LCT conductor (courtesy of Vakuumschmelze, Germany).

**The ETL Wind-and-React Flat Conductor.** An unusual example of a Nb<sub>3</sub>Sn flat cable-in-conduit, wind and react conductor was developed in the early eighties at ETL, Japan [29]. Two coreless flat cables, each made of 15 stabilized Nb<sub>3</sub>Sn strands 1.4 mm in diameter, are separated by a CuNi strip to reduce the coupling loss (lack of transposition). Two thick copper strips are rolled

to form ribbed, U-shaped profiles with large cooling channels for the helium flow (the coolant is also allowed to flow in the interstices between cables and stabilizers). The cable sandwich is assembled with the copper stabilizer and wrapped with a punched CuNi tape; the components are not bonded together. The 0.75 mm thick outer copper sheath is welded in a tube

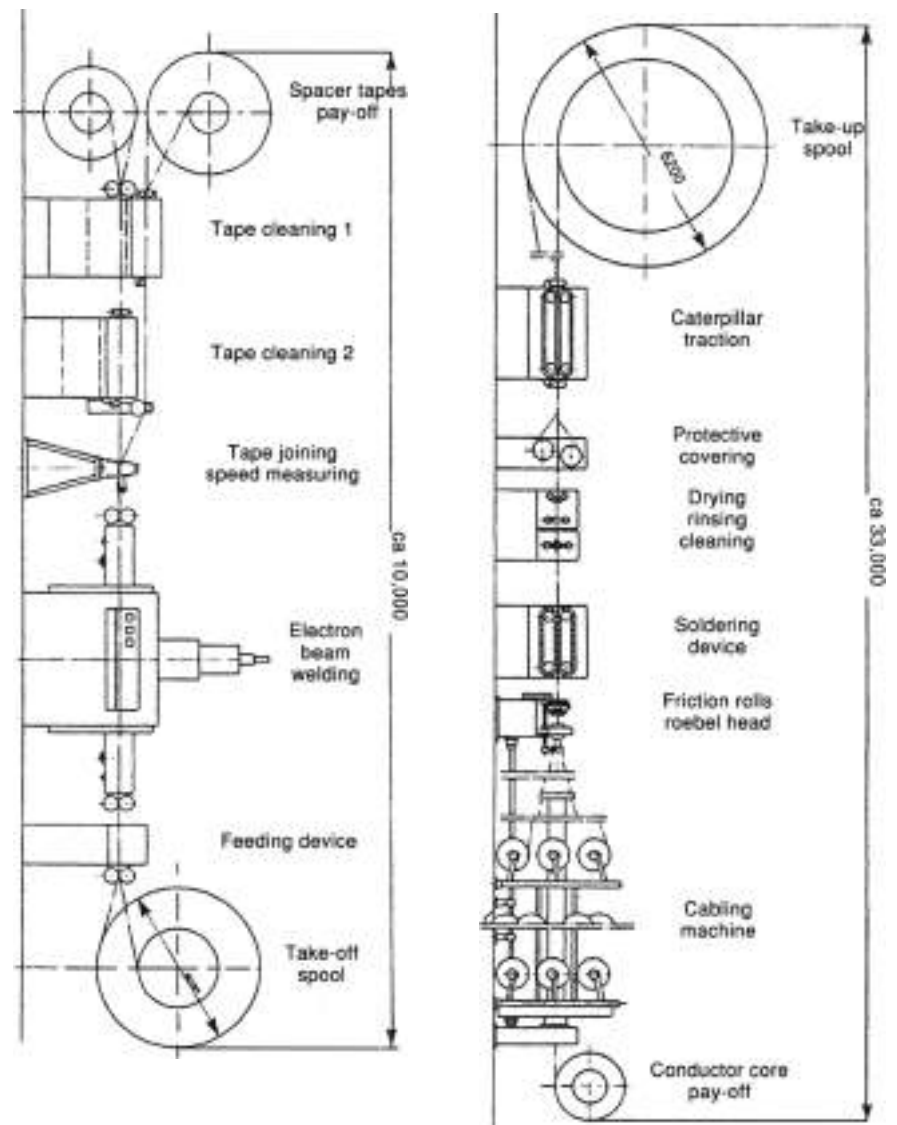
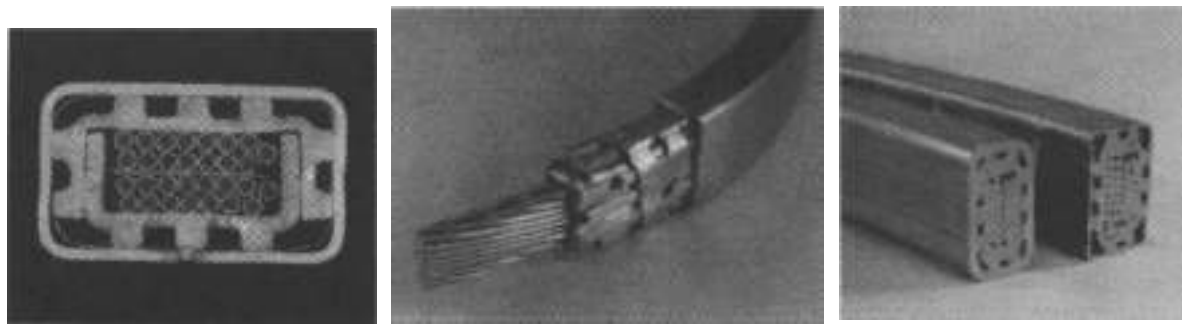


Figure 12. The cabling (soldering upper) and the sheathing (lower) processes for the EU-LCT conductor (courtesy of Vakuumschmelze, Germany).



**Figure 13.** The  $13 \times 23$  mm  $\text{Nb}_3\text{Sn}$  wind-react conductor developed at ETL: copper-sheathed conductors, left and middle; stainless-steel-sheathed conductors, right (courtesy of K. Agatsuma, ETL, Japan).

mill. The wrapped cable–stabilizer is pulled through the welded copper pipe and drawn down to the final dimension of  $13 \times 23$  mm; see Fig. 13. After the draw, the conductor passes through a high-frequency annealing furnace to soften the sheath. The production unit length ( $\approx 10$  m) is limited by the size of the drawing bench.

Eight conductor sections (total about 300 m) were insulated by glass tape, wound into a coil with 300 mm inner diameter (over 4% bending strain) and eventually heat-treated to form the  $\text{Nb}_3\text{Sn}$ .

The same manufacturing process was also applied for a stainless-steel-sheathed conductor (30). The selected steel was 316L, the wall thickness 0.6 mm. Wrinkling of the sheath was observed at the smallest winding radius. The properties of the longitudinal TIG welding seam were investigated before and after heat treatment (500 C, 50 hr). A pancake coil was wound, heat-treated, and tested with eight other pancakes based on Cu-sheathed conductors. The busbars for the coil are manufactured with a thicker (1.0 mm) steel sheath; see Fig. 13.

**The TMC-FF Conductor at JAERI.** The development of  $\text{Nb}_3\text{Sn}$  flat cable, react-and-wind conductors for toroidal field coils started at JAERI in the early eighties with a short prototype, called STEP-1 (31); see Fig. 14(a). A Nb layer was included at the outer surface of the strand as a thermal barrier, to delay the diffusion inside the strand of heat from friction at the strand surface. Eventually, the strand was drawn through a star-shaped die cutting four grooves in the Nb outer shell to allow heat exchange between coolant and stabilizer. The sheath was planned to be a thin CuNi strip with an outer steel

reinforcement. However, the short prototype was jacketed in a steel U profile with a welded lid.

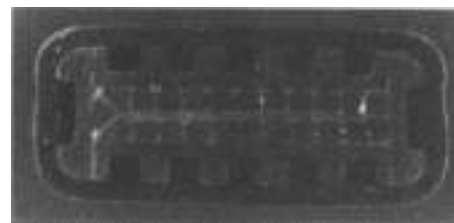
The actual toroidal model pancake was wound in 1991 with 90 m of monolithic flat cable manufactured by Hitachi Cable (32); see Fig. 14(b). The 4.5 mm thick  $\text{Nb}_3\text{Sn}$  cable consists of 28 monostabilized strands soldered after heat treatment to a number of C-shaped copper profiles, to build a monolithic, ribbed assembly. To increase the loop resistance and lessen coupling current losses, a few epoxy spacers are interleaved between the stabilizer elements. It was planned to build the conduit by longitudinal welding of two hot-rolled, 3 mm thick steel profiles to obtain sharp outer corners and improve the conductor's rigidity. However, this development was skipped, and the conduit for the 90 m model conductor was a 2 mm thick steel sheath, TIG-welded in a tube mill process.

**The OPC-EX Conductor at JAERI.** The  $\text{Nb}_3\text{Sn}$  conductor for the Demonstration Poloidal Coil (DPC-EX coil) was manufactured in 1988–89 at Mitsubishi Electric (33); see Fig. 15. Over 500 m of the  $10.2 \times 40.8$  mm conductor have been produced in two sections and wound after heat treatment into two double pancakes. From the design point of view, the DPC-EX represents the transition between the generation of the flat lead-in conductors and the emerging group of bundle conductors (cable-in-pipe or cable-in-conduit).

The 158 Cr-plated strands (0.61 mm diam) are cabled in three stages,  $3 \times 3 \times 17$ . The last cable stage has a steel strip core to cut the interstrand coupling loss. The helium void fraction in the cable space is in excess of 40%. At both sides of the cable, two stainless steel square pipes provide the main cooling channels. The helium in the cable space is quasistagnant.

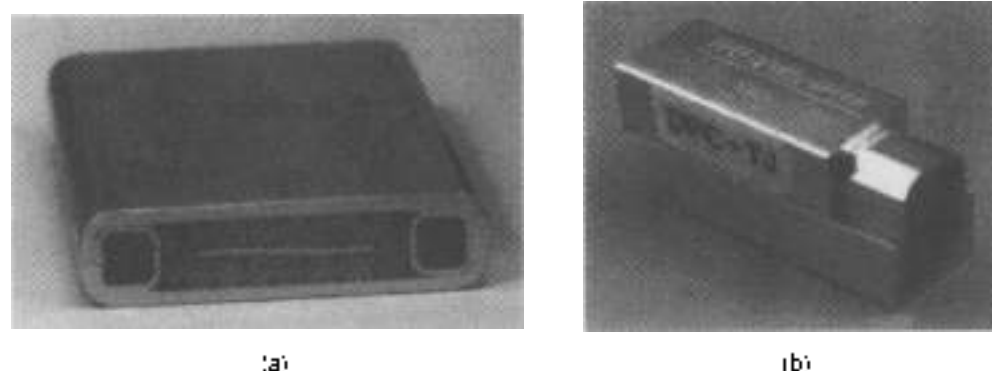


(a)



(b)

**Figure 14.** Heat-and-wind  $\text{Nb}_3\text{Sn}$  conductors developed at JAERI for toroidal field coils: (a) STEP-1, and (b) TMC-FF (courtesy of JAERI, Naka).



**Figure 15.** React-and-wind poloidal  $Nb_3Sn$  conductors developed at JAERI: (a) the DTY-EX 102 ( $\times 40$ ); (b) the DPC-TJ with the steel shell machined reinforcement welded longitudinally around the thin conduit. (Courtesy of JAERI, Naka.)

every 0.3 m, the steel pipes have *swirlnoles* (1 mm diam) to improve the heat exchange with the cable and to avoid a large pressure rise in case of quench. The outer sheath and the two pipes are made of a stainless steel (A1C-1) specially developed (34) to maintain high strength and ductility at low temperature after the  $Nb_3Sn$  heat treatment at 700 C, 30 h. The conduit has constant thickness, 1.5 mm, and is assembled from two strips, folded and welded by two simultaneous TIG seams at the edge side, in a modified tube mill (the steel pipes underneath the seams allow achieving full penetration without damaging the strands).

### CABLE-IN-CONDUIT CONDUCTORS

The development of cable-in-conduit conductors (CICC) started slowly in the late seventies, with only few significant applications till the late eighties. However, as soon as the manufacturing methods had been mastered by the industry, the CICC became very popular, and today they dominate the restricted market for forced flow superconductors. The main design motivation for introducing the cable-in-conduit conductor is its large wetted perimeter, improving the heat exchange and stability for pulsed field applications. The initial development was aimed at  $Nb_3Sn$  react-and-wind technology. In the nineties, the use of CICC was extended to the wind-and-react method, as well as to the  $NbTi$  technology.

Investigations of the stability and ac loss performance of the CICC stimulated a number of experiments with short-length, small-size conductors manufactured on purpose. This was in contrast with the former forced flow conductors, where the experiments were carried out only on full-size conductors, manufactured to wind a real magnet. The short conductors for small-scale experiments are ignored in this review, as they are mostly irrelevant for the issue of manufacturing methods.

Instead of a one-by-one description of each conductor, the manufacturing aspects are grouped into four categories, with cross-references that cannot always reflect the historical sequence.

#### Strand Coating

A coating of the strand surface may be applied to control the transverse resistance of the cable as well as the friction and

bonding at the strand contact points. The choice of coatings is broad, ranging from the bare strand surface (no coating) to insulating varnish. The design criteria for ac loss and stability (interstrand current sharing) drive the selection of the strand coating; see also the article Hysteresis and Coupling Losses in Superconductors. The coating options for  $Nb_3Sn$  strands are restricted by the requirement of withstanding the heat treatment.

The properties of metallic coatings for  $NbTi$  strands have been thoroughly investigated for accelerator cables (not forced flow conductors). The coating methods include galvanic deposition of Ni, Cr, and Zn, as well as dipping in molten SnAg alloy (35, 36). For  $NbTi$ -based CICC's, copper oxide has been used as a strand surface treatment in a SMES coil (37), and Cr plating for a prototype SMES conductor (38). The bare copper surface has been selected for the de-operated poloidal coils of the Large Helical Device (39), as well as for a 100 kA conductor of a flux pump (40). Organic-varnish-insulated strands have been used for the CICC of the DPC-U coils (41). In the  $NbTi$  CICC for the 45 T hybrid magnet (42), the drawing lubricant is intentionally not removed from the strands to enhance the contact resistance. The bare copper surface, as well as the oxidized copper surface, proved to be not satisfactory for pulsed operation (because of high interstrand coupling loss). Conductors with fully insulated strands showed severe instabilities due to their inability to redistribute the current among the strands.

In  $Nb_3Sn$ -based CICC's, an antibonding coating is highly recommended to avoid intermetallic diffusion at the strand crossovers during heat treatment. If the strands bond together without possibility of sliding during the winding process (react-and-wind), the performance degradation due to the bending strain and the interstrand coupling loss may become unacceptably high. An example of noncoated  $Nb_3Sn$  in a CICC, with very high coupling-current losses, is provided by the DPC-TJ coil (43).

A 5  $\mu$ m copper sulfide layer was originally specified, but eventually not applied, for the Westinghouse-LCT conductor (44). The antibonding properties of the copper sulfide have also been assessed in the JA development for  $Nb_3Sn$  CICC (45). In the US-DPC conductors, some lubricant was accidentally used during the cabling, providing an oil film on the C-coating (46). Based on this experience and on special ac loss experiments (47), the bare  $Nb_3Sn$  strands of the CICC for the 45 T hybrid magnet were intentionally coated with Mobil 1 oil; during

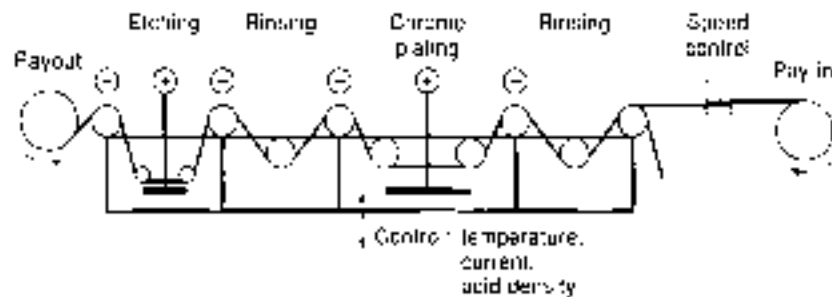


Figure 10. Scheme of a continuous Cr plating line (courtesy of ARD).

the heat treatment the organic film partly vaporizes, leaving a blackened strand surface. The organic vapors may have a detrimental effect on the other conductor components during the heat treatment, and the loose carbonized particles may affect the flow dynamics in operation.

**The Cr Electroplating.** Today, the large majority of  $Nb_3Sn$  strands for CFCs are Cr-coated in a continuous electroplating process. The plating of wire with Cr is not a standard process, and a number of new plating lines have been set up, especially for  $Nb_3Sn$  strands. Because of the environmental issues associated with Cr electrolyte, the plating lines are mostly assembled at Cr plating companies rather than at companies for continuous galvanic coating of wires.

Compared to other galvanic processes (e.g., Cu, Sn, and Ni), Cr electroplating is much slower (and more expensive too), at a current density of  $50 \text{ A/dm}^2$ , the growth rate of the Cr layer is in the range of  $1 \mu\text{m/min}$ . To obtain reasonable process speed and prevent the Cr plating from becoming bottleneck in strand production, several strands can be pulled in parallel through the same galvanic bath, or the same strand can make multiple passes through the bath. The process speed is obviously inversely proportional to the required thickness of the Cr layer. As the antibonding effect of the Cr plating is due to the surface oxide, a thin coating (in the range of  $2 \mu\text{m}$ ) is preferable (for cost reasons, to a thicker one). Typical process rates for a  $2 \mu\text{m}$  coating are in the range of  $1 \text{ m/min}$  for a single pass and can be substantially increased by multiple passes.

Prior to Cr plating, the strand needs to be cleaned, usually by galvanic etching. This step may be either integrated in the Cr plating line (46) (see e.g., Fig. 16) or done separately, together with a surface passivation (47).

The antibonding properties of the Cr plating are well established. However, the contact resistance at the strand crossovers (i.e., the interstrand coupling loss) and the worsening of the copper RRR (due to grain boundary diffusion during the heat treatment) may vary over a broad range, depending on the proprietary electroplating process of the different vendors (50). A discussion of the role of the electrolyte composition is given in Ref. 51.

### Strand Bundling

The way of bundling the strands in a CFC may have a crucial effect on the conductor's performance. For an even current distribution under time-varying current operation, the path, and thus the self-inductance, of each strand in the bundle must be identical. This requirement may also be formulated as full transposition with respect to the perpendicular applied field. Layered cables of superconducting strands, for example  $1 \times 6$

or  $1 \times 6 \times 12$  configurations, should be avoided. However, it is allowed to place a nonsuperconducting core at the center of a cable, as for example in the Swiss LCT cable (17); this option is satisfactory for transposition issues, but further investigations are required to assess the role of the central core ("segregated" copper) as a stabilizer for transient events. Among the coreless stranded bundles, the triplet is definitely the most stable geometry. With increasing number of strands (e.g., five or six), there is a tendency for one strand to slip to the center, making an imperfect, non-fully-transposed,  $1 \times 4$  or  $1 \times 5$  cable.

To bundle together a reasonable number of superconducting wires, multiple stranding (multistage cabling) is necessary. To preserve the transposition, the pitches of each cable stage must not be integer multiples. To avoid strand damage (45) and to reduce the interstrand coupling loss (52), the pitch direction must be the same in all the cable stages. The typical pitch lengths are in the range of 10 to 20 times the diameter of the cable stages. A sequence of tight pitches turns into a stiff cable with a large average angle of the strands to the conductor axis, leading to ineffective use of the cross section. On the other hand, a sequence of loose pitches may lead to strand slippage during the bundle compaction, with local transposition errors. The pitch of the last cable stage is crucial for coupling loss. In very large bundles, the subcables may be wrapped with a high-resistivity metal strip to cut the largest coupling-current loops (53). The wrap direction must be opposite to the pitch direction to avoid loosening and crinkling of the strip when cabling the next stage with back twist. The wrap may be spaced (no overlapping) to allow some coolant mass exchange. The use of the subcable wraps to cut the interstrand coupling currents must be carefully watched in the design, as it may lead to a severe penalty in interstrand current sharing and current redistribution.

The selection of the cable stage sequence, especially the number of elements in the last cable stage, is made according to the final cable shape—a rectangle, an annulus with central hole, a square bundle, or the like. Table 1 (54–59) lists a number of fully transposed cable configurations for large bundles (more than 200 strands).

The natural void fraction of a stranded cable prior to compaction is larger than 50%. However, the typical void fraction required in the cable spool is in the range of 30% to 45% (higher void fraction may allow strand movement in operation; lower void fraction may affect the integrity of the strands and limit the stability or the hydraulic performance). Compaction of the strand bundle may be achieved either with a die at the cabling point or with shaped rolls (either freely revolving or engine-driven). For large compaction loads, a die is not recommended, as abrasion may occur at the strand surface and large pulling loads may damage the cable. In large multistage cables, progressive compaction is advisable at each cable stage for an even

Table 1. Cable Pattern for Selected Large CICC's

| Conductor          | Ref | Strand                  | Coating   | Cable Pattern         | Strand<br>Diam (mm) | No. of<br>Strands | Void<br>Fraction (%) | Shape     |
|--------------------|-----|-------------------------|-----------|-----------------------|---------------------|-------------------|----------------------|-----------|
| WH-LCT             | 44  | Nb <sub>3</sub> Sn      | CuNi      | 3 × 3 × 3 / 3 / 6     | 0.7                 | 486               | 42                   | Square    |
| DPC-TJ             | 51  | Nb <sub>3</sub> Sn      | Bare      | 5 × 3 × 3 / 3 / 6     | 0.67                | 486               | 40                   | Rectangle |
| LHD-OV             | 53  | Nb <sub>3</sub> Ti      | Bare      | 5 × 3 × 3 / 3 / 6     | 0.69                | 486               | 38                   | Rectangle |
| US-DPC             | 46  | Nb <sub>3</sub> Sn      | Cr        | 5 × 3 × 5 × 5         | 0.78                | 225               | 45-56                | Square    |
| W7-X               | 56  | Nb <sub>3</sub> Ti      | Bare      | 5 × 4 × 4 × 4         | 0.55                | 192               | 37                   | Round     |
| SMES ISTEK         | 37  | Nb <sub>3</sub> Ti      | CuO       | 3 × 3 × 3 × 3 × 3 × 4 | 0.62                | 972               | 38                   | Rectangle |
| GEM                | 57  | Nb <sub>3</sub> Ti - Cu | Bare      | 3 × 5 × 5 × 4         | 0.73                | 450               | 37                   | Round     |
| ITER-MC2           | 53  | Nb <sub>3</sub> Sn - Cu | Cr        | 3 × 3 × 4 × 5 × 6     | 0.81                | 1080              | 17                   | Annulus   |
| ITER-MC1           | 53  | Nb <sub>3</sub> Sn      | Cr        | 3 × 4 × 4 × 4 × 5     | 0.81                | 1152              | 37                   | Annulus   |
| HTC-1 <sup>a</sup> | 58  | Nb <sub>3</sub> Ti      | Insulated | 3 - 3 - 3 / 3 / 6     | 1.1                 | 486               | 45                   | Rectangle |
| 45 T Hybrid-A      | 59  | Nb <sub>3</sub> Sn - Cu | Oil       | 11 × 6 / 3 × 5 / 5    | 0.43                | 525               | 39                   | Rectangle |

<sup>a</sup> Specified, but eventually not applied.

distribution of the voids and to avoid very high compressive loads at the outermost strands. The springback of the bundle after passing the compacting tool must be allowed for in setting the size of the rolls or die. The cross-section reduction of the bundle is preferably done in one step. A thin, overlapped metal wrapping is sometimes applied after compaction to protect the bundle in the subsequent handling and manufacturing processes. A final calibration-shaping pass may be necessary if tight dimensional tolerance is required.

The hydraulic diameter of a strand bundle with ~40% void fraction is in the same range as the strand diameter. To reduce the pressure drop, whenever large mass flow rate is required in operation, parallel helium channels, with larger hydraulic diameter, may be attached to the strand bundle. In the DPC-EX, the extra channels are two perforated steel pipes assembled at both sides of the flat cable (33). In the US-DPC, the four-corner area between the central, wrapped bundle and the outer square conduit provides high-speed helium channels (46).

A practical way to add a pressure-relieving helium channel is to cable the strand bundle as an annulus around a central hole. This also provides a sound, fully transposed layout for the last cable stage. The central hole can be obtained by preshaping the subbundles with rollers in the last but one cable stage, as wedges and fitting them together in the last cable stage (60). In this way, the subbundles form a vault and the central hole does not need a mechanical support. However, matching the preshaped subbundles correctly requires an accurate, continuous adjustment of the back twist of the individual components. The cabling rate is very low, and only a limited compaction load can be applied after cabling. In the ITER conductors (53) [Fig. 17(d), e], as well as in the prototype 200 kA SMES (61) [Fig. 17(b)], the QCELL conductor (62) [Fig. 17(c)], and the HT-70 (68) [Fig. 17(f)], the last cable stage is formed on a steel spiral that allows helium exchange between the central hole and the bundle. No preshaping is necessary in the subbundles of the ITER conductor, which are compacted against the spiral core, yielding a homogeneous filling of the annulus.

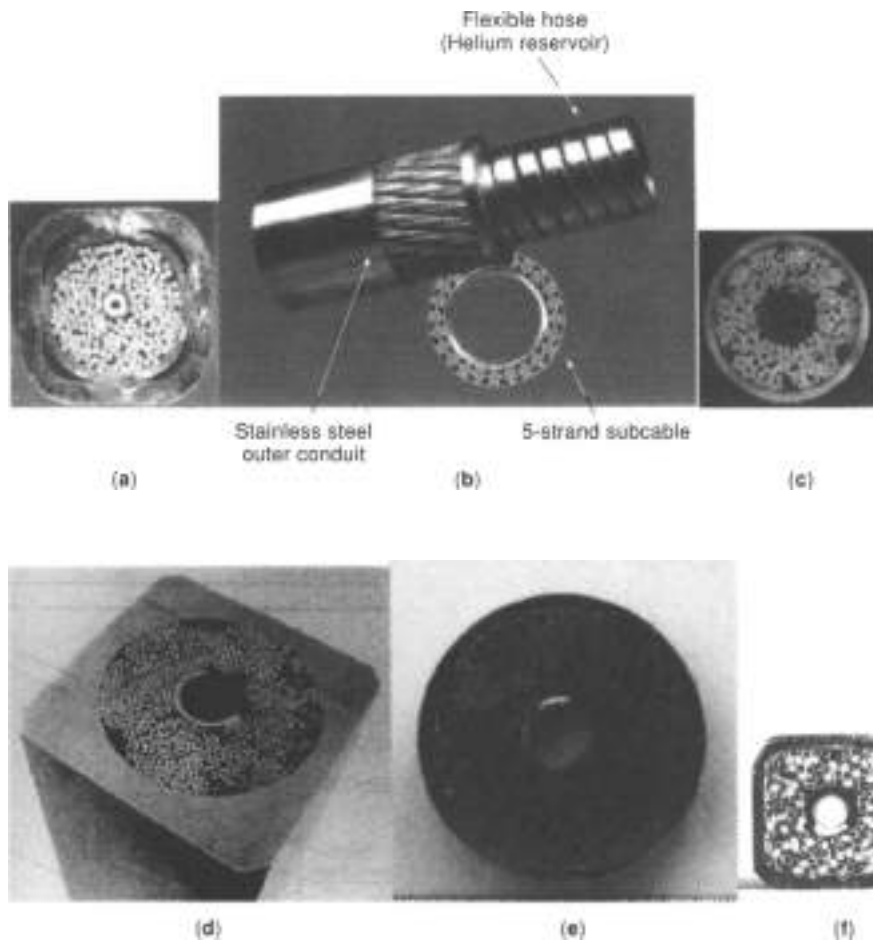
The subbundles of the Palo conductor (64) are cabled around a solid steel pipe (See Fig. 18), which forms a separate, central conduit for forced flow of two-phase helium at 0.12 MPa. The inner conduit is also wrapped with adhesive layers of Kapton and can be used in the quench detection system to balance the inductive voltage component. The annular space contains stagnant supercritical helium pressurized at 0.4 MPa. The 13

sub-bundles of six Nb<sub>3</sub>Ti strands around a CuNi core are spaced by an insulating wrap covering 70% of the surface. Neither instabilities nor ramp rate limitations have been observed in the perfectly transposed conductor, but the dc performance was degraded by 30% (65).

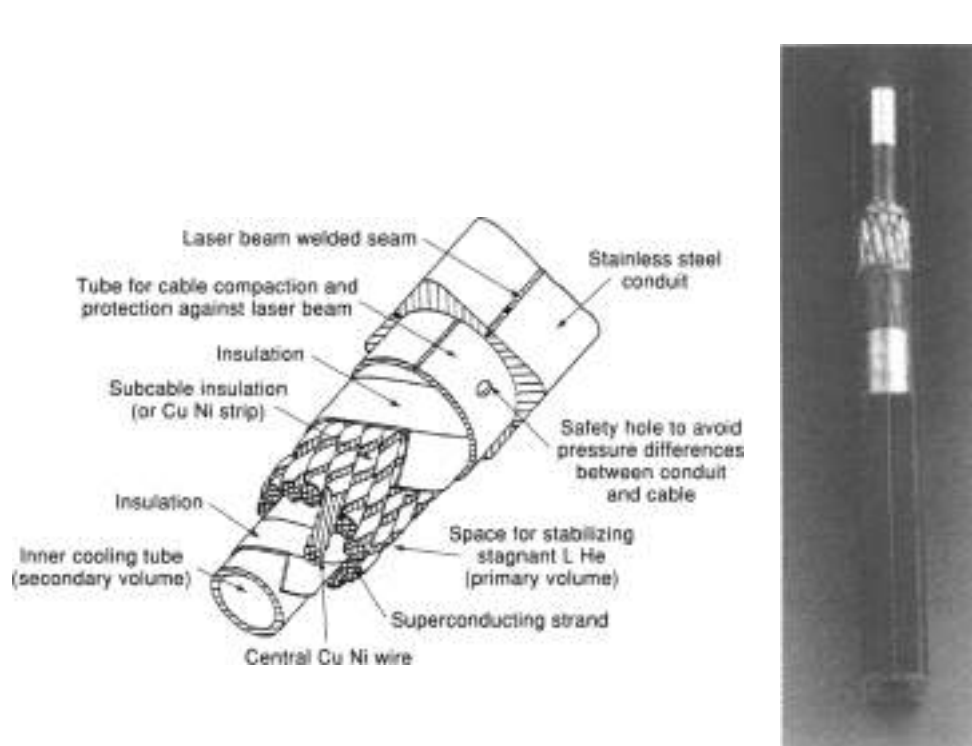
**Braided Bundles.** A powered braiding machine, able to handle 29 strands at once, was built in 1969; see Fig. 19 (48). The 29-strand lattice braid was used as a first cable stage for a rectangular Nb<sub>3</sub>Sn prototype CICC (67), with a total of 809 strands (29 × 3 × 7). Compared to a similar conductor with purely stranded cable configuration (3 × 3 × 73 × 3 × 3 × 7) and identical overall void fraction, it is evident from Fig. 20 that the braid-based conductor has a more homogeneous distribution of voids. Mechanical tests (68) also indicate a higher transverse modulus for the braided bundle because of the interlocked structure (no line contacts among strands).

Instead of multiple stranding, an alternative way to bundle together the superconducting strands for a CICC is braiding. The hollow braids in common use in electrical engineering (e.g., for high-frequency shields) consist of two interlocked multiple spirals with opposite pitch direction. Such hollow braids are transposed for perpendicular field, but are not suitable for a CICC, because the symmetry of the hollow geometry is destroyed during compaction and, in case of parallel time-varying field components, large voltage may be induced between the two opposite by running spirals. The flat braids proposed for superconducting accelerator dipoles are fully transposed for perpendicular and longitudinal applied field, but do not fit, because of the large aspect ratio, into a practical CICC.

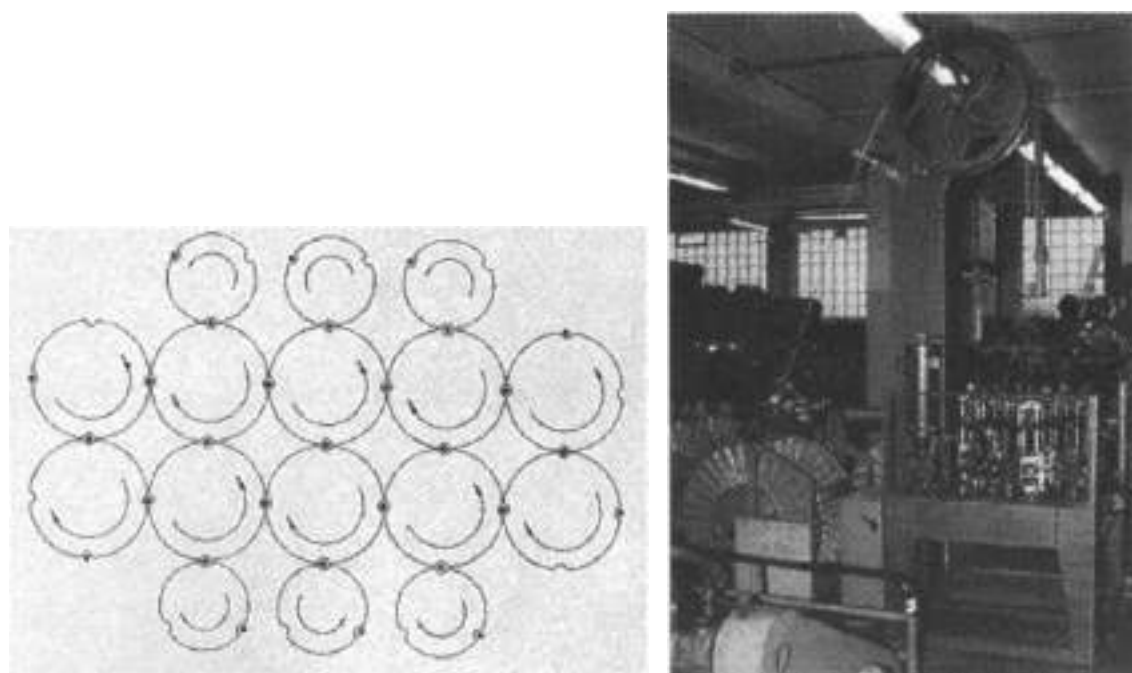
In *latticebraids*, similar to some type of alpine ropes, all the elements are interlocked in a bulk bundle. For selected layouts, the strand bundle is fully transposed for both transverse and longitudinal applied field (notice that stranded cables are never transposed for longitudinal applied field). Lattice braids were proposed as an option for the WH-LCT conductor (66). The full transposition, high flexibility, and mechanical stability of the bundle, as well as the large wetted perimeter, made very attractive the idea of a lattice braid as the first cable stage for a CICC bundle. If all the strands are braided at once into a large bundle, the aspect ratio of the braided bundle can be designed starting from the layout of the braiding platform (See Fig. 19) without the need of final shaping as with stranded bundles. The only drawback for practical application of the lattice braids is



**Figure 17.** Examples of CICCs with pressure-relieving helium channels: (a) the Nb<sub>3</sub>Sn US-DPC, 22.3 × 22.3 mm with central heater (courtesy of MIT); (b) the NbTi 200 kA SMES, 55.1 mm diam (courtesy of Bechtel); (c) the Nb<sub>3</sub>Sn QUELL CICC, 19.4 mm diam (courtesy of CRPP); (d) the Nb<sub>3</sub>Sn ITER-C5MC, 51 × 51 mm (courtesy of ITER); (e) the Nb<sub>3</sub>Al insert conductor, 45.7 mm diam (courtesy of JAEA); (f) the NbTi HT-9U conductor (from Ref. 63).



**Figure 18.** The Pils conductor, 22.5 × 22.5 mm, with double, separate cooling channels and insulated subbundle (courtesy of VACL).



**Figure 19.** The 20-strand lattice braid scheme of the braiding platform with the movements of the carriers (left) and operating braiding machine (right). (Courtesy of ABB.)

that such braiding machines are not industrially available and must be specially built. The complexity (and the cost) of the braiding machine sharply increases with the number of elements to be braided at once. For this reason, the braid option was abandoned in the WH-LCT, after a few meters of 34-strand bundle had been manually braided.

#### Conduit Materials

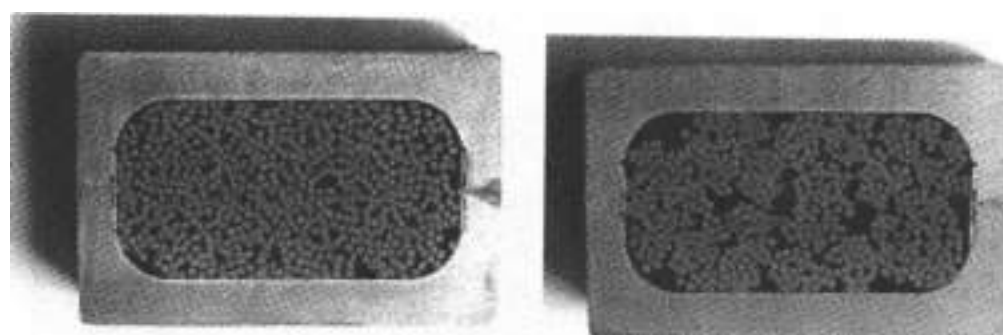
The selection of the conduit material is dictated by the mechanical requirement in operation, as well as by manufacturing issues (weldability, industrial availability of strips or pipes, and compatibility with the heat treatment, if any).

Copper is seldom used as a conduit material because of its low strength. The ETL wind-and-react  $Nb_3Sn$  flat cable (see above) has a 0.75 mm thick Cu welded sheath (29), manufactured in straight, 40 m long sections by longitudinal welding. A 200 m long section of NbTi CICC with 5/5 strands was manufactured in a continuous tube mill process, using

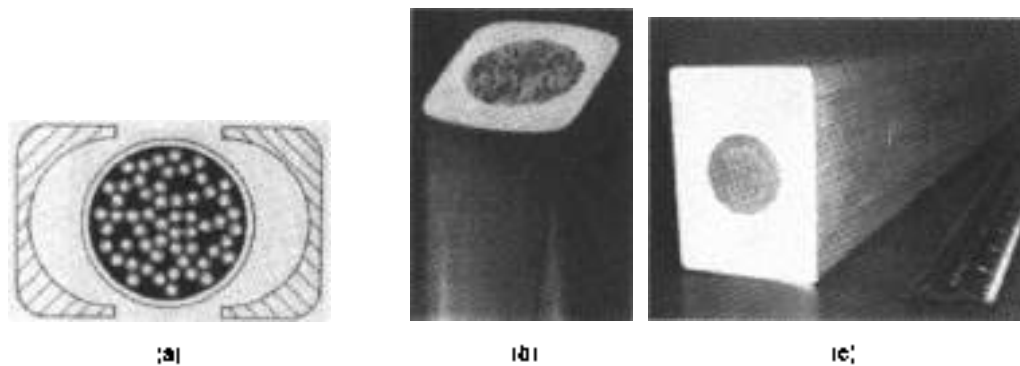
a 0.6 mm thick copper strip (69). Although the coils wound with the copper-sheathed conductors were small (inner radius 0.15 m and 0.20 m respectively), in both cases the sheath cross section was not designed to withstand alone the hoop load in operation.

Low-carbon-content, austenitic steel is the favorite material for CICC sheaths, because of its ductility, strength, broad availability in almost any shape, and low cost. Whenever higher strength is required, nitrogen-alloyed austenitic steels (e.g., AISI 316LN and AISI 304LN) may be used. Some precautions must be taken using the nitrogen-alloyed steels with  $Nb_3Sn$  conductors, due to the possible reduction of the fracture toughness after heat treatment. A large effort has been carried out, mostly in Japan, to develop and select suitable stainless steels to prevent degradation of the mechanical properties after heat treatment (70–73).

Among the relevant  $Nb_3Sn$ -based CICCs with austenitic steel conduits, we recall the 45 T-A and B hybrid (74), the DPC-FX (33), the DPC-TJ (54), the FNEA conductor (75), the ITER-



**Figure 20.** Cross sections of two prototype CICCs for NRT: braided bundle (29 × 3 × 7) (left) and stranded bundle (3 × 3 × 3 × 7) (right). (Courtesy of ABB.)



**Figure 21.** Use of Al alloys for CICC: (a) an MIT test conductor (courtesy of B. Montgomery), with Al primary conduit and co-wound profiles; (b) the W7-X conductor, with coextruded AL8060 alloy (courtesy of EM-LMD); (c) the GEM conductor with steel primary conduit and coextruded outer Al cladding (Courtesy of K. Martowatsky, LLNL).

TFMC (76), the HFTF (77), the FC-150 m coil (78), and the 12 T TFMC pinnake (79).

Most of the NbTi-based CICC's, have an austenitic steel conduit, for example, the poloidal LHD coils (66), and the DFC-U (58), the EU-LCT-24, the Italian MHD (60), the Polu conductor (64), the 45 T-C hybrid (42), the JF-30 (51), the 200 kA SMES (61), and the HT-7U (63).

**Aluminum-Based Conduit.** An Al-sheathed CICC [nontransposed cable of  $4 \times 1 + 6 \times 121 + 6 \times 12$  NbTi strands] was manufactured in short lengths in 1977–1978 (82). The 0.9 mm thick Al round pipe, 12 mm diam, was co-wound with two matching Al extruded and anodized profiles; see Fig. 21. The test conductor was bent to small radius and used for stability experiments.

For the W7-X coil system, consisting of nonplanar windings with small bending radius, a NbTi based CICC was proposed (83) with conduit made from a coextruded Al alloy; as extruded, the CICC is soft ( $\sigma_y \approx 100$  MPa) and can be easily bent to small radius. After aging at 130 to 150 C, which can be combined with the epoxy impregnation process, the Al 6060 alloy hardens to  $\sigma_y \approx 250$  MPa at 4 K, providing the required mechanical stiffness to the winding (86). The square conduit (14.8 × 14.8 mm) (see Fig. 21), is coextruded on the 10 mm diam cable, without any protective wrapping, in unit lengths of 600 m. To avoid degradation of the NbTi conductor performance, the temperature and speed of the coextrusion process must be carefully controlled; a decrease by 5.5% of the strand  $I_c$  performance was observed for the W7-X conductor (84). The coextrusion of superconducting cables and pure Al is well developed for particle detector magnets. For W7-X, the process was first used to produce an internally cooled conductor.

A coextrusion method was also used for the GEM detector conductor (85). However, here the cable (see Table 1) is primarily sheathed with a 1.5 mm thick 304L steel conduit and the thick pure Al cladding has no helium containment function. The bonding of Al to steel was proved to be effective for stability. With an overall size of  $49.8 \times 58.5$  mm (847) consisting of the pure Al cladding, the GEM conductor is the largest forced flow conductor manufactured in useful length; see Fig. 21.

**Incoloy 908 Conduit.** The use of a superalloy as a conduit material for CICC was first considered for the Nb<sub>3</sub>Sn WH-LCT conductor, due to the disappointing experience after heat treatment with the nitronic 40 (a steel with 0.33% nitrogen

content). The R&D work (86) led eventually to the choice of JBK-75, a 30% Ni alloy specially developed from the A-286 superalloy, as a sheath material.

The main drawback of the steel as conduit material for Nb<sub>3</sub>Sn CICC's is its large coefficient of thermal expansion: from the heat treatment down to the operating temperature, the differential contraction of steel and Nb<sub>3</sub>Sn causes a degradation of the superconducting properties, especially severe at high field. In the scope of the US-DFC coil R&D program, a new superalloy of the Incoloy group was developed as sheath material for Nb<sub>3</sub>Sn CICC's, with coefficient of thermal expansion matching that of Nb<sub>3</sub>Sn (87). Several test programs have been carried out to compare the mechanical properties of Incoloy 908 and 316 LN steel (see, e.g., Ref. 86). However, the most impressive advantage in the use of Incoloy 908 as a conduit material is the gain in the high-field performance of the Nb<sub>3</sub>Sn (89). On the other hand, a big concern in the use of Incoloy 908 is the possible occurrence of *stress - accelerated grain boundary oxidation* (SAGBO) during the heat treatment, leading eventually to catastrophic failure of the material. Relaxation of the surface stress in the conduit by shot peening, as well as tight control of the oxygen content (below 1 ppm) in the heat treatment atmosphere, are the key measures to avoid SAGBO (90, 91).

The US-DFC conductor was sheathed with a 2.35 mm thick strip of Incoloy 908, TIG-welded in a tube mill (46). The first large-scale application of Incoloy 908 as a conduit material for a Nb<sub>3</sub>Sn CICC is ITER-CS model coil conductor (53). Several tens of tonnes of round-in-square tubing, as 6 m to 11 m long sections, have been manufactured by combined hot, hollow extrusion and cold drawing (92).

Pure titanium has been proposed as an alternative low-coefficient-of-thermal-expansion material for sheathing of Nb<sub>3</sub>Sn-based CICC's. Its low-temperature mechanical properties after aging are reported in Ref. 93. The only important CICC with Ti conduit is the 160 m long QUELL conductor, with  $3 \times 3 \times 4 \times 6$  Nb<sub>3</sub>Sn strands and 1.2 mm thick pure Ti round pipe (62).

### Jacketing Methods

The selection of the jacketing method depends to a large extent on the dimension and cross section of the jacket. A key issue affecting the jacketing is the wall thickness. Conduits with constant wall thickness can be manufactured and assembled by a



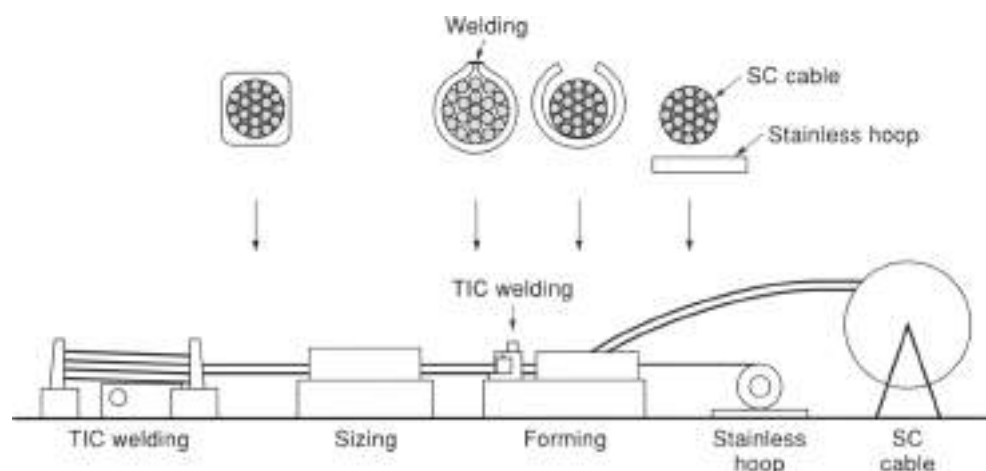


Figure 22. Scheme of the tube mill method for CICC, courtesy of JAFRI.

variety of methods, some of which are the cheapest. On the other hand, a conduit with variable wall thickness at the corners allows better packing of the conductor in the winding and offers superior mechanical performance.

**The Tube Mill.** The large majority of CICCs are jacketed by the tube mill method. The generic scheme of the process is shown in Fig. 22, from Ref. 54. The metal strip is formed by rolls around the cable in a large number of steps, see Fig. 23(a) for the WB-LCT conductor (94), Fig. 23(b) for the DPC-C (58), and Fig. 23(c) for the 45 T hybrid (59). The longitudinal TIG weld is usually done while the conduit is oversize, to limit the thermal load on the strand bundle underneath the welding seam and to avoid contact between the weld bead and the bundle. However, it is also possible to fold the strip exactly to the final size and weld it without damaging the superconductor, as for example in the DPC-EX (33) and the EU-LCT (24).

After forming and welding the round pipe (a highly standardized industrial process), the conduit is compacted in a number of steps to suppress the gap between pipe and bundle, and to give the square or rectangular shape required for winding, see Fig. 23. Typical conduit elongation during the

compaction process, by rolling and/or drawing, is up to 10%. It is crucial to control the cable elongation during the compaction at large cross section reduction and low void fraction, plastic axial deformation of the bundle may start, with fatal consequences for the stranded cable, which does not withstand elongation larger than 2% to 3% without permanent damage. On the other hand, good mechanical engagement between cable and conduit during the final shaping steps is highly desirable, to restrain dangerous strand movements under the operating load.

Starting from a metal strip, it is not possible to achieve sharp corners in the final, squared conductor. The typical range for the outer corner radius is 3 mm to 6 mm, depending on the wall thickness. The final cold work of the conduit (starting from the fully annealed strip) depends on the cross-section reduction, the conduit thickness, and the corner radius.

The thickness of the conduit (typically 1.5 mm to 2.5 mm) ranges from 0.8 mm (EU-LCT) up to 3.5 mm (LHD-OV). The tube mill method is used to form conduit from a variety of metals, including copper, titanium, Inconel 908, and austenitic steels of different composition. A list of selected CICC's sheathed by a tube mill is given in Table 2 (95). Some

Table 2. A Selection of CICC's Jacketed by the Tube Mill Method

| Conductor                       | Ref. | Outer Size (mm) | Wall Thickness (mm) | Conduit Material  |
|---------------------------------|------|-----------------|---------------------|-------------------|
| WB-LCT                          | 44   | 20.7 × 20.7     | 1.73                | Superalloy—JRK-75 |
| DPC-1J                          | 54   | 17 × 22.5       | 1.0                 | Steel—316 L       |
| LHD-OV                          | 55   | 27.5 × 31.6     | 3.6                 | Steel—316 L       |
| LHD-IV/IS                       | 55   | 23.0 × 27.6     | 3.0                 | Steel—316 L       |
| SMES—200 kA                     | 61   | 51 diam.        | 1.25                | Steel—304 LN      |
| SMES ISTEK                      | 37   | 25.4 × 27.6     | 2.3                 | Steel—316 L       |
| US-JPC                          | 46   | 22.3 × 22.3     | 2.45                | Inconel 908       |
| JP-33                           | 81   | 35 × 35         | 2.0                 | Steel—??          |
| HT-7U                           | 63   | 17.4 × 17.4     | 1.5                 | Steel—316 L       |
| EU-LCT                          | 23   | 10 × 40.0       | 0.6                 | Steel—316 LN      |
| DPC-EX                          | 33   | 9.4 × 40.0      | 1.5                 | Steel—JRK         |
| HFFP                            | 77   | 20.7 × 20.7     | 1.7                 | Steel—Nicro 46    |
| [TER—Nb <sub>3</sub> Al] insert | 95   | 45.7 diam.      | 2.0                 | Steel—Mod. JN1    |
| 12 T TF pancake                 | 78   | 15.7 × 14.3     | 1.0                 | Steel—316 L       |
| DPC-C                           | 58   | 29 × 35.5       | 2.0                 | Steel—JN1         |
| 45 T hybrid—A                   | 59   | 16.2 × 13.7     | 1.65                | Steel—316 LN      |
| 45 T hybrid—C                   | 69   | 15.4 × 13.5     | 2.0                 | Steel—316 LN      |

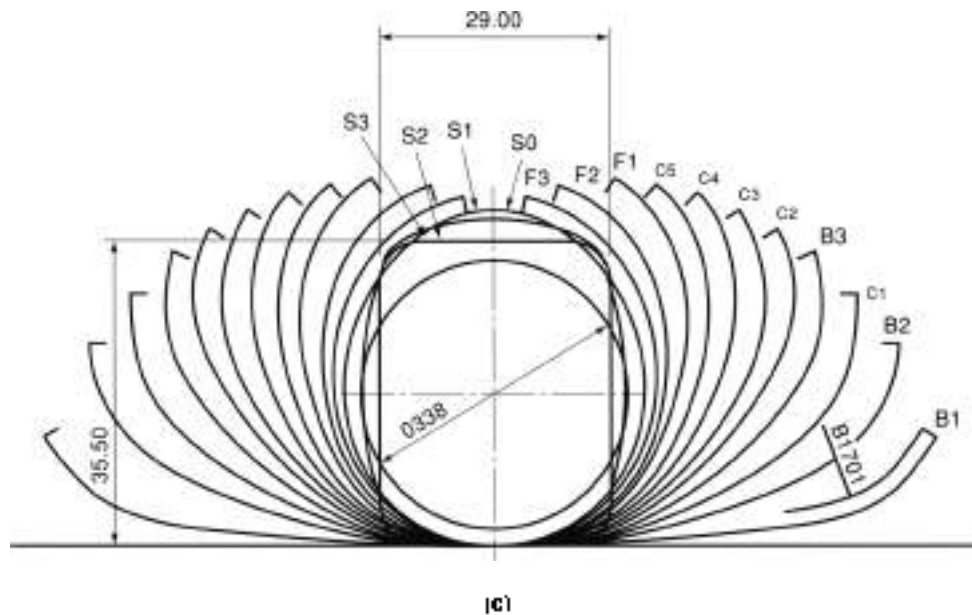
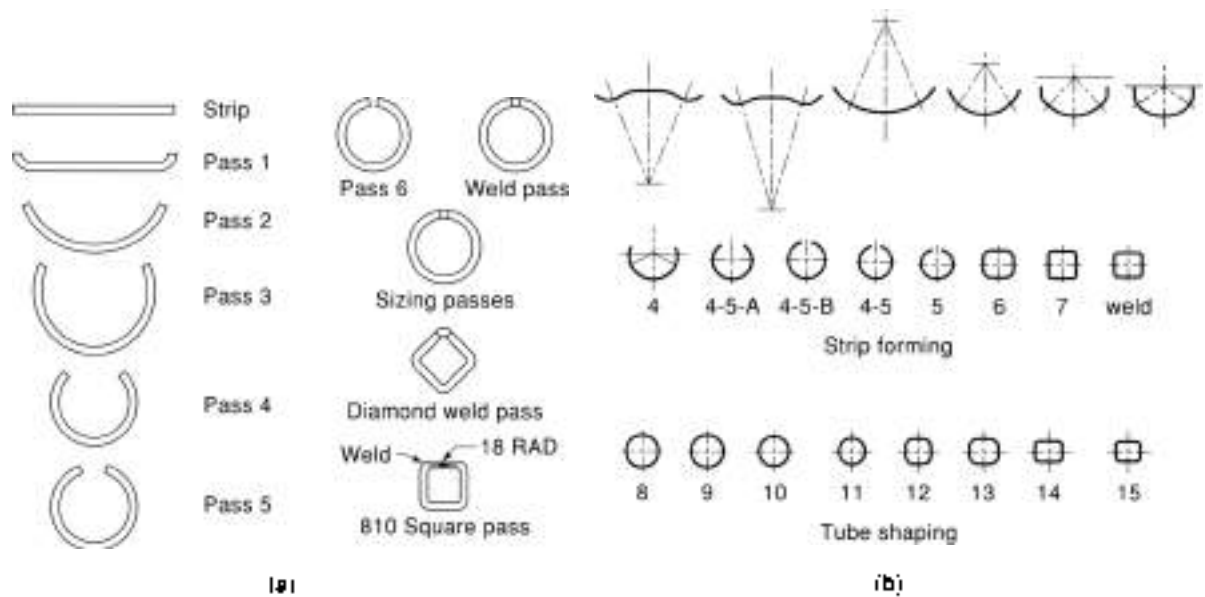


Figure 23. The forming steps for the metal strip in a tube mill: (a) the WH-LCF (courtesy of P. Sanger); (b) the 45 T hybrid (courtesy of J. Miller); (c) the DFC-L (courtesy of JAEHI).

cross sections of CFCs jacketed in tube mill are shown in Fig. 24.

Quality control of the weld (59) is obtained through tight monitoring of the process parameters. The presence of the strand bundle underneath the weld seam precludes a check by X rays or eddy currents after unjacketing and shaping. Stop and restart of the process (e.g., to replace the TIG electrode and the gas bottles) is common practice for the tube mill method.

**The Pullthrough Method.** In the pullthrough method (see scheme in Fig. 25) the sheath is first manufactured as a straight, oversize conduit of the same length of the final conductor section. In a second step, the bundle, wrapped with a protective metal foil, is pulled through the oversize conduit. Eventually, the conduit-cable assembly is compacted by rolling

or drawing and coiled on a take-up reel.

The pullthrough method can be applied to any conduit geometry, however, it is especially attractive for thick-walled conduit and conduit with nonconstant wall thickness (which cannot be welded in a tube mill). Compared to the tube mill, an advantage of the pullthrough method is that the welds of the conduit can be fully inspected (and repaired) before assembling the superconductor.

The straight, oversize conduit can be assembled by different methods, including longitudinal welding of strips or pipes. However, all the pullthrough applications to date have used the butt welding of extruded tubing sections (typical unit length 6 m to 12 m). In one case (96), the 1 mm thick Incoloy 908 pipe was first manufactured in a tube mill, then cut in 10 m sections, and eventually reassembled by butt welding at

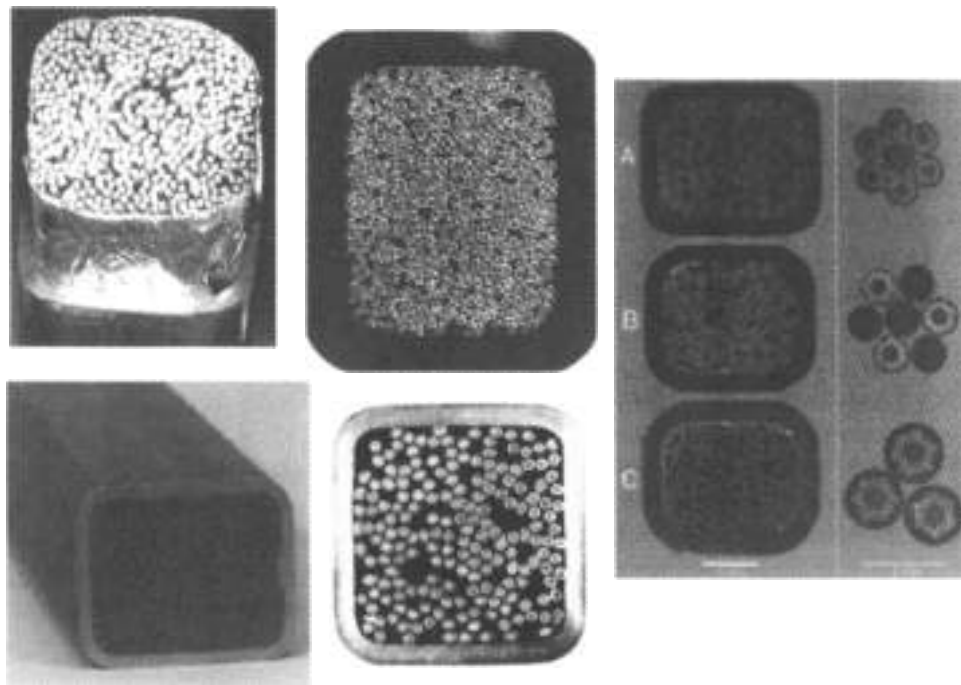


Figure 24. Examples of CICCs packed in a tube mill. Top left, the WJ-LCT-94; top middle, the LED IV (courtesy of NIFS); bottom left, the DPL-U (courtesy of JARRI), bottom middle, the JP-30 (courtesy of JARRI); right, the 45 T hybrid conductors (courtesy of NHMFL)

the jacketing site. The pullthrough method was first applied for short conductor length and prototype conductors (e.g., Refs. 30, 60). Table 3 reports the pullthrough jacketing applications with useful 1–50 m conductor length. Except for the ITER-CS conductor, the conduits quoted in Table 3 have constant wall thickness.

The butt welding of the thin-wall conduits ( $\leq 2$  mm) is done with a TIG single-pass orbital welder without filler. A second pass with filler was necessary for the 3 mm thick conduit of GEM (85). For the circle-in-square Inconel tubing of ITER-CS,

a number of orbital passes with filler were carried out after the first root pass. Eventually, the corner area was filled manually. Special attention has to be paid to the weld protrusions (drops) inside the conduit: these can be monitored either by an endoscope or by inserting a calibrated gauge. The quality assurance on the butt welds may include X rays, ultrasound, dye penetrant, and leak tightness.

The applicable pulling load,  $F_p$ , is limited by the yield strength of the cable (i.e., the threshold  $\sigma_y$  where strand damage occurs) and by the strength of the conduit, which must

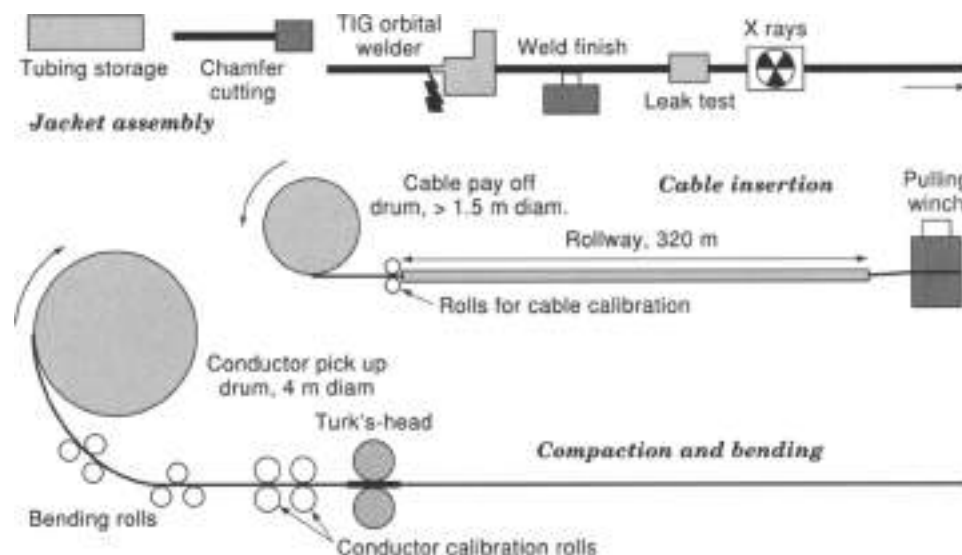


Figure 25. Scheme of a jacketing process according to the pullthrough method.

Table 3. Long CICC's Jacketed by the Pullthrough Method

| Conductor  | Ref. | Outer Size (mm) | Wall Thickness (mm) | Largest Unit Length (m) | Conduit Material |
|------------|------|-----------------|---------------------|-------------------------|------------------|
| ITER-CS    | 57   | 51 × 51         | 4-16                | 200                     | Inconel 908      |
| ITER-Dummy | 96   | 39.5 diam       | 1                   | 850                     | Inconel 908      |
| ITER TF    | 76   | 40.7 diam       | 1.6                 | 160                     | Steel—316 LN     |
| GEM*       | 65   | 26.08 diam      | 3.05                | 75                      | Steel—304 L      |
| ENEA-12T   | 75   | 13.8 × 13.8     | 1.3                 | 260                     | Steel—316 LN     |
| MHD—Italy  | 88   | 16.3 × 16.3     | 2                   | 340                     | Steel—316 LN     |

\* Without the outer Al cladding.

\* The entire conductor is manufactured with sections up to 400 m long for KSTAR model coil.

react the pulling load at the insertion point (this is an issue only for thin-walled jackets). The pulling load per unit length has been found (96) to be a function of the clearance between the cable and the oversize conduit. The coefficient of friction,  $\nu$ , defined as the pulling load divided by the cable weight can also be monitored. There is broad agreement (53, 75, 76, 96) that the insertion clearance must be larger than 1.5 mm for friction coefficients  $\nu < 1$ . However, according to the actual tolerances on the size of cable and conduit, a larger clearance may be necessary. The longer the conductor section to be jacketed, the more crucial becomes the issue of clearance. For short length, the risk of higher friction may be accepted. For long section jacketing, high initial friction may lead to an avalanche effect and the cable may eventually get stuck in the conduit (76).

After pullthrough, the conductor is compacted by rolling or drawing to suppress the insertion clearance. For conduits with nonconstant wall thickness, large cross-section reduction should be avoided to keep tight tolerance in the final conductor geometry. On the other hand, the compaction must achieve a positive engagement of cable and conduit to restrain strand movement in operation, which may result in dangerous instabilities.

The jacketing of long sections of CICC by the pullthrough method has been successfully demonstrated at VNIIEP, Moscow, where a 1 km long jacketing line has been set up and operated (96). The maximum conductor length,  $L_{max}$ , that can in principle be jacketed by the pullthrough method can be assessed by imposing the constraint that the pulling load at the head of the cable does not exceed the tensile cable strength where strand damage occurs,  $\sigma_A$ :

$$L_{max} \leq \frac{\sigma_A}{\gamma}$$

where  $\gamma$  is the density of the cable, typically in the range of  $7 \times 10^4 \text{ kg/m}^3$  to  $8 \times 10^5 \text{ kg/m}^3$ . The range of  $\sigma_A$  depends on the cable pattern and strand properties. For the friction coefficient  $\nu$ , a safety margin must be considered. Assuming, conservatively,  $\sigma_A = 50 \text{ MPa}$  and  $\nu \leq 1$ , the maximum allowed length would be about 700 m.

**Laser Beam Welding (CO).** Laser beam welding was proposed in 1986 (97) as an effective method to form thick conduit by longitudinal welding. The laser beam allows deep penetration and continuous joining with negligible deformation of thick-walled conduits in a single pass, without filler. Due to the small heat load on the underlying cable, laser welding can be applied to the

final conductor size without the final compaction and shaping steps, which are mandatory in the pullthrough method and in most tube mills to suppress the gaps between cable and welded conduit. Compared to the electron beam, also used for narrow welding of thick walls, the laser beam is most suitable for continuous processes, as it does not need a vacuum environment.

Conduit segments with variable wall thickness can be manufactured by hot rolling and cold drawing in sections up to about 20 m, butt-welded and coiled to the required conductor length. Two longitudinal welds are required to join the jacket segments, with the cable fed in between, as in the tube mill. This can be obtained either by two synchronized laser beams or by a single, optically split beam. For a process rate in the range of 2 m/min, the required beam power is about 1 kW per millimeter of wall thickness. The penetration depth of the seam can be controlled in practice to  $\pm 10\%$  (mostly due to the plasma oscillations in the weld). To guarantee full penetration of the weld and avoid contact of the strand bundle with the laser beam, protection underneath the weld is mandatory: this can be either a wrap on the cable or a longitudinal strip or gap under the seam.

Longitudinal laser beam welding has been applied only to two conductors: Polo (64) and the NET-ABB prototype conductors (48). The jacket of Polo is built from four quadrant profiles (see Fig. 18), preassembled by laser welding into two U profiles and eventually welded by two synchronized beams on the wrapped cable. The weld seam is 1.5 mm thick. Four conductor sections, each 150 m long, have been produced. The actual welding time for one section was less than 1 h, with several stops and restarts (97). Eddy currents were used to check the quality of the welds. Leaking spots were repaired by TIG welding.

In the NET-ABB prototype conductor (see Fig. 20) the weld seam is 3.9 mm thick. An air gap and a longitudinal steel strip underneath the weld protect the strand bundle from the contact with the beam. Quality assurance by ultrasound can be applied online in the process.

Despite the satisfactory employment and attractive features of laser beam welding (high rate, no deformation, minimum heat load for large penetration depth), no further application of it has been reported for CICC jacketing in the last decade. The main reasons are the investment cost for the laser device, which can be justified only for large series production, and the limited experience of industrial conductor suppliers with the laser welding technology (a similar situation to that for lattice braiding; see above).

**The Coextrusion Method.** Because of the need for temperature compatibility, only aluminum alloys (see above) can be used to jacket a CICC by the coextrusion method. For the W7-X

156), the square, coextruded conduit fulfills at once the functions of helium containment, structural support, and stabilization (see Fig. 21). In the case of the GEM conductor (185), the pure Al cladding is coextruded on the steel jacketed conductor and has only the function of limiting the hot spot temperature in case of quench.

**Double Conduit.** To avoid the problem of assembling a thick-walled jacket, a *double conduit* has been used in a few cases. The additional conduit has no helium containment function, but its cross section adds to the main conduit from the structural point of view.

The secondary conduit may be inside the main conduit, as in the US-DPC (see Fig. 17(a)), where a 1 mm thick Incoy 806 strip is formed in a tube mill and welded (not helium-tight) on the round cable (46). The inner conduit stabilizes the shape of the cable, defines the void fraction, and protects the cable during the welding of the thicker, outer conduit. During the development of the Palo conductor (64) (Fig. 18), an inner conduit with punched holes was initially foreseen as a protection for the cable during the outer conduit welding (the inner conduit was later replaced by a double wrap).

An outer secondary conduit can be used to give a structural reinforcement and provide a better shape for winding. This possibility is of special interest when the primary conduit is made in a tube mill, with constant wall thickness and large corner radii. In the MIT test conductor (82), both primary and secondary conduits are aluminum (see Fig. 21); the two extruded U profiles are fitted to the round NbTi conductor without welding. In the JPC TJ 199 (see Fig. 15), the Nb-Sn conductor, jacketed in a tube mill, pancake-wound, and heat treated, is eventually encased and spot welded into two steel U profiles (preformed armor). The thin-walled, sharp-cornered U profiles are manufactured by S-D rolling a spiral into a thick steel plate, in order to fit the shape of the heat-treated pancake.

## CHOICE OF MANUFACTURING METHODS

The conductor layout, and hence the manufacturing methods, are to large extent dictated by design considerations. However, the cost and the reliability of the manufacturing process should also be taken into account in the selection of the conductor layout.

Cable-in-conduit conductors are today more popular than soldered conductors, although many applications actually do not require the special features of the CICC. From the manufacturing point of view, a soldered conductor (e.g., a layer of NbTi strands cabled and soldered on a square copper pipe) is preferable and cheaper for low operating current and overall cross sections smaller than 60 mm<sup>2</sup>. On the other hand, a CICC is more effective for high current and cross sections larger than 100 mm<sup>2</sup>.

For cable-in-conduit conductors, a strand coating is mandatory whenever low ac loss is required, but may be omitted for de-superated conductors. The electroplated Cr coating is the only practical choice for Nb<sub>3</sub>Sn strands. More options are available for NbTi strands, including hot dipping in low-melting alloys and galvanic coating with Ni or Cr.

The requirement of transposition is applied currently to all the strand bundles for CICC's. However, nontransposed cables have been successfully operated in dc mode. An attractive, cost-

saving option is to cable a layer of strands on a central copper core as the first cable stage, reducing the copper ratio in the strands and hence the overall mass of superconducting strand, and saving one cabling stage compared to the multiple-triplets pattern. Lattice braids have proved to be an elegant and effective method for strand bundling, but require the use of dedicated machinery.

The most popular conduit material for CICC is low-carbon stainless steel (nitrogen-alloyed, whenever required), readily available in any shape, easy to weld, and cheap. For Nb<sub>3</sub>Sn conductors, the Incoy 908, due to the matching coefficient of thermal expansion, offers a superior performance at high field and competing mechanical properties. The higher cost, the precautions to be taken at welding and heat treating, and the single-source supplier to some extent balance the advantages of the Incoy 908 over stainless steel.

The tube mill is a highly developed industrial process for pipe production. For CICC's with thin, constant-thickness conduits, the tube mill is the cheapest, easiest jacketing method, with virtually no limit on the conductor length. For thick-walled conduit, as well as for nonconstant thickness, the tube mill and TIG welding can be replaced by laser beam longitudinal welding of preassembled profiles. The pullthrough method, definitely preferable for prototypes and short-length jacketing because of the small investment and the low risk, has become unexpectedly popular also for long length and series production, where its competitiveness is questionable.

Three decades of forced flow superconductors show clearly a trend toward easy, low-technology manufacturing methods despite the successful demonstrations of sophisticated layouts and advanced technology (e.g., EU-LHT and Palo). Initially, the research institutes were the leaders in conductor development. Today, many companies can supply forced flow superconductors without the support of design and R&D activities from the lab community.

## BIBLIOGRAPHY

1. H. H. Kuhn, A closed loop cooling system for superconducting bubble chamber magnets, *Proc. Int. Symp. Magn. Technol.*, Stanford, CA, 1965, p. 811.
2. M. Murpurgo, The design of the superconducting magnet for the "OMEGA" project, *Part. Accel.*, 1: 255, 1970.
3. M. Murpurgo, Construction of a superconducting test cell cooled by helium forced circulation, CERN Report 68-17, 1968.
4. M. Murpurgo, A superconducting solenoid cooled by forced circulation of supercritical helium, CERN Report 69-26, 1969.
5. C. Lesmond, J. C. Lottin, and S. Shimamoto, Experiment with hollow conductor superconducting magnet, *Proc. Magn. Technol. Conf.*, Hamburg, 1970, Vol. 3, p. 925.
6. G. Meyer and R. Maix, Superconductors and superconducting magnets, *Brown Bull. Rev.*, 57: 7, 91-355, 1970.
7. Y. Nagamura et al., The manufacture of hollow superconductors, *Proc. ICEC*, Kyoto, 1974, Vol. 8, p. 508.
8. V. N. Agareev et al., Electroplated stabilized multistrand superconductor, *IEEE Trans. Magn.*, 11: 309, 1975.
9. D. E. Ivanov et al., Test results of "tokamak-7" superconducting magnet system (SMS) sections, *IEEE Trans. Magn.*, 15: 550, 1979.
10. E. Yu. Klimenko et al., Superconducting conductor for T-15 toroidal magnet, *Sov. At. Energy (Eng. Transl.)*, 68: 258, 1977.
11. N. Schuerli, Superconductors for the magnet coils of the Omega

- spark chamber at CERN, *Revue Suisse Rev.* **50** (2,3): 75, 1972.
12. M. Margurio, A large superconducting dipole cooled by forced circulation of two phase helium, *Cryogenics*, **18**: 411, 1979.
  13. H. Jilkaert, Fabrication technology of superconducting materials, in S. Foner and R. B. Schwartz, eds., *Superconductor Materials Science*, Ser. 888, New York: 1981. NATO Adv. Study Inst., 1981.
  14. G. Paolotti et al., SULTAN: An 8 T, 1 m bore test facility, the outer solenoid, *IEEE Trans. Magn.*, **17**: 3007, 1981.
  15. V. Trinke et al., A prototype coil for the superconducting separated sector cyclotron SuSe, *J. Phys. C: Suppl.*, **1**: 217, 1984.
  16. K. Agatsuma et al., Braided multifilamentary Nb<sub>3</sub>Sn hollow superconductor and its magnet, *IEEE Trans. Magn.*, **16**: 787, 1979.
  17. H. Herz et al., The conductor for the Swiss-LCT coil, *IEEE Trans. Magn.*, **17**: 2213, 1981.
  18. H. Herz et al., Design and manufacture of the conductor for the Swiss LCT coil, *IEEE Trans. Magn.*, **10**: 711, 1980.
  19. W. M. P. Franken et al., Manufacture of the hollow superconducting He cooled conductor for the EUN-SULTAN project, *IEEE Trans. Magn.*, **10**: 368, 1980.
  20. B. Jakob and G. Pasztor, Fabrication of a high current Nb<sub>3</sub>Sn forced flow conductor for the 12 Tesla SULTAN test facility, *IEEE Trans. Magn.*, **23**: 914, 1987.
  21. J. A. Roederink et al., Design and construction of the EUN 12 T niobium tin magnet insert for the SULTAN facility, *IEEE Trans. Magn.*, **24**: 1420, 1988.
  22. R. Jakob, G. Pasztor, and R.G. Schindler, Fabrication of high current Nb<sub>3</sub>Sn forced flow conductors and coils for the SULTAN III test facility, *Fusion Technology 1992*, Amsterdam: Elsevier, 1993, p. 872.
  23. H. Kraeth et al., Manufacturing and testing of a forced flow cooled superconductor for tokamak magnets, *IEEE Trans. Magn.*, **17**: 918, 1981.
  24. H. Kraeth et al., Development and testing of a forced flow cooled superconductor for LCT, *Proc. 8th Symp. Eng. Probl. Fusion Res.*, 1979, p. 1451.
  25. B. Jakob et al., Design and fabrication of a 17 kA preprototype Nb<sub>3</sub>Sn conductor for the TF coils of the NET fusion project, *IEEE Trans. Magn.*, **24**: 1437, 1988.
  26. J. M. Platon et al., Development toroidal field conductor for NET, *IEEE Trans. Magn.*, **24**: 1435, 1988.
  27. K. Fläkiger et al., The Nb<sub>3</sub>Sn react and wind conductor for NET toroidal field coils and its boundaries, *Fusion Technology 1990*, Amsterdam: Elsevier, 1991, p. 1584.
  28. V. A. Glukhikh, Programme of the conductor development for the ITER toroidal field coils, *Proc. Magn. Technol. Conf.*, Tsukuba, 1989, Vol. 11, p. 896.
  29. K. Agatsuma et al., Fabrication and test of a forced cooled Nb<sub>3</sub>Sn superconducting coil, *IEEE Trans. Magn.*, **21**: 1040, 1985.
  30. K. Agatsuma et al., Stainless steel sheathed forced flow fully cooled Nb<sub>3</sub>Sn superconductor and its coil test, *IEEE Trans. Magn.*, **23**: 1535, 1987.
  31. T. Ando et al., -2T test coil from TMC-II for the first test program, *Proc. 10th Symp. Fusion Eng.*, Philadelphia, 1983, p. 1346.
  32. M. Sugama et al., Development of hollow cooling monolithic conductor for ITER TF coil, *IEEE Trans. Magn.*, **28**: 216, 1992.
  33. T. Ando et al., Fabrication and test of the Nb<sub>3</sub>Sn dema poloidal coil (DPC-EX), *Fusion Technology 1990*, Amsterdam: Elsevier, 1991, p. 243.
  34. H. Nakajima et al., Tensile properties of new cryogenic steels as conduit material for forced flow superconductors at 4 K, *Adv. Cryog. Eng. Mater.*, **24**: 113, 1988.
  35. M. D. Sumption et al., Contact resistance and cable loss measurements of coated strands and cables wound from Chem, *IEEE Trans. Appl. Supercond.*, **5**: 692, 1995.
  36. J. M. Depunt et al., Examination of contacts between strands by electrical measurement and topographical analysis, *IEEE Trans. Appl. Supercond.*, **7**: 793, 1997.
  37. T. Sataw et al., Present status of 480 MJ/40 MW SMES development project, *Proc. Int. Conf. Electr. Eng.*, Matsue, Japan, 1997.
  38. K. Iwasawa, A. Sultan, and S. Al-Wakeel, AC losses of a 10 kA NbTi cable-in-conduit superconductor for SMES application, *Cryogenics*, **36**: 27, 1996.
  39. K. Takahata et al., Stability tests of the NbTi cable in conduit superconductor with bare strands for demonstration of the large helical device poloidal field coils, *IEEE Trans. Magn.*, **30**: 1705, 1994.
  40. G. Pasztor et al., Design fabrication and testing of a 100 kA superconducting transformer for the SULTAN test facility, *Proc. Magn. Technol. Conf.*, Beijing, 1997, Science Press, 1998, Vol. 15, p. 639.
  41. Y. Takahashi et al., Experimental results of stability and current sharing of NbTi cable-in-conduit conductors for the poloidal field coils, *IEEE Trans. Appl. Supercond.*, **3**: 810, 1993.
  42. J. E. C. Williams et al., The development of a NbTi cable-in-conduit coil for a 4.5 T hybrid magnet, *IEEE Trans. Magn.*, **22**: 1633, 1986.
  43. K. Okuno et al., AC loss performance of 1 m bore, large-current Nb<sub>3</sub>Sn superconducting coils in JABRI demo poloidal coil project, *IEEE Trans. Appl. Supercond.*, **3**: 802, 1993.
  44. C. J. Hayes et al., Westinghouse design of a forced flow Nb<sub>3</sub>Sn test coil for the large coil program, *Proc. 5th Symp. Eng. Probl. Fusion Res.*, San Francisco, 1979, p. 1146.
  45. N. Aoki et al., Development of forced cooled Nb<sub>3</sub>Sn bundle conductor, *IEEE Trans. Magn.*, **19**: 735, 1983.
  46. M. M. Stevens et al., The US demonstration poloidal coil, *IEEE Trans. Magn.*, **27**: 2969, 1991.
  47. B. J. P. Brudnovy et al., AC loss measurements of the 4.5 T hybrid CIC conductor, *IEEE Trans. Appl. Supercond.*, **6**: 699, 1995.
  48. P. Bruzzone, Fabrication of a short length of wind and react conductor, ASEA Brown Boveri Rep. HIM 20420, Zurich, 1990.
  49. Y. Iguchi et al., Galvanic chromic plating of copper wire for the ITER program, *Proc. ICCC-ICMC 96*, Kitakyushu, Japan: Elsevier, 1996, 1996.
  50. P. Bruzzone, A. Nyhuss, and H. H. J. ten Kate, Effect of Cr plating on the coupling current loss in cable-in-conduit conductors, *Proc. ICCC-ICMC 96*, Kitakyushu, Japan: Elsevier, 1996, p. 1243.
  51. Y. Iguchi, P. Dolgoshnev, and V. Sytnikov, Prospective barrier coatings for superconducting cables, *Supercond. Sci. Technol.*, **10**: 507, 1997.
  52. K. Kawaguchi and I. Horvath, Experimental evidence for an interaction effect in the coupling losses of cabled superconductors, *Cryogenics*, **23**: 9, 1983.
  53. P. Bruzzone et al., Conductor fabrication for the ITER model coils, *IEEE Trans. Magn.*, **22**: 2300, 1986.
  54. N. Aoki et al., Fabrication of superconductor for the THW-TJ coil, *Cryogenics*, **23**: 581, 1983.
  55. K. Nakamoto et al., Design and fabrication of forced-flow superconducting poloidal coils for the Large Helical Device, *Fusion Technology 1989*, Amsterdam: Elsevier, 1995, p. 909.
  56. R. Heller, Superconductor for the coils of the modular stellarator Wendelstein 7 X, *IEEE Trans. Magn.*, **30**: 2393, 1994.
  57. J. V. Minervin et al., Conductor design for the GEM detector magnet, *Proc. ISSC*, San Francisco, 1983, Vol. 5, p. 585.
  58. Y. Kamano et al., Development of superconductors for the DEMO poloidal coils (DPC-OL, C2), *Proc. Magn. Technol. Conf.*, Tsukuba, 1989, Vol. 11, p. 841.

59. J. R. Miller et al., Experience in sheathing 10 km of cable-in-conduit conductor for the NIMFI hybrid, *Adv. Cryog. Eng. Mater.*, **41**: 449, 1994.
60. D. Besaette et al., Fabrication and test results of the 40 kA CEA conductors for NET/ITER, *Fusion Technology 1992*, Amsterdam Elsevier, 1993, p. 789.
61. D. L. Walker et al., Design of a 200 kA conductor for superconducting magnetic energy storage (SMES), *Adv. Cryog. Eng.*, **35**: 573, 1990.
62. A. Anghel et al., The ITER quench experiment on long length at the SULTAN facility, *Fusion Technology 1994* Amsterdam Elsevier, 1995, p. 931.
63. B. Z. Li et al., Conductor fabrication for the HT-7U model coil, *Fusion Technology 1998*, Amsterdam Elsevier, 1998, p. 775.
64. S. Forster, U. Jeske, and A. Nyilas, Fabrication of a 15 kA NbTi cable for the 150 T/s high ramp rate poloidal coil, *Fusion Technology 1988*, Amsterdam Elsevier, 1989, p. 1557.
65. R. Heller et al., Stability of a poloidal field coil under rapidly changing magnetic field, *IEEE Trans. Magn.*, **32**: 2326, 1996.
66. M. A. Janacko, Lattice liquid superconductors, *IEEE Trans. Magn.*, **15**: 797, 1979.
67. P. Piazzone, Fully transposed leads for the prototype cable-in-conduit conductors of NET, *IEEE Trans. Magn.*, **28**: 190, 1992.
68. P. Bruzzone, N. Mitchell, and J. Kübler, Mechanical behavior under transversal load of the 40 kA Nb<sub>3</sub>Sn cable-in-conduit conductor for the NET inner poloidal coil, *Proc. Magn. Technol. Conf.*, Tsukuba, 1989, Vol. 11, p. 926.
69. C. R. Walters et al., Quench transients in internally cooled conductors, *IEEE Trans. Magn.*, **19**: 640, 1983.
70. M. Shinoda and S. Tazoe, Effect of solution on cryogenic mechanical properties of aged stainless steel, *Adv. Cryog. Eng.*, **34**: 131, 1988.
71. J. Kübler, H. J. Schneider, and W. J. Mueser, Influence of aging on the fracture toughness of cryogenic austenitic stainless steels, evaluated by a simple method, *Adv. Cryog. Eng. Mater.*, **36**: 191, 1992.
72. R. P. Reed, R. P. Walsh, and C. N. McCowan, Effect of Nb<sub>3</sub>Sn heat treatment on the strength and roughness of 316 LN alloys with different carbon content, *Adv. Cryog. Eng. Mater.*, **38**: 45, 1992.
73. H. Nakajima et al., Development of high strength austenitic stainless steel for conduit of Nb<sub>3</sub>Al conductors, *Adv. Cryog. Eng.*, **42**: 323, 1996.
74. R. P. Walsh, L. T. Summers, and J. R. Miller, The 4 K tensile and fracture toughness properties of a modified 316LN conduit alloy, *Proc. ICFE/ICMC 96*, Kitakyushu, Japan, Elsevier, 1996, p. 1891.
75. B. Bruzzone et al., The cable-in-conduit Nb<sub>3</sub>Sn conductor for the EURATOM ENEA 12 T wind and react magnet, *IEEE Trans. Appl. Supercond.*, **3**: 315, 1993.
76. A. della Torre et al., Successful completion of the conductor manufacture for the ITER TF model coil, *Fusion Technology 1996*, Amsterdam Elsevier, 1996, p. 941.
77. M. M. Almstrud and M. G. Heenig, Constructing, fabricating and forming of internally cooled cabled superconductors, *IEEE Trans. Magn.*, **17**: 922, 1991.
78. M. Nishi et al., A 12 T forced flow type superconducting magnet, *IEEE Trans. Magn.*, **23**: 1531, 1987.
79. M. Yamaguchi et al., Development of a 12 T forced cooling toroidal field coil, *Proc. ICFE*, **10**: 108, 1984.
80. F. Negami et al., First results of the Japanese program on superconducting MRI magnets technology, *IEEE Trans. Magn.*, **30**: 2086, 1994.
81. Y. Takahashi et al., Development of a 30 kA cable-in-conduit conductor for pulsed poloidal coils, *IEEE Trans. Magn.*, **19**: 346, 1983.
82. M. O. Huenig, A. G. Montgomery, and S. J. Waldman, Experimental evaluation of a 1 m semi-D-shaped test coil fabricated from a 23 m length of internally cooled cabled superconductor, *Adv. Cryog. Eng.*, **25**: 251, 1979.
83. *Mecherkelestudie über einen Supraleiter für Wendelstein VII-X*, ABB Rep. HIM 20-142, 1989.
84. H. Kinnhardt, O. Darmochi, and J. Snpper, Design and manufacture of a Wendelstein 7-X demonstration coil, *Fusion Technology 1999*, Amsterdam Elsevier, 1999, p. 735.
85. N. N. Murtyvetsky et al., GEM detector conductor manufacturing experience, *IEEE Trans. Appl. Supercond.*, **5**: 761, 1995.
86. R. E. Gold et al., Evaluation of conductor sheath alloys for a forced flow Nb<sub>3</sub>Sn superconducting magnet coil for the large coil program, *Adv. Cryog. Eng. Mater.*, **24**: 750, 1982.
87. I. S. Hwang et al., Mechanical properties of Incoloy 908—an update, *Adv. Cryog. Eng. Mater.*, **38**: 1, 1992.
88. A. Buscha, R. L. Taylor, and J. B. Berger, Superconductor conduits: Fatigue crack growth rate and near threshold behavior of three alloys, *Adv. Cryog. Eng. Mater.*, **38**: 167, 1992.
89. W. Specking, J. L. Duchateau, and P. Deenal, First results of strain effects on critical current of Incoloy jacketed Nb<sub>3</sub>Sn CICC's, *Proc. Magn. Technol. Conf.*, Beijing, 1997, Science Press, 1998, Vol. 15, p. 1216.
90. J. S. Smith, J. H. Weber, and H. W. Szok, Control of stress-accelerated oxygen assisted cracking of Incoloy alloy 908 sheath for Nb<sub>3</sub>Sn cable-in-conduit, *Adv. Cryog. Eng.*, **42**: 407, 1996.
91. N. Mitchell et al., Avoidance of SAGBO in Incoloy<sup>®</sup> 908 used as a jacket material for Nb<sub>3</sub>Sn conductors, *Proc. Magn. Technol. Conf.*, Beijing, 1997, Science Press, 1998, Vol. 15, p. 1163.
92. J. H. Weber and J. M. Prots, Production and properties of Incoloy<sup>®</sup> alloy 908 tubing for sheathing of Nb<sub>3</sub>Sn superconducting cables, *Adv. Cryog. Eng.*, **42**: 339, 1996.
93. H. Nakajima et al., 4 K mechanical properties of pure titanium for the jacket of Nb<sub>3</sub>Sn superconductors, *Proc. ICFE/ICMC 96*, Kitakyushu, Japan, Elsevier, 1996, p. 1895.
94. P. Sanger et al., The trials and tribulations of fabricating the pipe for the "rape in pipe" Nb<sub>3</sub>Sn superconductor, *Adv. Cryog. Eng.*, **28**: 751, 1982.
95. N. Kazuro et al., Design of the Nb<sub>3</sub>Al insert to be tested in ITER central solenoid model coil, *IEEE Trans. Magn.*, **32**: 2336, 1996.
96. V. Sytnikov et al., RFPackaging line for manufacturing ITER cable-in-conduit conductor, *Proc. ICFE/ICMC 96*, Kitakyushu, Japan, Elsevier, 1996, p. 799.
97. S. Forster et al., Development of components for poloidal field coil-wind the KFK Pol-project, *Fusion Technology 1990*, Amsterdam Elsevier, 1991, p. 1706.
98. B. Garré et al., Development and manufacture of superconducting cable-in-conduit conductors, *Cryogenics*, **34** (Suppl.): 619, 1994.
99. M. Nishi et al., Development of high current density, large superconducting coil for fusion machines—The DUC-TJ program, *Cryogenics*, **33**: 573, 1993.

PIERLUIGI BRUZZONE

Centre de Recherches en Physique  
des Plasmas

## HTS FILM GROWTH

Thin films of superconductors are primarily of interest for electronics applications. High performance superconducting thin films are essential, for example, for high-frequency passive devices and Josephson junction circuits. Thin films are also important for fundamental studies of superconducting materials,

**Table 1. HTS Materials Systems, Methods Used to Fabricate Thin Films, and Applications of the Thin Films**

| Material   | $T_c$ (K) | Thin Film Deposition Technique                        | Applications                    |
|--|-----------|---|---------------------------------|
| YBaCuO (2)   | 85–95     | In situ: sputtering, PLD, evaporation; MOCVD, ex situ | RF devices, Josephson junctions |
| BiSrCaCuO<br>2201, 2212, 2223                      | 10–115    | In situ: sputtering, PLD, evaporation                 | Josephson junctions             |
| TlBaCaCuO<br>1201, 1212, 1223,<br>2201, 2212, 2223 | 20–125    | PLD, In situ sputtering; ex situ                      | RF devices                      |
| HgBaCaCuO<br>1201, 1212, 1223                      | 95–138    | Ex situ   |                                 |
| LaSrBaCuO  | 20–40     | In situ   |                                 |
| NdCeCuO  | 20–30     | In situ   |                                 |

where specially designed experiments can take advantage of the thin layer geometry and the capability of layering various combinations of materials, as well as lithographically defining fine features in them. Additionally, there is currently growing interest in using thin film deposition technology for deposition of thick films on polycrystalline substrates for power applications.

In the decade since the discovery of high-temperature superconductivity (HTS) in cuprate compounds, thin film materials have played an important role in the evolution of this field. The discovery of HTS has led to a rapid development of many different thin film deposition techniques. HTS thin films are now routinely made in hundreds of laboratories around the world, and thin film research is actively pursued in areas such as new superconducting materials and epitaxial oxide heterostructures. More recently, advances in deposition processes have brought about the synthesis of larger-area HTS films, making fabrication less expensive and allowing for high-throughput manufacturing of single superconducting-layer films.

In this article, we review the physical vapor deposition (PVD) techniques most frequently used for HTS thin film synthesis. Chemical vapor deposition techniques are reviewed in another article of this encyclopedia. Here, we review only the methods and issues that are relevant for film deposition. This does not cover the details of the film growth processes, nor the physical properties of HTS films and their applications. The reader is referred to other references for a more complete discussion of HTS thin films and their applications (1,2).

## MATERIALS

The HTS material of choice for applications is still YBa<sub>2</sub>Cu<sub>3</sub>O<sub>7</sub> (YBCO), the first cuprate discovered to superconduct above 77 K. There are several compelling reasons why this material is still the most suitable for many of the applications. Some of the fundamental ones are its smaller conduction anisotropy, higher superconducting critical currents in a magnetic field, and greater chemical stability than in other HTS compounds. However, the most important reason is the ease of fabricating high-quality, single-phase YBCO thin films.

In this review, YBCO will be considered the canonical high- $T_c$  superconductor, and the discussion will be limited to this example, partly for reasons of conciseness and partly because

of the prominence of this compound in the HTS community. Additionally, unless otherwise mentioned, one is generally considering growth of *c*-axis-oriented films, i.e. where the CuO<sub>2</sub> layers are parallel to the substrate.

Table 1 reviews the HTS compounds and methods used to date to prepare these thin films. For a review of materials, we refer the reader to Refs. 2 and 3. There has been a substantial effort related to Bi–Sr–Ca–Cu–O thin films (2212 and 2223 phases) by a variety of techniques, particularly in Japan. Some industrial work in the US has focused also on Tl–Ba–Ca–Cu–O (mostly 2212), primarily for passive electronics applications. One of the obvious reasons for looking at these other compounds is their higher critical temperatures. Nevertheless, the advantages of YBCO have been hard to surpass, particularly at temperatures below 65 K. Substantial thin-film work has also been devoted to La–Sr–Ca–O, Nd–Ce–Cu–O, Hg–Ba–Ca–Cu–O, and infinite-layer compounds, mostly driven by academic interest in basic properties of these HTS materials.

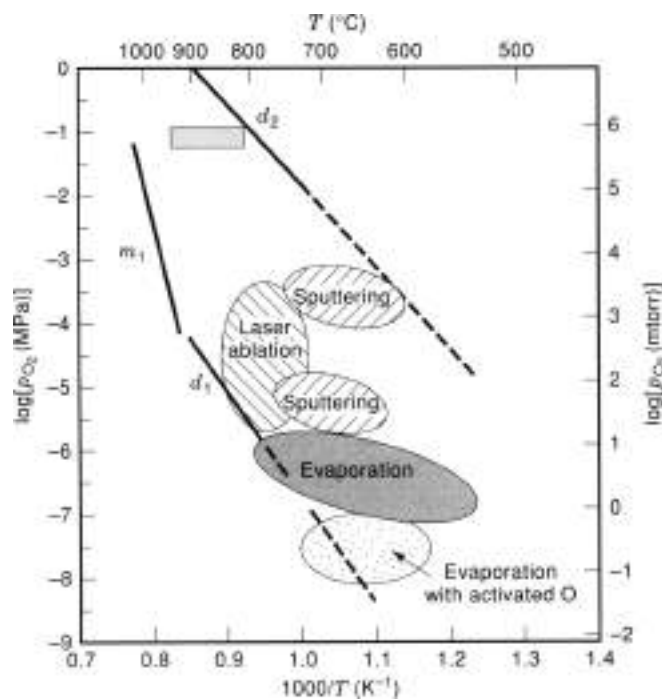
## THERMODYNAMIC ISSUES

Film growth is inherently a nonequilibrium process. Nevertheless, thermodynamic stability is important as a driving force for the reactions taking place during growth. HTS compounds possess large unit cells which can have additional complexities, such as a wealth of metal-atom defects and oxygen non-stoichiometries. These complex structures require temperatures for formation close to their melting points, typically  $(1.6–0.9)T_m$ , much higher than is usual for growth of epitaxial metals and semiconductors. The intricacies of film growth and the thermodynamics of phases produced are certainly not sufficiently well understood at present. Below, we summarize some of the key thermodynamic issues as they are currently accepted.

### Oxygen Stability

HTS compounds are oxides and, as such, require the presence of oxygen during synthesis. Although it is technically possible to deliver oxygen through the substrate, in practice, oxygen is delivered as a gas impinging on the growth surface together with the cation species. Figure 1 shows the pressure-temperature thermodynamic stability diagram for the bulk YBCO compound. One can see from the diagram that a certain



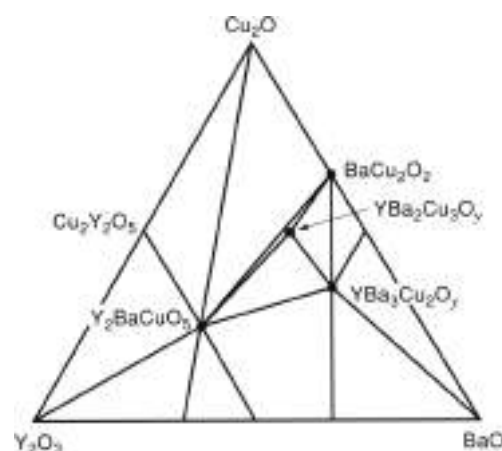


**Figure 1.** Thermodynamic stability diagram for YBCO in oxygen pressure and temperature. The shaded elliptical areas indicate regimes where in situ film growth for YBCO has been reported successful. The rectangles show ex situ film growth regimes.

minimum pressure must be maintained for the stability of YBCO. Below that oxygen pressure, depicted by the line labeled  $d_1$ , the YBCO compound is not stable and will decompose. For the typical temperatures during film formation, this translates into a minimum oxygen pressure of about 100 mPa (1 mtorr). At higher oxygen pressures, there is another decomposition line for YBCO, labeled  $d_2$ , usually not reached during in situ growth of thin films. During cooldown of films to room temperature, this second decomposition is usually not observed because it is strongly limited by diffusion kinetics. However, the stability diagram also implies that the YBCO compound is metastable at room temperature and pressure. In fact, most HTS compounds are believed not to be thermodynamically stable at lower temperatures. References 4 and 5 have a more detailed discussion of the YBCO stability lines. Hammond and Hornum (6) argued that growth of thin films is optimal close to the high-temperature thermodynamic decomposition line of the compound. In fact, the empirical data confirm this, as can be seen also in Fig. 1, where successful film growth regions are indicated. For processes that possess a more reactive form of oxygen, such as atomic oxygen or ozone, the equivalent decomposition line is shifted compared with that in Fig. 1. Such activated oxygen species have been proven helpful for growth of YBCO when the total pressure is lower than about 1 Pa (10 mtorr).

#### Compositional Phase Diagram

The HTS compounds contain typically three to five different metal species in addition to oxygen. Some of these materials are line compounds (and this is presumed to be the case for YBCO) and, as such, do not accept a solid solution of atoms in their chemical formula. This means that one will always be synthesizing a number of phases in addition to the desired HTS



**Figure 2.** Thermodynamic ternary phase diagram for Y-Ba-Cu oxides at temperatures below 850°C and oxygen pressures below one atmosphere.

material. For YBCO, the Gibbs phase rule implies two other compounds as impurity phases. Figure 2 shows the present understanding of the ternary phase diagram for Y-Ba-Cu oxides at the low oxygen pressures relevant for in situ thin film growth (7,8). The corners of the triangles define the phases which will be present for any composition in that triangle.

There are some caveats to the description given above. First of all, as already stated, film growth is not an equilibrium process, and therefore metastable phases could be formed during synthesis. Secondly, the HTS material that is formed during film growth may not be, and probably is not, exactly like the bulk material. This has now been well documented in the literature for YBCO films. In fact, it is also fairly well established that YBCO film growth does not match exactly the thermodynamic phase diagram shown in Fig. 2. In particular, the  $Y_2BaCuO_5$  phase has not been observed in in situ grown films. On the other hand,  $CuO$  and  $Y_2O_3$  second phases are commonly observed in YBCO films prepared close to 1:2:3 stoichiometry and are not predicted by the phase diagram.

#### EPITAXY

Due to the significant anisotropy of the HTS materials, the easy direction for electrical current flow is in the  $ab$  plane, along the  $Cu-O_2$  planes. For devices which require current to flow parallel to the substrate, the growth direction of the HTS films has to be in the  $c$ -axis direction. In order to achieve  $c$ -axis-oriented growth, good epitaxial growth of all grains is necessary. Furthermore, a high-angle, in-plane grain boundary of two  $c$ -axis-oriented grains has been found to behave as a superconducting weak link. This implies that one needs to align all grains in the plane of the substrate as well as out of the plane. The requirements of in-plane epitaxy are very well illustrated in the case of yttria-stabilized zirconia (YSZ) substrates where the YBCO films grow in the  $c$  direction, but the  $c$ -axis grains can have two major different orientations in the  $ab$  plane, 45° apart (10). In the case of applications which require the current flow to be perpendicular to the substrate,  $a$ -axis growth has been implemented. Many multilayer Josephson devices have been tried in this way.  $a$ -axis growth is typically initiated with a lower substrate temperature, which in turn reduces the mobility of ions

Table 2. Substrates and Buffer Layer for HTS Thin Films

| Substrate                      | Orientation          | Buffer Layer  |
|--------------------------------|----------------------|---|
| SrTiO <sub>3</sub>             | (100)                | none<br>CeO <sub>2</sub>                                  |
| LaAlO <sub>3</sub>             | (100)                | CeO <sub>2</sub>  |
| MgO                            | psuedocubic<br>(100) | none<br>SrTiO <sub>3</sub>                                |
| YSZ                            | (100)                | CeO <sub>2</sub><br>Y <sub>2</sub> O <sub>3</sub>         |
| Al <sub>2</sub> O <sub>3</sub> | (1102)               | CeO <sub>2</sub><br>MgO                                   |
| Si                             | (100)                | YSZ/Y <sub>2</sub> O <sub>3</sub><br>YSZ/CeO <sub>2</sub> |

necessary for the *c*-axis growth. Due to the lower growth temperature, films are formed which are structurally less perfect. To overcome this problem, many workers have used a technique of starting growth at a lower temperature for *a*-axis nucleation, making a template, and then raising the temperature for the remainder of film deposition (16).

### Substrates

The most frequently used substrate for fundamental research is SrTiO<sub>3</sub>, because of its good thermal and lattice constant match with most HTS materials. However, high dielectric loss, small wafer size, and high cost of SrTiO<sub>3</sub> render it unfavorable for technological applications. LaAlO<sub>3</sub> has emerged as an alternative to SrTiO<sub>3</sub> as a low-loss perovskite substrate, although its critical dielectric properties are not under complete control. MgO substrates, on the other hand, have an even lower loss for microwave applications. However, the reproducibility of the MgO surface quality has been a problem in the earlier days, but more recently, the surface quality has improved. For applications where very low loss is required, MgO is the best HTS substrate at present. MgO and LaAlO<sub>3</sub> are commercially available in wafer sizes larger than 5 cm (2 in.). More conventional substrates such as Si and sapphire have been used and found to be incompatible with thick HTS thin films, mainly due to microcracking in the HTS films resulting from the thermal expansion mismatch between the substrates and the HTS materials. The maximum thickness of YBCO one can grow without significant microcracking is about 50 nm on Si and about 600 nm on sapphire. Table 2 shows a list of substrates most commonly used today for growth of HTS films. References 17 and 18 have a more extensive discussion of HTS compatible substrates and buffer layers.

In most cases, excellent epitaxial alignment can be achieved by depositing oxide buffer layers, such as CeO<sub>2</sub>, before HTS growth. The CeO<sub>2</sub> buffer layer has been found to suppress the unwanted *a*-axis nucleation for the *c*-axis growth on substrates such as SrTiO<sub>3</sub> and LaAlO<sub>3</sub>. On substrates such as YSZ and *c*-plane sapphires, the CeO<sub>2</sub> buffer layer promotes in-plane alignment of each *c*-axis grain, thereby eliminating potential high-angle grain boundaries. In fact, by using an appropriate buffer layer and a seed layer, one can create a process to produce 45° angle grain boundaries in a controlled fashion. On Si substrates, the YSZ buffer layer was found to grow epitaxially and to be very effective in preventing chemical reaction between HTS materials and Si.

### Ion-Beam-Assisted Deposition

Recently, a new technique, ion-beam-assisted deposition (IBAD), has been implemented to grow HTS layers on polycrystalline substrates. By bombarding the growing film surface with an ion beam incident at a specific angle to the substrate, one is able to achieve partial in-plane crystalline alignment of the deposited material where the substrate has none. Iijima et al. (19) first utilized this technique to grow in-plane textured zirconia buffer layers on substrates made of polycrystalline Ni-based alloys. YBCO was then deposited by PLD. Due to the absence of high-angle grain boundaries in YBCO, the critical currents attained are much higher than without the IBAD process. Other groups, including the one at Los Alamos National Laboratory, has followed on this work and achieved critical currents over 10<sup>4</sup> A/cm<sup>2</sup> at 77 K (20). More recently a group at Stanford University has demonstrated very good IBAD results with 10 nm thick MgO films (21). While more development is needed to establish whether the IBAD process is a viable one for large-scale applications, it certainly opens up a whole new range of substrate materials for HTS film growth.

### FILM GROWTH METHODS

Most of the methods currently employed for fabrication of HTS thin films involve formation of the cuprate crystal structure during film deposition. These are usually called in situ processes, and they will comprise most cases covered in this review. Chemical vapor deposition processes, which are also in situ growth, are covered in another chapter of this volume. Historically, the processes that were developed first, required a post-deposition anneal, or simply postanneal, in order to crystallize the material. Such methods are now less frequently utilized, since they are not well suited for fabrication of multilayer structures and are thus technologically more limiting. However, they are still used for compounds such as TlBaCaCuO or HgBaCaCuO, where the in situ processes are often not practical due to the required high vapor pressures of Hg and Tl. Finally, there are other processes such as sol-gel, plasma spray deposition, and liquid phase epitaxy that are generally used for fabrication of thick films (thicker than 1 μm); these also will not be covered in this brief review.

### Postanneal Growth Methods

In the post-anneal, or ex situ, growth process, the metallic elements are deposited on a substrate in the correct composition as an amorphous or multilayer film, usually in compound form with oxygen and possibly fluorine. Subsequent annealing in air or oxygen at a high temperature, typically ≥850°C for YBCO, forms an epitaxial film of the desired HTS phase (9). A very commonly used postanneal method for growth of YBCO films is achieved by using BaF<sub>2</sub> in the precursor film. Water is then required in the annealing step in order to eliminate the fluorine and start the HTS growth process. Since the critical growth step takes place separately from deposition, the actual technique used to deposit the layers is of secondary importance. Sputtering and evaporation are typically used.

Ex situ methods have initially focused on synthesis under atmospheric oxygen pressure and correspondingly high temperatures. This region of phase space is depicted in Fig. 1 with the rectangle in the upper left corner. Later work has shown

that lower-temperature ex situ growth can also be achieved if the oxygen pressure is reduced at the same time (see also the lower rectangle in Fig. 1). Such films grown under lower oxygen pressures were observed to have properties closer to the in situ grown films (10,11).

#### In situ Growth Methods

Dominant methods in use today for physical vapor deposition of HTS are sputtering and laser ablation (also referred to as pulsed laser deposition, or PLD). Both of these methods are most commonly done from a single target and as such, because rather popular mostly because they are relatively simple to implement and fairly reproducible in the films they produce. In addition, targets for PLD are relatively inexpensive and easy to fabricate, and hence, the technique is well suited for investigation of many different materials. Less widespread today, but currently growing in popularity, are coevaporation and molecular beam epitaxy. After the initial slow start in contending with a sufficiently oxidizing environment in high vacuum, evaporation methods have now emerged as an established way to grow high-quality HTS films.

The following is a list of in situ physical vapor deposition techniques used for deposition of HTS compounds, which will be covered in this review.

#### 1. Sputtering

- On-axis magnetron
- Off-axis magnetron
- Inverted cylindrical magnetron
- Ion-beam sputtering

#### 2. Laser ablation

- On-axis
- Off-axis
- Laser-MBE

#### 3. Evaporation

- Flash evaporation
- Reactive coevaporation
- Reactive-MBE

For more discussion on various PVD techniques, we also refer the reader to Refs. 12, 13, and 14.

## SPUTTERING

Sputtering is a very commonly used technique for metal deposition in semiconductor and magnetic storage industries. As such, it was applied early on to HTS materials. The first event of significance was the use of a single composite target, which became commonplace for HTS materials because of the inherent difficulty of reproducing the metal-atom stoichiometry. The major complication compared to more conventional sputtering of metals is the energetic negative-ion (oxygen) bombardment of the substrate, due to the ionic nature of the oxide target. This causes resputtering and degradation of the sample. This has been recognized even before the advent of HTS and has been fully investigated by Rosanigel and Cuomo (22). Two

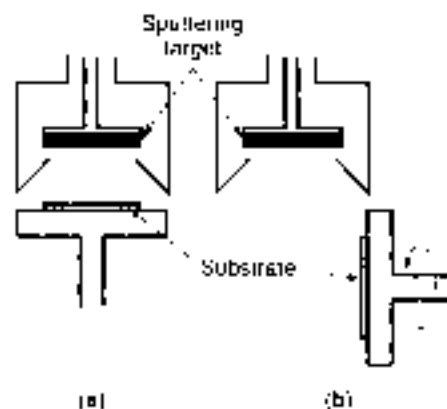


Figure 3. Schematic for the (a) on-axis and (b) off-axis sputter deposition.

types of solutions had emerged: (1) using a higher gas pressure in order to thermalize the energetic species (23), and (2) off-axis sputtering, where one eliminates the high-energy particles coming directly from the target (24,25). Figure 3 shows the schematic of the two processes. Other variations include on-axis unbalanced-magnetron sputtering (26) and inverted cylindrical magnetron (ICM) sputtering (27).

The energies of ejected atoms from the sputtered targets are in the range of tens of electron volts, and they get scattered by the background Ar and O<sub>2</sub> gas. The angular distribution of each cation species is different, and the *scand spot* for the proper cation composition is relatively small. In order to cover a 5-in. (12-in.) wafer uniformly from a single 5-cm target in an off-axis geometry, one has to resort to some sort of scanning method, such as rotation of the wafer. The deposition rate is very low, not only because the deposition rate for oxides is much lower than for metals, but also because of the off-axis geometry. It takes several hours to deposit a few hundred nanometers of film. Larger targets and a number of sputter guns depositing simultaneously have been used at several laboratories to increase the deposition rate. When all the deposition conditions are optimum, sputtering has demonstrated a capability to produce YBCO films of excellent crystallinity and surface condition. However, the deposition conditions have to be changed as the targets erode, because the changing target surface geometry results in a different plasma distribution. In addition, the substrate heating method for off-axis sputtering is not as simple as in the case of the on-axis laser ablation or evaporation which will be described later.

## LASER ABLATION

Laser ablation, or pulsed-laser deposition (PLD), is a relatively new technique that gained much popularity because it is ideally suited for deposition at a high oxygen pressure. The relative ease of this technique in depositing multicomponent oxides (and nitrides) has made it especially effective in exploring new materials for HTS electronics, such as epitaxial dielectrics or barrier layers. A short-wavelength (150 to 260 nm) excimer laser is focused onto a rotating target of the material to be deposited. Under the energy of the laser beam (0.1 to 2 J per pulse), the matter emitted from the target forms a plume that carries it to the substrate at supersonic velocities. In general, a higher gas pressure is required during laser ablation, due to

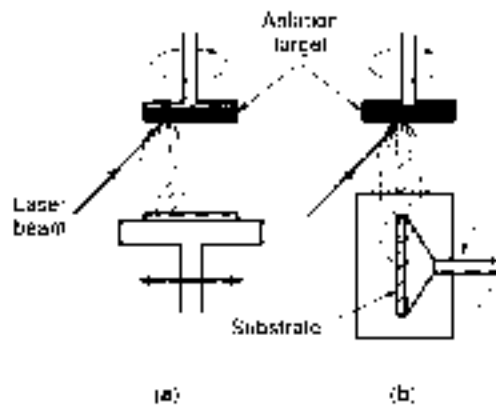


Figure 4. Schematic for the (a) on-axis and (b) off-axis laser evaporation processes.

the very high energy of the vaporized material from the surface of the target. The laser plume glows brightly from the target, and deposition is usually done near the end of the plume, about 5 cm away; see Fig. 4. Again, as in the case of sputtering, the various species scatter differently, and therefore the sweet spot of the deposition process is relatively small, usually one to two centimeters, depending on the geometry. The deposition rate per laser pulse ranges from a fraction of an angstrom to a few angstroms. The technique was found to be fairly reproducible from the early days and has been used extensively for research and developmental purposes during the past ten years. Even though other techniques are potentially more manufacture-friendly, laser ablation is still very actively used for prototyping devices made up of several complex materials, such as superconductors, ferroelectric oxides, and magnetic oxides.

A particular problem associated with PLD is the deposition of micron-size droplets, so-called *boulders*, on the grown film. These particles originate at the target and are emitted from the action of the laser pulse. A variety of procedures have been utilized to reduce this problem, so that the boulder density can be very low in the best films. Such procedures include target preparation (such as frequent polishing), defocusing the laser spot, mechanically chopping the plume, and spatially filtering the beam.

Although PLD can produce high deposition rates (up to tens of nanometers per second), the area on which one deposits is small. A straightforward way to increase the deposition area is to scan wafers over the plume either by moving the substrate vertically and horizontally or by rotating the substrate. This method has been tried in several laboratories and is still pursued. Another scheme utilizes a rotating cylindrical target with a linear laser profile to obtain larger deposition areas.

Another major difficulty of these techniques is uniform heating of large wafers. Most of the heating methods for small-area laser ablation have been to mount a substrate on a heated metal surface with silver paste, which is difficult to extend to larger sizes. In order to overcome this difficulty of heating a large wafer, an off-axis laser ablation technique has been developed; see Fig. 4(b). A large wafer (5 to 8 cm) can be mounted parallel to the direction of the plume inside a relatively simple blackbody-like heater. The deposition takes place when the atoms collide with the background pressure and are scattered to the surface of the wafer. By rotating the wafer and selecting an appropriate pressure for the geometry, a fairly uniform deposi-

tion was achieved over 5 cm wafers. This technique allows for simultaneous deposition on both surfaces of the wafer, which is an important benefit for microwave applications requiring a ground plane. A drawback of this technique is its low deposition rate because of the off-axis geometry. Typical conditions for deposition of a 5 cm wafer resulted in a deposition rate about a factor of 10 lower than for the on-axis deposition.

By using laser fluence just enough to evaporate a few atomic layers of the target in a low-oxygen environment (less than  $10^{-2}$  Pa =  $10^{-4}$  torr) and at the same time using a sequence of metal or metal oxide targets, one can obtain a process similar to molecular beam epitaxy (MBE) by evaporation, here called laser MBE. Some in situ diagnostic tools can then be used to characterize the growth of the materials. This technique has mainly been used to grow artificially layered superconducting materials, such as the infinite-layered superconductor (28,29).

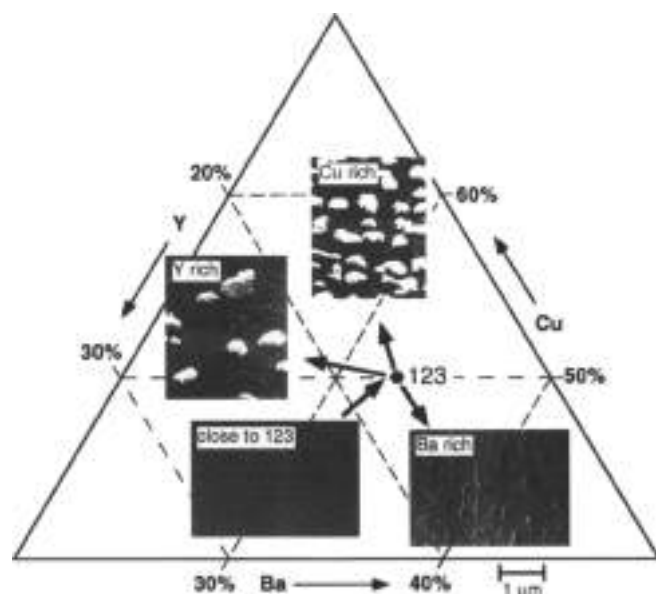
## EVAPORATION AND MOLECULAR BEAM EPITAXY

Deposition of HTS thin films by evaporation follows a tradition of such deposition of metal films. The added complexity here is the required partial pressure of oxygen during growth. One can distinguish several variations in the evaporation approaches: flash evaporation, reactive coevaporation of metals, and sequential deposition by MBE. Except for the first process, these evaporation techniques utilize individual metallic sources. However, in attempting to control individual sources, two difficulties arise. One is the need for very fine control of individual sources, including development of species-specific sensors. The second is the need to work at a low enough pressure to minimize beam scattering and at the same time achieve the highly oxidizing thermodynamic conditions required for the growth of these compounds. For an in situ process, one is confined to work at a pressure of molecular oxygen above 100 mPa (1 torr) or to use a more reactive form of oxygen supply than molecular oxygen.

### Flash Evaporation

The simplest approach to evaporation of HTS materials is to evaporate the compound in small batches, i.e. in flashes of evaporant material. Since the material does not melt congruently, it is not possible to establish a continuously constant rate of metal fluxes, but for short enough intervals one can average out the compositional variations. Usually, the evaporant material is a powder of YBCO located in a feed mechanism that drops small quantities onto an evaporation source (30). In most cases, the films require a post-annealing treatment to oxidize the film sufficiently, since the fast deposition does not allow for sufficient incorporation of oxygen. This method is at present less significant.

Probably the most technologically significant advance in HTS thin film evaporation has been the large-area heater developed for coevaporation by the group of H. Kinder at the Technical University of Munich (Germany) and depicted in Fig. 6. They used a blackbody-type rotating disk heater, similar to the one used by others in PLD deposition, but they added an *oxidation pocket* (31). This heater has a narrow slit that allows for a differential pressure between the oxidation pocket and the rest of the chamber of about 1000:1. The deposition and the oxidation processes are therefore separated, and a low pressure in the chamber improves the stability of thermal evap-



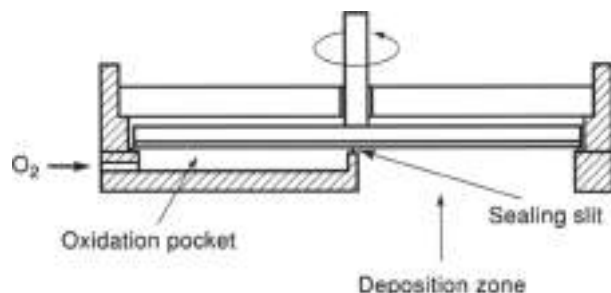
**Figure 5.** Morphology of YBCO films as a function of their composition during reactive coevaporation.

eration sources. Uniform YBCO films were grown on wafers up to 20 cm in diameter using this technique. More importantly, this method is significantly more cost-effective and has higher throughput than any of the other PVD processes.

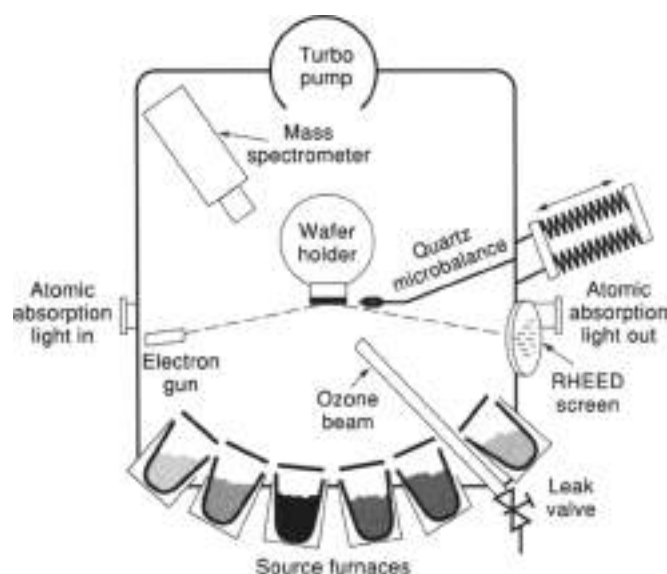
#### Reactive Coevaporation

Historically, the first attempts to make films by evaporation utilized a high molecular oxygen pressure. The problems associated with a high background oxygen pressure are rate control of the individual species and the degradation of the sources. In order to circumvent this difficulty, several approaches were taken. One was to accommodate a high differential pressure between the sources and the sample by introduction of nozzles in close proximity to the sample. Another method is to utilize a more reactive species for oxygen incorporation, such as atomic oxygen or ozone.

The evaporation technique allows one to tune the composition of the film by adjusting the relative rates of the sources. Several groups have worked over the past years on developing process control for coevaporation and studying YBCO film properties as a function of metal atom composition. Figure 5 shows SEM micrographs of films with various metal compositions. The general finding is that films that are grown slightly Y- and Cu-rich have better performance than films with exactly 1:2:3



**Figure 6.** Schematic of the heater for large-area HTS deposition by evaporation.



**Figure 7.** Schematic for the reactive MBE process used in HTS film growth.

stoichiometry. This is true for a majority of *in situ* techniques and not just evaporation. The reasons for this are still not completely clear—in particular, whether it is a materials issue or a process control issue. It is known that superconducting properties of films degrade significantly as one goes into the Ba-rich composition. It is possible that process fluctuations around the desired metal-atom stoichiometry are responsible for degradation of films close to 1:2:3 composition.

#### Molecular Beam Epitaxy

In general, one can distinguish between reactive molecular beam epitaxy (MBE) approaches to HTS film deposition and more reactive evaporation by the lower background pressure of the former technique. Another difference is that MBE deposition is usually done sequentially rather than simultaneously. Typically, MBE systems have multiple sources with individual shutters, as well as some *in situ* monitoring tools; see Fig. 7. Reference 32 has a thorough discussion of relevant issues in reactive MBE of HTS films. The work by a number of groups in this field has focused on careful atomic layering to produce very smooth films, as well as customized growth of new materials and heterostructures containing these phases. The reactive MBE technique also lends itself to careful control of chemical doping of these materials, an issue that is very important for the superconducting properties of the cuprates.

#### *In situ* Monitoring and Diagnostics

Evaporation and, especially, MBE, with their inherent low background pressure, lend themselves well to vacuum techniques for *in situ* deposition monitoring and film diagnostic tools such as reflection high-energy electron diffraction (RHEED). RHEED has been shown to be particularly helpful in controlling deposition of Bi-Sr-Cu-O films, where the growth occurs in blocks of subunit cells, but where many similar phases are possible. RHEED has been less helpful to date in the growth of YBCO, which occurs in blocks of unit cells and where intergrowths are more difficult to tailor due to the

higher stability of the primary phase. Terashima and coworkers (33) have shown that growth of YBCO can exhibit oscillations in the RHEED pattern commensurate with unit cell deposition, suggesting that growth proceeds in a unit-cell-by-unit-cell fashion, rather than in smaller building blocks. More recently, RHEED has also been extended to high-pressure processes such as PLD (34).

Another area of technological development has been in the use of optical absorption techniques for measurement and control of atomic fluxes. Both hollow-cathode lamp systems (35) and tunable diode lasers (36) have been used to monitor fluxes in situ in close proximity to the substrate. Such monitoring of individual fluxes is a prerequisite for careful control of HTS growth.

## CONCLUDING REMARKS

In spite of a wealth of research, growth of cuprate films has remained a complicated matter. This is due to the materials' rather complex multicomponent crystal structures. They are prone to a variety of defects and growth morphologies. Much work still remains to be done until films are better understood and more reproducible.

As already described, HTS films are now routinely made with out-of-plane as well as in-plane alignment on single-crystalline substrates. Aligned films are still not routine on polycrystalline substrates, but there has been progress in this area as well. This has been manifested in high critical currents in films, particularly for YBCO. Typical critical currents in good quality YBCO films are over  $2 \times 10^6$  A/cm<sup>2</sup> at 77 K and above  $10^7$  A/cm<sup>2</sup> at low temperatures. There has been much work in characterizing defect structures, but a good understanding of their influence on physical properties is still lacking. One growth structure that seems to be nearly ubiquitous in YBCO films is the spiral, which forms during growth and at the core of which is a screw dislocation (36). Figure 8 shows a scanning

tunneling microscopy (STM) image of a surface of a YBCO film with a spiral. Steps in this image have a height of one unit cell of YBCO. Such spirals have been identified as contributing to pinning of vortices, important for high critical currents.

Maximally oxygen-doped YBCO bulk materials, made at higher temperatures than thin films, have a superconducting transition temperature of 92 K and a resistivity of  $35 \mu\Omega/\text{cm}$  at 100 K. Although having nominally the same crystal structure as the bulk material, YBCO films often have different electronic properties. For example, lower  $T_c$  (85 to 90 K) and expanded  $c$ -axis lattice constants (1.170 to 1.172 nm) are frequently found in thin films. Additionally, flux-pinning centers appear to be highly dispersed in in situ grown films of YBCO, although the structural origin of these centers has not yet been identified. It is important to have a close feedback loop between film deposition and structural and electrical characterization in order to further improve HTS films.

Technology for deposition and monitoring of HTS thin films has developed significantly over the past decade. At the same time, the materials understanding of film nucleation, growth, and oxide interfaces has started to unfold. Together, these developments in the future will bring new capabilities for HTS film growth.

## BIBLIOGRAPHY

1. J. M. Powell, A decade of progress towards a high temperature superconducting electronics technology, *Solid State Commun.*, **102**: 269-282, 1997.
2. T. R. Lemberger, *Films of high-temperature oxide superconductors*, in D. M. Unalberg, ed., *Physical Properties of High Temperature Superconductors III*, Singapore: World Scientific, 1992.
3. C. P. Poole, H. A. Farach, and P. J. Creswell, *Superconductivity*, San Diego: Academic Press, 1985.
4. T. B. Lindemer et al., Decomposition of  $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$  and  $\text{YBa}_2\text{Cu}_3\text{O}_6$  for  $p_{\text{O}_2} < 0.1$  MPa, *Physica C*, **178**: 93, 1991.
5. R. K. Williams et al., Oxidation-induced decomposition of  $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$ , *J. Appl. Phys.*, **70**: 3008-3013, 1991.
6. R. H. Hammond and R. Borjesson, Correlation between the in situ growth conditions of YBCO thin films and the thermodynamic stability criteria, *Physica C*, **162-4**: 753-764, 1989.
7. R. Beyers and B. T. Ahn, Thermodynamic considerations in superconducting oxides, *Rev. Mater. Sci.*, **21**: 335, 1991.
8. P. Kjaer, O. Braaten, and A. Kjekshus, Chemical phase diagrams for the  $\text{YBa}_2\text{Cu}_3\text{O}_x$  family, *Acta Chem. Scand.*, **48**: 845-846, 1992.
9. M. Naito et al., Thin film synthesis of the high  $T_c$  oxide superconductor  $\text{YBa}_2\text{Cu}_3\text{O}_x$  by electron beam evaporation, *J. Mater. Res.*, **2**: 513-525, 1987.
10. R. Frenken et al., Effect of oxygen pressure on the synthesis of  $\text{YBa}_2\text{Cu}_3\text{O}_x$  thin films by post-deposition annealing, *J. Appl. Phys.*, **69**: 6569-6575, 1991.
11. J. M. Phillips et al., Comparison of  $\text{Ba}_2\text{YCu}_3\text{O}_{7-x}$  thin films grown on various perovskite substrates by coevaporation, *J. Mater. Res.*, **7**: 2650-2657, 1992.
12. R. C. Humphreys et al., Physical vapor deposition techniques for the growth of  $\text{YBa}_2\text{Cu}_3\text{O}_x$  thin films, *Supercond. Sci. Technol.*, **3**: 34-52, 1990.
13. J. A. Alarco et al., High quality YBCO thin films—laser deposition, co-evaporation, and device fabrication, *Phys. Scripta*, **44**: 95-101, 1991.
14. C. H. Stoezel et al., Thin film processing of high- $T_c$  superconductors, *J. Supercond.*, **0**: 1-1-17, 1993.

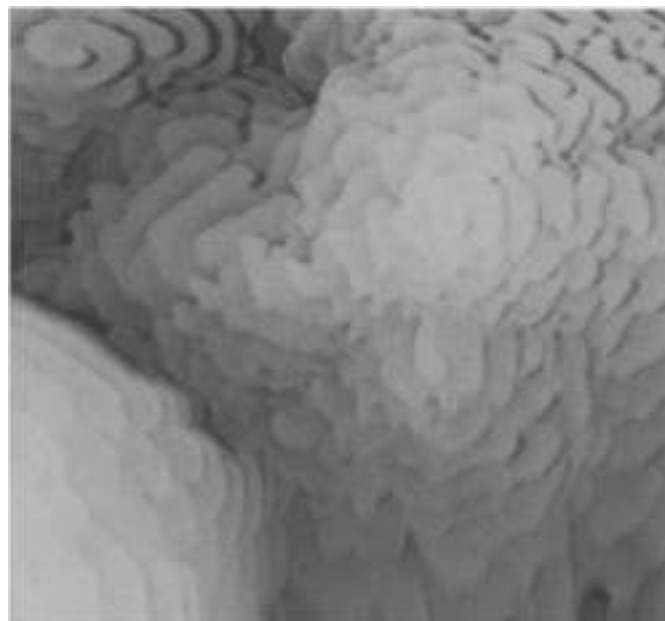


Figure 8. An STM image of a laser-ablated YBCO film showing spiral growth structures. Image courtesy of Prof. Harold Sorenson.

15. S. M. Garrison et al., Observation of two in-plane epitaxial states in  $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$  films on yttria-stabilized  $\text{ZrO}_2$ , *Appl. Phys. Lett.*, **58**: 2168-2170, 1991.
16. H.-U. Habermeier et al., Preparation and properties of YBCO thin films with the *c*-axis aligned in the film plane, in L. Carrera (ed.), *High Temperature Superconductor Thin Films*, Amsterdam: North-Holland, 1992, pp. 343-352.
17. A. Parrini and M. Sergent, in A. Narlikar (ed.), *Studies of High Temperature Superconductors*, vol. 7, Commack, NY: Nova Science Publishers, 1991.
18. J. Qian and C. Y. Yang, High- $T_c$  superconductors on buffered silicon: Materials properties and device applications, *Mater. Sci. Eng.*, **B14**: 177-202, 1993.
19. Y. Iijima et al., In-plane aligned  $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$  thin films deposited on polycrystalline metallic substrates, *Appl. Phys. Lett.*, **60**: 769-771, 1992.
20. X. D. Wu et al., High current  $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$  thick films on flexible nickel substrates with textured buffer layers, *Appl. Phys. Lett.*, **65**: 1961-1963, 1994.
21. C. P. Wang et al., Deposition of in-plane textured  $\text{MgO}$  on amorphous  $\text{Si}_3\text{N}_4$  substrate by ion-beam-assisted deposition and comparisons with ion-beam-assisted deposited yttria-stabilized-zirconia, *Appl. Phys. Lett.*, **71**: 2955-2957, 1997.
22. S. M. Rossnagel and J. J. Cuomo, Negative ion effects during magnetron and ion beam sputtering of  $\text{YBa}_2\text{Cu}_3\text{O}_x$ , in *Thin Film Processing and Characterization of High-Temperature Superconductors*, American Institute of Physics Conf. Proc. 165, New York: AIP, 1989, pp. 108-113.
23. L. Poppo et al., Direct production of crystalline superconducting thin films of  $\text{YBa}_2\text{Cu}_3\text{O}_x$  by high-pressure oxygen sputtering, *Solid State Commun.*, **66**: 661-665, 1988.
24. C. B. Fom et al., Synthesis and properties of  $\text{YBa}_2\text{Cu}_3\text{O}_x$  thin films grown in situ by 90 degree off-axis single magnetron sputtering, *Physica C*, **171**: 354-362, 1990.
25. J. R. Gavaler et al., Critical parameters in the single-target sputtering of  $\text{YBa}_2\text{Cu}_3\text{O}_x$ , *J. Appl. Phys.*, **70**: 4383-4391, 1991.
26. N. Savvides and A. Katsaris, In situ growth of epitaxial YBCO thin films by on-axis unbalanced direct current magnetron sputtering, *Appl. Phys. Lett.*, **62**: 528-530, 1993.
27. X. X. Xi et al., Preparation of  $\text{YBa}_2\text{Cu}_3\text{O}_x$  thin films by inverted cylindrical magnetron sputtering, *J. Less-Common Metals*, **161**: 349-355, 1989.
28. T. Kurwai et al., Superconducting artificial lattices grown by laser MBE, in L. Carrera (ed.), *High Temperature Superconductor Thin Films*, Amsterdam: North-Holland, 1992.
29. H. Kojima and M. Yoshizawa, Controlled formation of oxide materials by laser molecular beam epitaxy, *Appl. Surface Sci.*, **75**: 308-319, 1994.
30. Y. Yasuda et al., Y-Ba-Cu-O superconducting thin films prepared by plasma-assisted flash evaporation, *Appl. Phys. Lett.*, **55**: 307-309, 1989.
31. P. Berberich et al., Homogeneous high quality  $\text{YBa}_2\text{Cu}_3\text{O}_x$  on *g* and *g'* substrates, *Physica C*, **219**: 497, 1994.
32. D. G. Schlom and J. S. Harris, Jr., MBE growth of high  $T_c$  superconductors, in R. E. C. Farnwell (ed.), *Molecular Beam Epitaxy: Applications to Key Materials*, Park Ridge: Noyes, 1996.
33. T. Yeraslimskii et al., Reflection high-energy electron diffraction oscillations during epitaxial growth of high-temperature superconducting oxides, *Phys. Rev. Lett.*, **65**: 2684-2687, 1990.
34. G. J. H. M. Rijnders et al., In situ monitoring during pulsed laser deposition of complex oxides using reflection high energy electron diffraction under high oxygen pressure, *Appl. Phys. Lett.*, **70**: 1863-1865, 1997.
35. J. N. Eckstein, I. Bezovic, and G. E. Virshup, Atomic layer-by-

layer engineering of high- $T_c$  materials and heterostructure devices, *Mater. Res. Soc. Bull.*, **19**(9): 44-50, 1994.

36. W. Wang et al., Direct atomic flux measurement of electron-beam evaporated yttrium with a diode-laser-based atomic absorption monitor at 668 nm, *Appl. Phys. Lett.*, **71**: 31-33, 1997.
37. C. Gerber et al., Screw dislocations in high- $T_c$  films, *Nature*, **350**: 279-280, 1991.

KROKIN CHAE  
Centre de Recherches en Physique  
des Plasmas  
VLADIMIR MATJASEVIC  
Conductus, Inc.

## HTS JOSEPHSON JUNCTION DEVELOPMENT

Josephson junctions are the fundamental building blocks for a variety of superconducting electronics applications, including high-speed, low-power digital logic, and sensitive magnetic field and high-frequency electromagnetic detectors. A Josephson junction consists of a "weak" connection between two superconductors which exhibits the Josephson effects (described below). While low-temperature superconductor (LTS) Josephson junction technology is well-developed, high-temperature superconductor (HTS) Josephson junctions are still relatively immature. Nonetheless, extensive HTS junction fabrication efforts are in progress due to the possibility of applying Josephson effects at temperatures compatible with reliable, low-cost refrigerators. In this article we discuss the more common approaches to HTS junction fabrication and optimization, with a focus on potential digital circuit applications.

### Introduction to Josephson Junctions

According to the Bardeen-Cooper-Schrieffer (BCS) theory of superconductivity, the electrons in a superconductor are weakly bound into Cooper pairs and can be described mathematically by a complex wave function with spatially dependent amplitude and phase. The fact that all of these electrons occupy a macroscopic quantum state leads to several profound properties of superconductivity, such as zero resistance, quantization of the magnetic flux threading a hole in a superconductor, and the Josephson effects.

If two pieces of superconductor, each characterized by their own wave function, are brought very close together, but separated by a thin insulating layer, then the two wave functions can overlap. It was predicted by Brian Josephson [1] that this overlap would lead to novel phenomena associated with the dissipationless tunneling of Cooper pairs through the insulating barrier. The first of these, the direct current (dc) Josephson effect, is that a dc current can flow through this "Josephson junction" without the appearance of a voltage drop across the junction, and furthermore that the magnitude of this "super-current" is a function of the phase difference between the two electrodes. The maximum value of the supercurrent is called the critical current,  $I_c$ . Josephson's second prediction was that if a dc voltage were applied across the junction, then the phase difference would evolve at a rate proportional to the voltage, leading to a time-oscillating Cooper pair current. This is the alternating current (ac) Josephson effect.

Most of the above discussion can also be applied in cases where two superconducting electrodes are separated not by an insulator, but by other nonsuperconducting or weakly superconducting regions. These include normal metals, semiconductors, or even extremely narrow superconducting constrictions. This broader class of Josephson devices is generally known as "weak links," and essentially all HTS Josephson devices being developed fall under this heading. In particular, the use of a normal metal, or of a superconductor above its transition temperature, is the focus of much of the development work, and thus of the discussion in this article.

### HTS versus LTS Materials

Fabrication of HTS Josephson junctions is complicated by the materials properties of the high-temperature superconductors. Low-temperature superconductor junctions are typically produced using polycrystalline metallic Nb as the superconductor. In contrast, the new HTS materials are multicomponent ceramic oxides with four or more elements, such as yttrium-barium copper oxide (YBa<sub>2</sub>Cu<sub>3</sub>O<sub>7-x</sub>) or YBCO. The oxide superconductors have perovskite-based crystal structures with two to three copper-oxygen planes in a layered structure within a unit cell. As a consequence, these materials are anisotropic, with higher critical current densities and longer superconducting coherence lengths parallel to the CuO<sub>2</sub> planes. The superconductive coherence length is the characteristic length over which the Cooper pair density decays, as for example in approaching an interface between a superconductor and a normal metal. It also corresponds to the average physical separation of the two electrons which make up a Cooper pair. For a clean material (mean free path much longer than coherence length) it is given by the approximate expression  $\xi \approx 0.18\hbar v_F / k_B T_c$ , where  $v_F$  is the Fermi velocity and  $T_c$  is the superconducting transition temperature (1). In contrast to most LTS materials, the coherence lengths in the oxide superconductors are on the scale of atomic dimensions ( $\approx 30$  Å parallel to the copper oxide planes, and  $\approx 4$  Å perpendicular to the planes). Since damaged layers thicker than the coherence length result in degraded superconducting properties, surface cleaning and film growth at device interfaces are especially critical in the HTS materials.

Also unlike in most low temperature superconductors, grain boundaries in the HTS oxide superconductors lead to reduced critical current densities, and in fact can behave like Josephson junctions (see the section entitled "HTS Josephson Junction Types"). Because of the anisotropy and grain boundary problems associated with the high-temperature superconductors, epitaxial (single crystal) HTS films grown at high temperatures (850°C to 800°C) are required for many applications. High-temperature growth and the need to maintain high epitaxial quality in each layer of multilayer structures make HTS junction and circuit fabrication considerably more complex than the corresponding LTS processes.

### Selection of YBCO for HTS Junctions

Nearly all Josephson junction development in HTS materials has been based on YBCO or nearly identical compounds with another rare-earth element in place of yttrium. The superconducting critical temperature,  $T_c$ , for this family of compounds is approximately 90 K. The main advantage of using YBCO compared with other HTS material families is the low volatility of all cation elements which permits films to be grown in a

single step from a source—for example, a laser ablation target or sputtering target—fabricated with a stoichiometric ratio of yttrium, barium, and copper.

In contrast, films of the HTS families with the highest critical temperatures, Tl-Ba-Ca-Cu-O (TBCCO) and Hg-Ba-Ca-Cu-O (HgBCCO), are typically prepared in a two-step process in which a Tl- or Hg-deficient as-deposited film is annealed to the point of partial melting in a container sealed to obtain a high pressure of Tl or Hg vapor. Although grain boundary junctions can be formed by using this technique with bicrystal substrates or substrates with patterned steps, it is not extendible to multilayers. In the case of the Bi-Sr-Ca-Cu-O (BSCCO) or Ba-K-Bi-O (BKBO) families, films can be produced in a single step, but the volatility of K and Bi has limited film reproducibility and has discouraged junction development efforts. The TBCCO, HgBCCO, and BSCCO materials are also more highly anisotropic than YBCO, which introduces additional complications in 3-D structures.

### Basic HTS Josephson Junction Characterization

The current-voltage characteristic for an ideal Josephson junction described by the Resistively-Shunted-Junction (RSJ) model, is shown in Fig. 1. The RSJ model adequately describes the behavior of many HTS junctions, as explained in more detail in the section entitled "Josephson Effects." The primary junction parameters of interest are the critical current,  $I_c$ , the normal state resistance,  $R_n$ , and the product of these two factors,  $I_c R_n$ . The critical current is the maximum current that can flow through the junction without a voltage drop, while the normal state resistance is given by the inverse of the slope of the  $I$ - $V$  characteristics at a few times  $I_c$ , as shown in Fig. 1.  $R_n$  is due to dissipation in the device, for example, by current flow through the normal metal in a superconductor/normal metal/superconductor weak link.  $I_c$  and  $R_n$  are often expressed in terms of area-normalized quantities: the critical current density  $J_c = I_c/A$ , and the resistance-area product,  $R_n A$ . Characterization for an HTS Josephson junction typically includes (a) measurement of the current-voltage ( $I$ - $V$ ) characteristics as a function of temperature and (b) measurement of  $I_c$  modulation in a magnetic field. In some cases, the  $I$ - $V$  characteristics are also studied under microwave irradiation as a measure of the

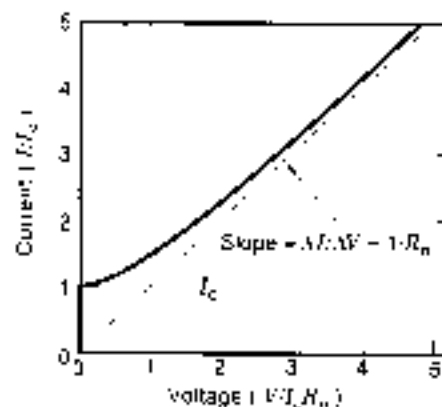


Figure 1. Current-voltage characteristic of a Josephson weak link in the zero-resistance limit, without thermal noise, showing definitions of the critical current,  $I_c$ , and normal resistance,  $R_n$ . The dashed line is the high-current asymptote of the  $I$ - $V$  curve.



ac Josephson effect. In a well-behaved HTS junction the  $I$ - $V$  shape will show the concave-up curvature consistent with the resistively shunted junction model. The temperature dependence of  $I_c$  and  $R_n$  gives information on the nature of coupling across the Josephson junction (e.g., proximity-effect coupling or tunneling), while the dependence of  $I_c$  on magnetic field provides a gauge of the uniformity of Cooper pair transport across the weak link. More comprehensive Josephson junction studies can also include determination of the  $J_c$  dependence on tunnel barrier or interlayer thickness, as well as measurement of  $J_c$  uniformity across multiple junctions on a chip (important for circuit applications):

### Survey of Junction Configurations

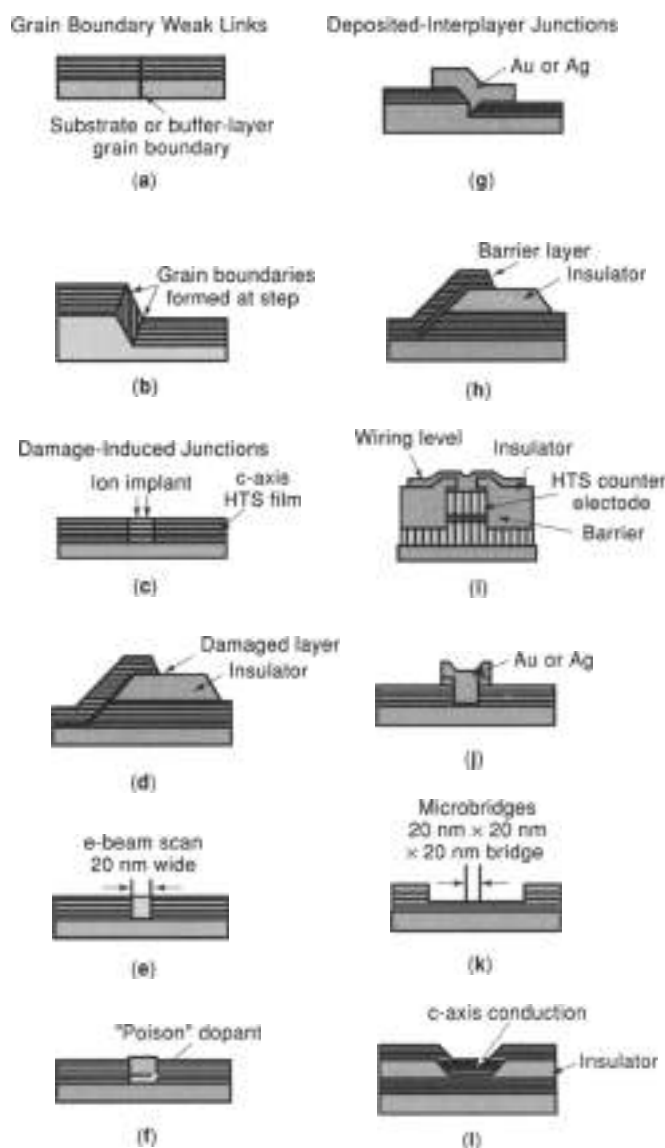
Josephson effects in HTS materials were first observed at naturally occurring grain boundaries in films and in weak links between two HTS samples created by either breaking a sample in vacuum and coupling across the vacuum gap or bringing two pieces in contact and coupling through a surface layer that was degraded by reaction in air. Most of the configurations used since that time to fabricate junctions with controllable and reproducible properties are shown in Fig. 2. All configurations require epitaxial films to prevent naturally occurring and randomly placed grain-boundary junctions from interfering with the engineered junctions. The fill pattern for HTS film layers in Fig. 2 indicates the orientation of Cu-O planes.

The junction configurations shown in Fig. 2 are grouped in the four categories of grain-boundary weak links, damage junctions, deposited interlayer junctions, and microbridges. The chronology of development was that single-HTS-film configurations that did not require special substrate treatments were explored first—for example, the junctions in Figs. 2(c), 2(f), and 2(j). These were followed by more complex processes where substrates were fabricated to promote the formation of a junction during HTS film growth as in Figs. 2(a), 2(b), and 2(g). Since the configurations in Figs. 2(d), 2(h), and 2(i) are direct analogues of LTS junctions, they were identified as candidate structures soon after discovery of HTS, but development did not begin until multilayer growth and patterning techniques were developed.

### Application Requirements

The key electrical parameters of an individual junction, which determine its suitability for a given application, are principally its  $I_c$  and  $R_n$ , and secondarily such parameters as capacitance  $C$  and inductance  $L$ . For applications which require many junctions, such as digital circuits, the uniformity of these parameters is of critical importance, while for sensor applications, such as superconducting quantum interference devices (SQUIDs), low intrinsic noise is a key characteristic.

For digital circuits a number of constraints dictate the range of parameters required. First, the product of  $I_c$  and a typical gate inductance  $L_g$  should be not much more than a flux quantum,  $\approx 2.07 \times 10^{-5} \text{ Wh} \approx 2.07 \text{ pH}\cdot\text{nA}$ . Since it is difficult to fabricate gates with inductance less than about 4 pH in HTS (see the section entitled "Circuit Integration"),  $I_c$  can be no more than about 0.5 nA. At the same time,  $I_c$  should be large enough that the Josephson energy,  $I_c^2/2e$ , is much larger (say by a factor of 100 to 1000) than the thermal energy,  $k_B T$ , otherwise there will be too many thermally induced, erroneous switching events. For operation at 40 K this dictates that  $I_c$  be greater



**Figure 2.** The wide usage of fabrication methods for HTS Josephson junctions can be loosely grouped into four classes: grain boundary junctions, damage-induced junctions, deposited-interlayer junctions, and microbridges. The deposited-interlayer approaches are thought to be the most promising for applications requiring more than a few junctions.

than 0.04 nA to 0.4 nA. At the same time, the product of  $I_c$  and  $R_n$  establishes the maximum reliable operating frequency of the circuit,  $2eI_c R_n/h$ . For Josephson devices to be competitive, this frequency must be at least tens of gigahertz. This means that  $I_c R_n$  must be greater than approximately 0.1 mV, and  $R_n$  must be of order 1  $\Omega$  to 10  $\Omega$ . While different arguments are applied for sensor applications of SQUIDs, such as magnetometry, the resulting requirements for  $I_c$  and  $R_n$  are quite similar.

### JOSEPHSON EFFECTS

The dc and ac Josephson effects introduced in the section entitled "Introduction to Josephson Junctions" can be stated

mathematically as follows:

$$\theta = \theta_1 - \theta_2 \quad (1a)$$

$$I = I_c \sin \theta \quad (1b)$$

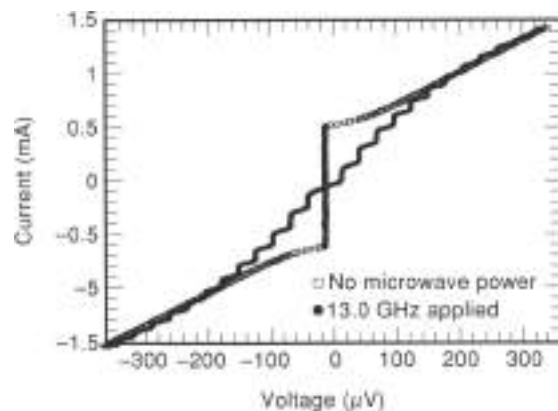
$$V = \frac{\hbar}{2e} \frac{d\theta}{dt} \quad (1c)$$

where  $\theta_1$  and  $\theta_2$  are the phases of the wavefunction in the two electrodes. In the presence of a magnetic field, with vector potential  $A$ , Eq. (1a) is generalized as follows:

$$\theta = \theta_1 - \theta_2 + \frac{2e}{\hbar} \int_1^2 \vec{A} \cdot d\vec{l} \quad (1d)$$

where  $\theta_1$  and  $\theta_2$  refer to two specific points on opposite electrodes, and the integral is taken along a straight line between those two points (1). While the phase difference across the junction cannot be directly measured, Josephson's predictions have several measurable consequences. For example, as a result of the periodic dependence of supercurrent on phase difference, and the field-dependence of the phase difference given by Eq. (1d), the critical current of a spatially extended junction displays a dependence on the magnetic field described by the Fraunhofer pattern characteristic of single-slit diffraction in optics:  $I_c \propto |\sin(\Phi_c/2\Phi_0)/(\Phi_c/2\Phi_0)|$ , where  $\Phi$  is the magnetic flux through the junction, and  $\Phi_0 = hc/2e$  is the superconducting flux quantum. The ac Josephson effect is responsible for Shapiro steps, which are constant voltage steps in the  $I$ - $V$  characteristics of the junction in the presence of microwave radiation. Such steps occur when the average voltage across the junction has values such that the oscillation frequency of the supercurrent is a multiple of the frequency of the applied radiation. This effect is the basis for the definition of the standard volt, which uses a precisely known frequency to generate a fixed voltage across a series array of many junctions. Figure 3 shows an example of Shapiro steps induced in a YBCO-Cobalt-YBCO/YBCO edge superconductor/normal metal/superconductor (SNS) junction (see the section entitled "Edge-Geometry Weak Links") by 13 GHz radiation.

Real-Josephson junctions can be modeled by the so-called resistively and capacitively shunted junction (RCSJ), or, for small



**Figure 3.** Current-voltage characteristics of a YBCO-Co-YBCO/YBCO edge SNS junction, showing Shapiro steps in response to 13 GHz radiation. Such steps are a manifestation of the ac Josephson effect.

capacitance, just RCSJ model, which takes the ideal Josephson element described by Eq. (2) and shunts it with a resistor and a capacitor. The resistor models the path by which quasiparticles cross the junction. For HTS junctions we can use the RCSJ model with a linear resistor and zero capacitance for comparison with the data. In this case the RCSJ model, without thermal noise, predicts a hyperbolic shape for the  $I$ - $V$  characteristic:

$$V = 0 \quad \text{for } I < I_c$$

$$V = R_c \sqrt{I^2 - I_c^2} \quad \text{for } I > I_c \quad (2)$$

where  $V$  is the time average of the instantaneous voltage,  $V(t)$ , which oscillates at the Josephson frequency with amplitude  $I_c R_c$ . The dc behavior is illustrated in Fig. 1. In the presence of thermal noise the sharp voltage onset at  $I_c$  is smeared out.

## HTS JOSEPHSON JUNCTION TYPES

As described in the section entitled "Survey of Junction Configurations," there are a wide variety of HTS Josephson junctions. In fact, it has proven relatively easy to fabricate HTS devices which exhibit Josephson effects, with the technology ranging in sophistication from naturally occurring grain boundary weak links to all-epitaxial structures incorporating superconductors, insulators, and deposited interlayers. However, it has proven difficult to meet the requirements of some of the more demanding applications. For example, single flux quantum (SFQ) digital circuits require junctions with high resistance, high  $I_c R_c$  products, and  $1 - \theta_c$  spreads less than  $10^{-3}$ . No junction technology to date has consistently met these constraints. At present the most widely used HTS junctions are grain boundary weak links and SNS edge junctions with doped YBCO interlayers. The grain boundary devices have excellent characteristics but relatively poor  $I_c$  spreads, while the SNS edge junctions are more difficult to produce but appear to be the most viable candidate for fabrication of complex circuits.

### Grain Boundary Junctions

**Naturally Occurring Grain Boundary Junctions.** Within a few months of the discovery of YBCO, Josephson effects were measured in bulk ceramic samples. An example of these measurements can be found in Robbins et al. (2), where a polycrystalline ceramic pellet was bonded to a glass slide, polished to a thickness of 0.25 mm, and constrictions were engraved in pattern a dc SQUID. Shapiro steps and modulation of the critical current by a magnetic field were observed.

Although randomly oriented grain boundary junctions in polycrystalline material are too poorly controlled to be practical for electronics, they are the most important obstacle in fabrication of HTS conductors for carrying large currents for large-scale, high-power applications. Many of the possible mechanisms for weak link formation at grain boundaries, such as impurities or second phase formation, have been eliminated by careful synthesis. However, the grain boundary continues to be the subject of research on topics such as the relationship between local strain fields and oxygen deficiency (3).

**Bicrystal Grain Boundary Junctions.** Dumas et al. (4) were the first to fabricate grain boundaries in HTS materials with controlled angles of misalignment. They cut and polished bulk

SrTiO<sub>3</sub> single crystals with symmetric  $\{100\}$  tilt boundaries. The two crystals were sintered together with parallel  $[001]$  axes. The bicrystal was then cut and polished to obtain  $[001]$  surfaces for growth of epitaxial *c*-axis-oriented YBCO films. At least on a macroscopic scale, the misorientation angle in the basal plane of YBCO matched the misorientation angle in the SrTiO<sub>3</sub> bicrystal.

The bicrystal experiments elegantly quantified the need for biaxial alignment in YBCO high-current conductors by demonstrating the rapid decrease in  $J_c$  with increasing grain misalignment. They also provided a route for synthesis of grain boundary junctions, so-called bicrystal junctions, with controlled misalignment between grains and controlled placement of junctions along a single line. Strontium titanate bicrystals with 24 and 37 tilt rotations—angles where particularly well-defined, clean interfaces can be formed—are commercially available.

Bicrystal junctions are still in use for low-junction-count applications, usually where an integrated HTS groundplane is not required. Specifically, most commercially available dc SQUIDS are fabricated in this way. Bicrystal junctions are the simplest and least expensive way for a university laboratory to pattern a few junctions for research. However, in addition to the obvious disadvantage imposed by placement of junctions along a single line, junction uniformity has not matched the level obtained by other junction configurations. Several different explanations for the lack of uniformity are based on the observation that the junction interface meanders along the substrate bicrystal boundary on a microscopic scale as YBCO grains overgrew the boundary from each side.

**Biepitaxial Grain Boundary Junctions.** Biepitaxial junctions are a variant of bicrystal junctions in which the bicrystal template for junction formation is provided not by a bulk bicrystal substrate but by the tendency of particular buffer layers to grow with different epitaxial orientations depending on the presence or absence of an intermediate 'seed' layer. The first implementation was by Char et al. (8), who observed that SrTiO<sub>3</sub>(001) grew on *R*-plane sapphire, Al<sub>2</sub>O<sub>3</sub>(1102), when there was no seed layer with in-plane parallel directions, SrTiO<sub>3</sub>(110) // Al<sub>2</sub>O<sub>3</sub>(1120). However, when there was an MgO(001) seed layer present, the in-plane orientation was SrTiO<sub>3</sub>(100) // MgO(100) // Al<sub>2</sub>O<sub>3</sub>(1120). So, by patterning the MgO layer with conventional lithography, 45° grain boundaries in the SrTiO<sub>3</sub>(001) buffer layer could be placed in arbitrary positions.

Other combinations of substrates, seed layers, and buffer layers were later found to work. All of them had in common with the original concept that the materials were oxides with dissimilar crystal structures and large lattice mismatches so that  $\sqrt{2}$  times one lattice constant provided as close a match with a 45° in-plane rotation as could be achieved by cube-on-cube growth. While the virtue of this technique compared to bicrystal junctions is the ability to arbitrarily place grain boundary junctions on a mask, its disadvantage is that 45° is larger than the angle normally selected for bicrystal junctions. At this large angle,  $J_c R_n$  products at 77 K were on the order of just 10  $\mu$ V to 20  $\mu$ V and critical currents per unit junction width approximately 3  $\mu$ A/ $\mu$ m. Although some multilayer SQUIDS have been based on biepitaxial junctions, there has been little work on their integration with an HTS groundplane for digital circuits.

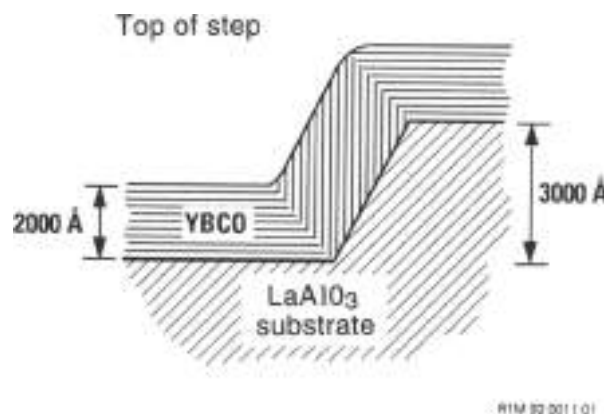
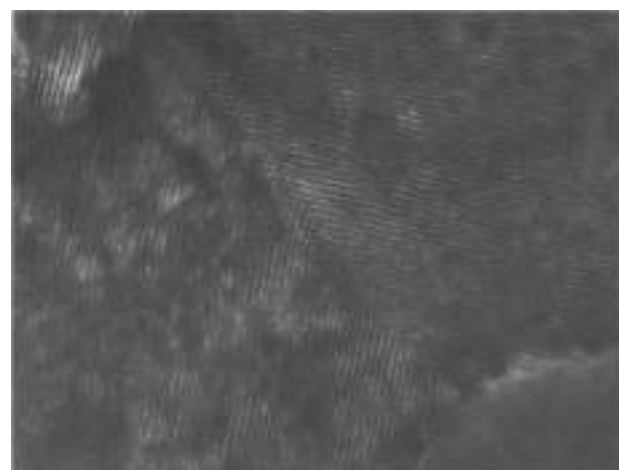


Figure 4. TEM cross-section of the upper, symmetric grain boundary in a step-edge grain boundary junction formed at a step in a LaAlO<sub>3</sub> substrate. (Courtesy of Claire Petteville-Hall of TRW.)

**Step-Edge Grain Boundary Junctions.** When a *c*-axis YBCO film is grown over a sharp step in a cubic (or nearly cubic) substrate or deposited insulator, the YBCO grows such that the *c* axis is perpendicular to the local principal crystalline axes of the substrate material (6), forming a pair of grain boundaries (at the top and bottom of the step), which behave electrically as one or two Josephson junctions. Figure 2(b) shows the simplest example, where each grain boundary (GB) is of the symmetric,  $\{103\} \times \{103\}$  type. Figure 4 shows a transmission electron micrograph of a cross section of such a boundary, formed at the top of a step in a single-crystal LaAlO<sub>3</sub> substrate. In practice the lower GB often consists of a mixture of orientations, including  $\{010\} \times \{001\}$ .

Such devices show  $I$ - $V$  characteristics with rather ideal RSJ shape, and they also show the signatures of true Josephson behavior, including critical current modulation with field (though often nonideal), Shapiro steps, and the emission of Josephson radiation. Values of  $J_c$  are fortuitously close to those required for digital and SQUID applications in the 65 K to 77 K range, with  $J_c R_n$  values of up to several hundred microvolts at 77 K (7). Electrical characteristics, and stability with thermal cycling, are improved by paying close attention to forming a sharp,

clean step—for example, by the use of ion milling with a very hard mask, such as amorphous carbon. While there have been reports of sets of such junctions with  $J_c$  spreads as narrow as 5% (7–9), typical spread values are 30% or above, making the long-term viability of this junction technique for complex circuits doubtful. For SQUIDs this is not an issue, and the simplicity of fabrication, along with the low measured noise of such junctions, makes them attractive. Their incorporation into multilayers is relatively straightforward.

The nature of the Josephson junction in this geometry is controversial. For example, GBs similar to those observed in step edge junctions have been deliberately fabricated in planar YBCO films, by the use of seed layers to control film orientation (see the section entitled "Bicrystal Grain Boundary Junctions"), and have been found to *not* exhibit weak link behavior. Also, while some of the most definitive work on these junctions has attributed the weak link to a (010)001 GB at the bottom of the step (6), recent measurements which probe the GBs individually (by use of a narrow YBCO lead, formed by shadowing, along the step face) suggest that it is the symmetric GB at the top which is the weakest.

#### Damage-Induced HTS Josephson Junctions

**Ion-Damaged Weak Links.** Weak links based on ion damage can be classed into two categories: (i) single-layer devices in which a focused ion beam (FIB) creates a weak link [Fig. 2(c)], and (ii) multilayer structures where ion surface damage of a base electrode produces a Josephson junction after deposition of a counterelectrode [Fig. 2(d)]. FIB junctions have been produced by a number of groups including Zani et al. (8), who used a 300 keV Si ion beam. The main attraction of the FIB approach is its relative simplicity, since only a single HTS film is required. However, because the typical FIB spot size is much greater than HTS coherence lengths, weak link behavior in these devices is presumably due to local variations in damage which lead to a parallel array of filamentary connections across the damaged region. Consequently, FIB-defined weak links often show non-ideal  $I$ - $V$  characteristics, as well as excess SQUID inductance. While this technology has produced high-resistance junctions and working SQUIDs up to 60 K, it has been largely superseded by other junction processes.

The second type of HTS Josephson junctions based on ion damage rely on ion bombardment at relatively low energies to create surface damage on a base electrode, which is then overlaid by an epitaxial counterelectrode. Work in this area has included room-temperature oxygen, argon and oxygen-fluorine-based plasma treatments (9), as well as low-energy Ar and Xe ion mill processing (10). The surface ion damage approach avoids the complication of a deposited interlayer and has produced high quality Josephson junctions with  $I_c R_n$  products up to 250  $\mu$ V at 77 K. However, the weak links working at 77 K also typically have current densities and resistances which are not well suited to digital and SQUID applications ( $J_c > 10^6$  A/cm<sup>2</sup> and  $R_n A < 2 \times 10^{-4}$   $\Omega$ -cm<sup>2</sup>). More recently it has been found that radio-frequency (RF) plasma exposure at 400–500 C can produce higher resistance weak links with  $J_c$  1- $\sigma$  spreads as small as 8% at 4.2 K (11).

**Electron-Beam-Damaged Weak Links.** Electron-beam-damaged junctions are fabricated by writing a line across a narrow, 2  $\mu$ m to 10  $\mu$ m wide bridge with an electron beam in

a transmission electron microscope with 120 keV or 350 keV beam energies (12). The electrons are thought to disorder oxygen ions located in the chains of copper and oxygen which are present in YBCO. The higher damage energy, 350 keV, results in a weak link that is stable at room temperature, whereas disorder caused by 120 keV electrons starts to heal even at room temperature. In either case, the evidence for oxygen disorder is mainly found in (a) the annealing behavior of these junctions which tends to a recovery of the initial, undamaged bridge properties over time with temperatures < 400 C and (b) energy barriers of 1.1 eV which are characteristic of the chain oxygen sites.

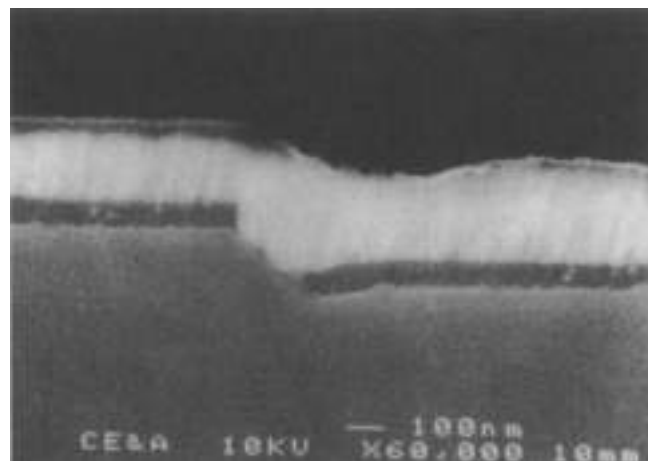
Targeted junction properties can be achieved by adjusting the total damage dose and the junction length. Critical current modulation in a magnetic field indicates that the damage is uniform compared to the scale of the junction width. The most sophisticated digital circuit demonstrations that have been performed without integrated HTS groundplanes have used this type of junction. A possible long-term limitation of any sequential process of writing to form junctions using either electron or focused ion beams is that writing time can become significant as circuit complexity and junction counts increase. The special equipment required, the lack of junction stability, and the fact that groundplane integration has not yet been demonstrated are other probable reasons why this technique has been used in university laboratories but has not been adopted by industrial labs.

**"Poisoned" Weak Links.** An alternative technique for weakening a small area is to dope an HTS film with a small patch of a material that degrades the superconducting properties. In the example of Simon et al. (13), a 10 nm thick, several-micron-wide Al stripe was patterned on a substrate before deposition of a 200 nm thick YBCO film. A weak link formed where a bridge patterned in the YBCO film crossed the Al stripe. The Al dopant depressed the critical current of the YBCO by two orders of magnitude.

This technique has the benefit of a simple fabrication process but has not been used in recent years. It is included here for completeness and to illustrate two of the factors that influence junction reproducibility: control of bridge length and control of interface thickness. Since the ideal length of a coupling region between YBCO banks is no greater than tens of nanometers, it is preferable that the length is determined by a reproducible scale such as a film thickness or step height rather than the width of a patterned line. The deposited interlayer junction configurations described below are designed to use these better-controlled length scales and to maintain a thinner interface layer between undamaged YBCO banks and a more weakly superconducting region than one can achieve by an interdiffusion process.

#### Deposited-Interlayer Josephson Junctions

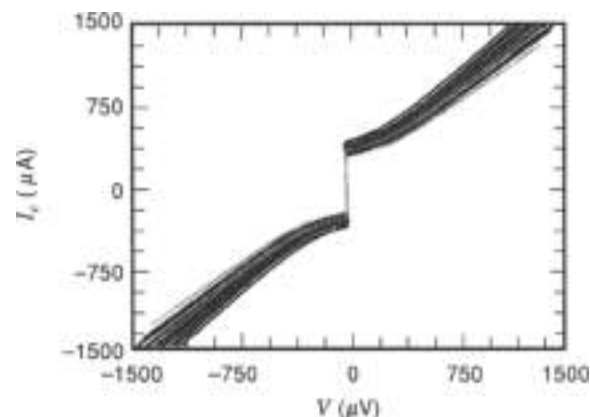
**Step-Edge SNS Junctions.** Like the step-edge grain boundary (SEGB) junction, the step-edge SNS junction is formed at a step, either in the substrate or a deposited insulator, but in this case the deposition is performed directionally, from behind the step, so that the YBCO is discontinuous. This discontinuity is bridged by an in situ deposited noble metal such as Ag (14,15), making the nominally SNS structure illustrated in Fig. 2(g). A cross-sectional scanning electron micrograph of an actual



**Figure 5.** Cross sectional electron micrograph of a step edge SNS junction of the type shown in Fig. 2(g), clearly showing the discontinuity of the YBCO film (dark) over the step, bridged by the Au-Ag alloy (light). Courtesy of Mark D'Iorio of Magnasensors Inc.

device is shown in Fig. 5, which clearly shows the YBCO to be discontinuous across the step. In practice the resistance of such a device is dominated not by the normal metal itself, but by the interfaces with the YBCO. By using a directional ion mill from behind the step, it is possible to remove most of the normal metal, leaving only that which is shadowed by the step. In this case,  $R_n$  can be as much as  $10 \Omega$ , and  $I_c R_n$  as high as several millivolts, at 4.2 K. Useful SQUIDs operating at 77 K, with either "flip-chip" or integrated pick-up coils, are routinely fabricated, with modulation voltage of several microvolts. Statistics for large numbers of junctions are not available to evaluate the critical current spreads. The directionality of the fabrication process means that this junction type does not scale well for complex circuit fabrication.

**Edge-Geometry Weak Links.** The HTS edge-geometry weak link is also known as an edge junction or ramp junction and is shown in Fig. 2(h). This structure typically consists of a c-axis-oriented HTS base electrode film overlaid by a thick insulator (not necessarily epitaxial) with an edge produced in the bilayer by ion milling or more rarely by wet etching. An epitaxial normal metal interlayer is deposited on the exposed edge followed by growth and patterning of the HTS counterelectrode. The SNS edge junction configuration is the most widely used HTS junction approach, because this geometry offers a number of advantages, including the fact that the critical superconductor/normal metal (SN) interfaces are located on the longer-coherence-length surfaces of the superconducting electrodes. The edge structure also enables the fabrication of very-small-area devices using conventional photolithography because one of the device dimensions is determined by the thickness of the base electrode film. Not surprisingly, however, successful fabrication of uniform sets of high-quality SNS edge junctions requires great care in base electrode edge formation. Shallow edges ( $<40 \mu\text{m}$ ) are used to avoid grain boundary formation in the counterelectrode, and the edge properties must be independent of edge orientation for ease of circuit layout. The most common approach to edge formation uses argon ion milling with rotating substrates and a tapered photoresist mask. Base electrode edge cleaning before growth of the nor-



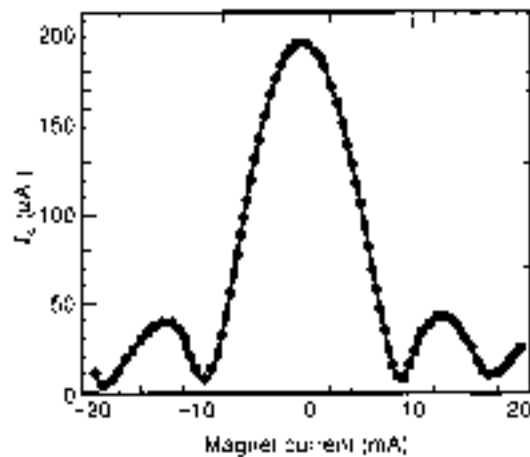
**Figure 6.**  $I$ - $V$  characteristics at 65 K for a chip with junction parameters suitable for SFQ logic. There are nineteen  $4 \mu\text{m}$  wide junctions with 50 Å Ce-doped YBCO interlayers and an average resistance of  $0.97 \Omega$  ( $1 - \alpha = 6\%$ ). The average  $I_c$  is  $327 \mu\text{A}$  ( $1 - \alpha = 13\%$ ) and the average  $I_c R_n$  product is  $315 \mu\text{V}$  ( $1 - \alpha = 9\%$ ).

mal metal and counterelectrode is also important and is usually done by low-energy ion milling or by etching in a dilute ( $<1\%$ ) solution of bromine in alcohol.

As with trilayer junctions (discussed below), the edge junction approach utilizes an all-epitaxial stack of base electrode, normal metal interlayer, and counterelectrode which places tight constraints on potential interlayers. Suitable interlayers must be lattice-matched to the HTS electrodes, grow without pinholes, and be chemically compatible with the superconductors at the elevated temperatures necessary for epitaxial growth. These requirements point to the use of materials with similar structures and compositions to YBCO such as  $\text{PrBa}_2\text{Cu}_3\text{F}_6$ ,  $\text{PBCO}$  or  $\text{YBaCu}_3$ ,  $\text{U}_0\text{O}_7$  (Ce-doped YBCO) (16), the two most commonly used interlayers. Note that the interlayers are often referred to as "normal metals," but, in fact range from semiconductors (PBCO) to superconductors operating above their transition temperature (Ce-doped YBCO).

$I$ - $V$  characteristics for a set of Ce-doped YBCO edge junctions produced at Northrop Grumman are shown in Fig. 6 (17). The normal metal layer is 70 Å of  $\text{YBaCu}_2$ - $\text{Ce}_0.2\text{O}_7$ . For this chip the average junction parameters, at 65 K, are  $I_c = 4.1 \times 10^5 \text{ A/cm}^2$  with  $1 - \alpha = 13\%$ ,  $I_c R_n = 315 \mu\text{V}$  with  $1 - \alpha = 9\%$ , and  $R_n A = 7.7 \times 10^{-7} \Omega\text{-cm}^2$  with  $1 - \alpha = 6\%$  ( $R_n = 0.97 \Omega$ ), values which are suitable for small-scale SFQ circuit applications. Note that this resistance value is surprisingly high for SNS junctions, as will be discussed in more detail in the section entitled "Advanced Issues." As shown in Fig. 7, the critical current modulation in edge-geometry weak links can approach the ideally expected  $I \sin \pi \phi / \pi$  behavior, indicating fairly uniform pair current transport through the normal metal layer (17). Good results have also been obtained for junctions utilizing PBCO and Ce-doped PBCO interlayers. For example, Verhoeven et al. (18) have produced Ce-doped PBCO edge junctions with  $I_c R_n$  products up to 5 mV at 4.2 K. The data suggest that transport in these devices takes place by resonant tunneling.

**Trilayer-Geometry Weak Links.** Trilayer-geometry Josephson junctions (Fig. 2(i)) consist of a layered sandwich structure of HTS, interlayer, and HTS epitaxial films, as well as associated wiring and insulator layers. This approach most closely resembles the highly successful LTS Nb-Al tunnel junction



**Figure 7.** Critical current modulation at 50 K for a 4- $\mu\text{m}$  wide SNS junction with La-YBCO base electrode, with the magnetic field normal to the substrate. The normal metal layer is 59 Å of Co-doped YBCO, and the junction resistance is 1.1  $\Omega$ .

process and consequently has attracted considerable attention. However, the trilayer process does require additional epitaxial insulator and HTS layers for wiring. Another potential disadvantage of this approach is the relatively large area of the devices, which can lead to unacceptably low values of  $R_n$  in many SNS processes, as well as a greater sensitivity to defects such as pinholes. Some of the earliest work in HTS deposited-interlayer weak links utilized *c*-axis oriented YBCO electrodes with a PBCO interlayer (19). Researchers at Varian have used a molecular beam epitaxy approach to engineer *c*-axis trilayers on a layer-by-layer basis (20). Most recent studies of trilayer junctions have focused on *a*-axis-oriented or (100)-oriented trilayers due to the longer superconducting coherence length parallel to the *a*-*b* planes, which in principle should lead to larger  $I_c R_n$  products. While it is not clear how an *a*-axis trilayer epitaxial wiring scheme will deal with the inherently lower  $I_c$  for wiring runs along the *c*-axis direction, some promising results have been obtained with this approach. Sato et al. (21) have produced (100)-oriented trilayer junctions with 350 Å Pr-BaCuO interlayers that exhibit RSJ  $I$ - $V$  characteristics at 50 K with  $J_c = 440 \text{ A/cm}^2$ ,  $1-\alpha = 38\%$ ,  $R_n A = 1.2 \times 10^{-7} \Omega\text{-cm}^2$ ,  $1-\alpha = 21\%$  (1.3  $\Omega$  for a  $3 \times 3 \mu\text{m}^2$  junction), as well as sensible magnetic field modulation.

**Planar or In-Line Junctions.** These junctions, illustrated in Fig. 2(j), consist of a narrow gap (<1  $\mu\text{m}$ ) in a YBCO film, bridged by a normal conducting film either situated below the YBCO or deposited on top after the gap is etched by, typically, a focused ion beam. The normal layer is usually Au or Ag (above the YBCO) or PBCO (above or below). While the normal metal coherence length of Au is long enough that gaps of order 1  $\mu\text{m}$  should support a critical current, the estimated values of  $\xi_c$  for PBCO and similar materials are so short that even devices as short as 0.1  $\mu\text{m}$  should not exhibit a critical current. The fact that they do has been used as evidence to suggest the existence of a "long-range proximity effect" in these materials, although the existence of this is controversial. For example, it is possible that the critical current is partly due to YBCO not fully removed from the gap.

There is little data on the reproducibility of these devices, but it is generally believed that they have little potential for

applications requiring many junctions. In addition, their characteristics are relatively nonideal, often exhibiting large excess current for example.

### Microbridges

If a superconducting bridge is made with a width comparable to the coherence length,  $\xi$ , it behaves as a Josephson junction. Since  $\xi$  is only a few angstroms in YBCO, this is not a practical way to make a Josephson device. However, if a bridge is wider than  $\xi$ , but still less than the penetration depth  $\lambda$ , it can behave in many ways as a Josephson junction, for example, exhibiting constant voltage steps in response to applied microwaves, in this case as a result of magnetic field vortices moving across the bridge in synchronization with the microwave field. As such they may be useful for SQUID applications, but not for SFQ circuits.

These devices have been fabricated both in the *a*-*b* direction [Fig. 2(k)], using focused ion beam etching for example, and in the vertical direction in a *c*-axis film [Fig. 2(l)]. The latter case is interesting in that the corresponding penetration depth is long enough (>1  $\mu\text{m}$  at 77 K) that conventional photolithography suffices. The *c*-axis microbridges have demonstrated encouraging  $I_c$  uniformity (13% for a few devices (22)), although critical currents are typically somewhat high, and resistances low, compared to typical application requirements, suggesting that more aggressive lithography may still be required. Low-inductance multilayer SQUIDs have been demonstrated, and the fabrication process is quite compatible with the needs of a multilayer process.

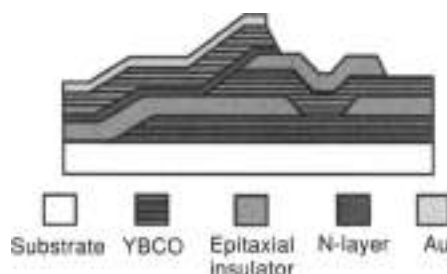
## CIRCUIT INTEGRATION

### Multilayer Circuit Requirements

Multilayer structures are required for many important applications of HTS junctions including digital circuits, SQUIDs with integrated pickup coils, voltage standards, and phase shifters, among others. Junction-specific requirements on  $I_c$  and  $R_n$  for some of these applications have already been discussed in the section entitled "Overview." It is desirable that these junction properties be independent of location in a multilevel structure—for example, on or off of groundplanes. This is a non-trivial constraint because junction performance is often intimately related to details of film microstructure which can be affected by growth over the underlying layers in a multilayer stack. Multilayer circuits also require good electrical isolation (>10<sup>4</sup>  $\Omega\text{-cm}$ ) between superconducting layers, using insulators with low dielectric constants and low losses, if high-frequency applications are the objective. High-critical-current-density vias and crossovers (>10<sup>6</sup> A/cm<sup>2</sup>) are essential for most HTS circuit applications. Digital circuits also call for the integration of superconducting groundplanes to produce the low-inductance SQUIDs and interconnects needed for SFQ logic. More complex circuits can require additional epitaxial insulators and wiring levels as well as integrated resistors.

### Fabrication Issues

A cross-sectional schematic view of the multilayer structure used to integrate edge SNS junctions on an HTS groundplane is shown in Fig. 8. A minimum of four mask levels and six



**Figure 8.** Schematic cross section of a multilayer process incorporating edge SNS junctions over an HTS groundplane. The horizontal shading of the YBCO layer indicates the direction of the copper oxide planes. All layers are epitaxial except for the Au, used for contacts.

epitaxial oxide film layers are needed for this process. Additional epitaxial film layers are often used such as buffer layers between the substrate and groundplane, or passivation layers above the groundplane, but these additional layers do not alter the basic processing steps.

The substrate selected for the most complex multilayer structures is single-crystal, perovskite-structure  $\text{NdGaO}_3$ , which is representative of the other candidate substrates. The (110) and (001) faces of  $\text{NdGaO}_3$  are virtually identical in providing a square two-dimensional lattice with approximately a 1% lattice mismatch for growth of *c*-axis-oriented epitaxial YBCO films.

The most important requirements for YBCO groundplane films are magnetic penetration depths,  $\lambda(T)$ , close to intrinsic values and smooth surfaces. These properties can be achieved, in principle, by films grown by a number of different techniques. In practice, pulsed laser deposition, sputtering, and co-evaporation are the three techniques most commonly used for multilayer film growth. The effect of the penetration depth on critical inductance is discussed in the next section. The smoothest films, those with approximately 1 nm root-mean-square (rms) roughness, have less than  $10^8/\text{cm}^2$  of the outgrowth that are commonly found in YBCO films. Most of the outgrowths are second-phase copper oxide particles that are 0.5 nm to 1.0 nm in diameter and grow higher than the film surface by a distance comparable to the film thickness.

As long as a copper oxide outgrowth does not form in a location where a junction is patterned, it does not have a deleterious effect on YBCO film properties. For applications of YBCO films requiring a single layer, such as bandpass radio frequency (RF) filters, no effort is made to minimize them. In fact, their presence is a sign that excess copper available at the growing film surface has been consumed, leaving behind a matrix of stoichiometric YBCO which has optimized properties. However, outgrowth density must be minimized by a slightly copper-deficient film composition for multilayer circuits to maintain electrical isolation between layers by the epitaxial insulators.

The epitaxial insulator that grows with the best edge coverage and smoothest surfaces is  $\text{SrTiO}_3$ . Its high dielectric constant, on the order of  $\epsilon = 500$ , is not a problem for low-frequency applications such as SQUID magnetometers. However, there is no consensus materials choice at present for high-speed digital circuits where a dielectric with  $\epsilon = 10$  to 30 is a practical maximum. Examples of relatively low- $\epsilon$  materials used as epitaxial insulators are  $\text{Sr}_2\text{AlNbO}_6$  and  $\text{Sr}_2\text{AlTaO}_6$ .

Since the integration level is very low at this stage of the development of HTS circuits, demands on lithography are min-

imal, and this is usually performed with contact lithography and standard resists. However, to avoid formation of step-edge grain-boundary junctions where interconnects cross steps in underlying films, an etch process is required that results in sidewalls that are sloped just 5° to 30° from the plane of the substrate—similar to the angles used for edge SNS junction fabrication. Tapered edges are obtained by ion milling with Ar or Ar/O<sub>2</sub> mixtures with the ion beam at an angle with respect to the substrate and the substrate rotating about its normal.

Because the surface of YBCO reacts with air and photoresist to form an amorphous layer of carbonates and hydroxides 2 nm to 10 nm thick, the surface must be cleaned after lithography and before a subsequent epitaxial film layer can be grown. In the case of edge SNS junctions shown in Fig. 8, the most important exposed surface is the edge cut in the base electrode which will serve as the template for growth of an N-layer and YBCO counterelectrode. Some combination of oxygen plasma ashing, blanket removal of the reacted layer by ion milling, and wet chemical etching, typically with bromine-alcohol mixtures, is performed to prepare surfaces for the next film deposition.

Figure 8 shows that a Au contact layer is typically used to complete this multilayer structure. The contact layer is normally deposited in situ—that is, after the YBCO counterelectrode layer has been deposited and cooled to room temperature in oxygen, but before it has been exposed to air. In situ Au deposition lowers contact resistance and improves adhesion. Other room-temperature film layers may be required for resistors and for electrical isolation of the resistor layers, but they do not require the same care that must be taken for film layers grown at high temperature.

### Groundplanes

As discussed in the section entitled "Application Requirements," digital circuits require low inductances, of order a few picohenries, to ensure that the  $LI_c$  product is about a flux quantum. The standard way to keep inductances low is to incorporate a superconducting ground plane above or below the active devices, which tends to confine the magnetic field in the relatively small volume between the two or more superconductor layers. Calculation of inductance is also simplified for this geometry since the problem is essentially reduced to counting squares and using the expression for the inductance per square of such a microstrip line:

$$L_{\infty} = \mu_0 dx \left( 1 + \frac{\lambda_1}{d} \coth\left(\frac{b_1}{\lambda_1}\right) + \frac{\lambda_2}{d} \coth\left(\frac{b_2}{\lambda_2}\right) \right) \quad (3)$$

where  $d$  is the insulator thickness,  $b_1$  and  $b_2$  the superconductor thicknesses, and  $\lambda$  is a factor that determines the field strength at the center of the finite width microstrip. The long penetration depth of YBCO, compared to Nb for example, means that the inductance per square is much higher. For example at 65 K we have  $\lambda_1 = \lambda_2 = \lambda = 0.23$  nm, so for  $d = b_1 = b_2 = 0.2$  nm and  $x = 1$ , we have  $L_{\infty} = 0.64$  pH. Contrast this with the case for Nb at 4.2 K, where  $\lambda = 0.04$  nm and  $L_{\infty} = 0.35$  pH. Thus the circuit-driven constraints of low inductance are particularly difficult to meet for HTS materials. Given the need for high  $I_c$  for thermal stability, the requirement that  $LI_c$  be about a flux quantum is particularly difficult to achieve, suggesting the need for novel gate layouts in HTS.

The integration of a YBCO groundplane has been demonstrated for several HTS junction types, including step-edge SNS, step-edge grain boundary, and edge SNS junctions. Measured inductances are consistent with other measurements of the penetration depth.

### Example Circuits

The key factor in determining the yield of working HTS circuits is the degree of control over junction critical currents—especially the on-chip spread. The more complex the circuit, the tighter the  $I_c$  spread must be, although the numerical relationship between  $I_c$  spread and circuit yield depends on the circuit margins. While LIS SFQ circuits often exhibit wide margins, of order 30% or more, extrapolating such values to HTS operation at 40 K to 86 K is controversial. If these margins do hold out, possibly in cases where bit-error rate is not a driving concern, then spreads of 15% should allow, based purely on statistical arguments, yields of 50% for circuits with up to about twenty junctions (24). In fact a number of such circuits have been demonstrated at up to about 65 K, at least at low speed, bearing out some optimism.

Demonstrated HTS gates or circuits include digital devices such as logic gates (OR, AND, etc.), set-reset flip-flops, toggle flip-flops, sampling circuits, and shift registers, as well as analog devices such as SQUID amplifiers. For digital devices, in which all of the junctions need to "work" for correct operation, the highest junction count is about 10. Figure 9 shows an example of a 10-junction circuit—the first stage of a low-power analog-to-digital converter, fabricated with edge SNS junctions over a ground plane.

So far there has been very little work to quantify the experimental margins of such HTS circuits. Should the more pessimistic estimates of HTS SFQ margins be borne out, then it may be necessary to rely on voltage-stair logic, where increased power dissipation will erase one of the major advantages of

superconductivity, or on multi-flux-quantum schemes, several of which have been proposed.

Manufacturability of HTS circuits will also require that chip-to-chip parameter spreads be well controlled. For a single circuit it may be feasible to tune the operating temperature to, for example, adjust the average critical current to the desired value—but this is clearly not practical for a system consisting of several separately manufactured superconducting circuits. Adjustment of overall circuit biases is a more practical solution to poorly targeted critical current values but will significantly increase costs due to the need for increased circuit testing and qualification. Thus ultimately it will be necessary to control chip-to-chip parameter reproducibility to the level of a few percent at most, which will be a significant challenge for these complex materials.

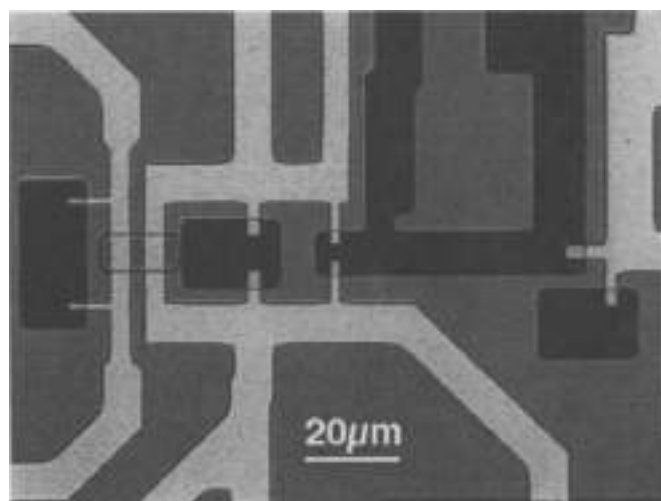
There is no single answer to the question of what is the best application for HTS Josephson junctions. However, a consensus of those working in the field believes that applications that utilize the analog precision of Josephson devices constitute the most promising niche where Josephson devices can surpass semiconductor circuits. Thus there is an emphasis on high precision analog-to-digital and digital-to-analog converters in such applications as radar, communications, and precision instrumentation. For example, an A-to-D converter with 20 bits of accuracy on a 10 MHz signal bandwidth would leapfrog semiconductor A-to-D converters, based on their historical rate of progress, by about 10 years. A potential application of such a converter is in radar, where it is desired to pick out a small target (airplane, missile) from a large "clutter" signal (rain, mountainous terrain, waves). Such a circuit should contain, depending on the details of the architecture, anywhere from several hundred to several thousand Josephson junctions. Assuming circuit margins of 30% (which may be optimistic), the required junction critical current spread for 50% yield, based on the calculations of Ref. 23, would be approximately 10% to 8%. Should the margins be reduced to, for example, 15% due to thermal noise issues, then the required spread would be more like 5% to 4%.

The issue of operating temperature is also of crucial importance, and is influenced by the potential circuit performance, as limited by thermal noise, and the availability, reliability, size, weight, and cooling power of the cryocooler. For example, a typical Stirling-cycle cooler with 0.3 Watts of heat lift at 4.2 K would have an input requirement of some 1500 W, and would weigh about 250 pounds. This power and weight probably rules out the application of HTS circuits in most airborne platforms. On the other hand a Stirling cooler with 4 W of heat lift at 77 K would only require about 100 W of input power, and should weigh only about 10 pounds, making airborne deployment much easier. Operating at 40 K, should thermal noise require it, might reduce the available heat lift to 0.4 W, which should still be sufficient circuits of several thousand junctions.

### ADVANCED ISSUES

#### Proximity Effect

When a superconductor and normal metal are brought into contact, Cooper pairs from the superconductor can diffuse into the normal metal. Due to phonon-induced pair breaking, the pair amplitude (also known as the superconducting order parameter or wavefunction) in the normal metal decays exponentially over



**Figure 9.** An example of a simple multilayer HTS circuit, a 1-bit analog-to-digital converter, fabricated with edge SNS junctions over an HTS groundplane. The lightest areas are the gold-coated counter electrodes, while the darkest are the base electrode. The process used to produce the circuit is extendible to much more complex circuits, provided that junction critical current uniformity can be improved sufficiently.



a decay length defined as the normal metal coherence length,  $\xi_n$ . In the clean limit where  $l_n$ , the mean free path in  $N$ , is much greater than  $\xi_n$ , the coherence length is given by

$$\xi_n = \frac{\hbar v_F}{2\pi kT} \quad (14)$$

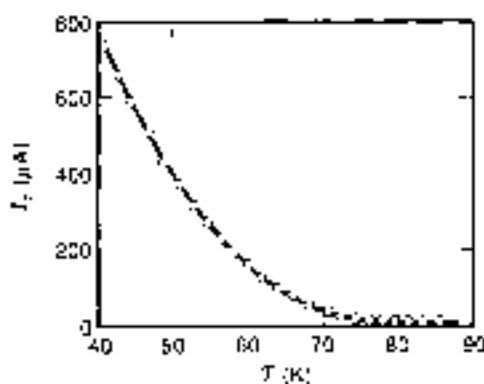
while in the dirty limit with  $l_n \ll \xi_n$ , the coherence length is

$$\xi_n = \frac{l_n}{\sqrt{\frac{\hbar D_n}{2\pi kT}}} \quad (15)$$

where the diffusion constant,  $D$ , equals  $v_F l_n/d$  and  $d$  is the dimensionality. Theories describing the details of the superconductor-normal-metal "proximity-effect" interaction have been developed for a variety of cases including back-to-back SN contacts—that is, the SNS weak link (24,25). In the SNS Josephson junction, pairs from each superconducting electrode "leak" into the normal metal interlayer and the overlap of the exponentially decaying pair amplitudes determines the strength of interaction between the superconductors. Consequently, the magnitude of the Josephson critical current scales as  $\exp[-L/\xi_n(T)]$ , where  $L$  is the normal metal bridge length. More specifically, in the dirty limit for long SNS bridges ( $L \gg \xi_n$ ) relatively close to  $T_c$  ( $T > 0.8T_c$ ), it is found that

$$I_c(T; L) \cong \frac{\pi}{2eR_n} \frac{|\Delta_c|^2}{kT_c} \frac{L}{\xi_n(T)} e^{-L/\xi_n(T)} \quad (16)$$

where  $\Delta_c$  is the superconducting gap at the superconductor-normal-metal interface. This equation indicates that the critical current of an SNS weak link should also vary exponentially with temperature, because of the  $(T)^{-2}$  temperature dependence of the dirty-limit normal-metal coherence length. The exponential length and temperature dependence of the critical current are the distinguishing signatures of true proximity effect devices. Indeed, there are a number of examples of HTS SNS devices which are largely consistent with proximity effect theory, most notably the junctions using Co- or Ca-doped YBCO as the normal metal layer. An example of exponential critical current dependence on temperature for a Co-doped YBCO SNS edge junction is shown in Fig. 10, along with a proximity theory



**Figure 10.** Critical current versus temperature for a YBCO:Co-YBCO edge-SNS junction. The solid line is a fit to the proximity effect theory of DeGennes. Despite the fact that the junction is non-ideal, in that it exhibits a large interface resistance, it still appears to exhibit behavior consistent with the proximity effect.

fit to the data (26). However, it is often found that HTS devices with a nominal SNS configuration do not show an exponential critical current dependence on temperature. In fact, such devices commonly exhibit a quasilinear temperature dependence, which may indicate that pinhole conduction through the normal metal is dominating the electrical characteristics (25).

#### Control of Resistance in SNS Devices

The normal state resistance of an SNS weak link is given by the sum of the normal metal resistance plus the resistance of each of the two SN interfaces. In the "ideal" SNS device the interface resistances are zero and the total device resistance is just  $R_n = \rho_n l_n/A$ , where  $\rho_n$  is the normal metal resistivity,  $L$  is the normal metal thickness, and  $A$  is the cross-sectional area. For typical values of these parameters in an SNS edge junction with a  $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$  normal metal layer at 65 K ( $\rho_n = 250 \mu\Omega\text{-cm}$ ,  $L = 100 \text{ \AA}$ , and  $A = 4 \times 0.2 \mu\text{m}^2$ ) we find  $R_n = 0.03 \Omega$ . In practice, such low values of resistance are often undesirable. For example, for SFQ digital applications at 65 K junctions are biased at a fixed current of order  $500 \mu\text{A}$  so that the available  $I_c R_n$  product with  $R_n = 0.03 \Omega$  is only  $15 \mu\text{V}$ , far less than the required value of approximately  $300 \mu\text{V}$ . Increasing SNS device resistances to a practical level requires adding interface resistance without degrading the inherent  $I_c R_n$  product. In principle, this can be done in at least two ways: (1) by incorporating an inhomogeneous interface resistance to reduce the effective device area; or (2) by producing a thin insulator at one SN interface to form an SIN structure. In practice, different groups have seen widely varying values of SNS resistance, ranging from the ideal but impractical case of very low  $R_n A$  (16) to the more technologically interesting case of high- $R_n A$  devices (27).

For SNS weak links using Co-YBCO as the normal metal, it has been found that the interface resistance is sensitive to a variety of factors including the base electrode material and the normal metal and counterelectrode deposition conditions (17). For example, SNS devices using  $\text{YBa}_2\text{Cu}_3\text{O}_7$  base electrodes grown by pulsed laser deposition (PLD) exhibit more than an order of magnitude lower resistance than devices with La-doped YBCO base electrodes ( $\text{YBa}_{1-x}\text{La}_x\text{Cu}_3\text{O}_7$ ) or  $\text{GdBa}_2\text{Cu}_3\text{O}_7$  or  $\text{NdBa}_2\text{Cu}_3\text{O}_7$  base electrodes. Varying the normal metal and counterelectrode growth parameters can also have a dramatic effect on device resistance: High-pressure PLD growth in an  $\text{Ar-O}_2$  atmosphere results in  $R_n A$  products over a factor of 10 smaller than for devices produced using the more conventional PLD deposition conditions in a pure oxygen background. While the detailed nature of the interface resistances is not understood at this point, the base electrode material dependence suggests that cation disorder (e.g., Y and Ba exchange) is affecting device resistance. The fact that the growth conditions of the normal metal also have a strong effect on SNS resistance indicates that defects "frozen in" in the early stages of normal metal growth may also play an important role in determining interface resistances. Because SNS interface resistances are strongly affected by a number of material and fabrication parameters, it is possible to control SNS device resistances over two to three orders of magnitude, with  $R_n A$  products ranging from  $0.03 \Omega\text{-}\mu\text{m}^2$  to more than  $10 \Omega\text{-}\mu\text{m}^2$ . Importantly, even in the relatively high  $R_n A$  limit required for SFQ applications ( $0.6 \Omega\text{-}\mu\text{m}^2$  to  $2 \Omega\text{-}\mu\text{m}^2$ ), Co-doped YBCO SNS devices incorporating significant interface resistance still behave like true proximity effect devices

(see, for example, Fig. 10) with parameter uniformity suitable for small-scale SFQ circuits.

### Limits on Reproducibility

Speculation on the origins of parameter spreads have led to many experiments in fabrication of edge SNS junctions. Surprisingly, the fabrication parameters which result in junction resistances greater than  $R_J = \rho_n L_j/A$  do not appear to systematically contribute to larger spreads in critical currents or other junction parameters. Similarly, the uniformity of current flow through a junction which can be inferred from  $I_c/R_j$  indicates that junctions with significant interface resistance maintain uniform current distributions.

Poor control over many fabrication parameters will certainly result in junction spreads worse than the state of the art. A good example is the roughness of YBCO base electrodes which gets transferred into the edge by patterning with Ar ion milling. While improvements from 10 nm to 2 nm rms surface roughness provide a measurable benefit for junction reproducibility, further improvements in smoothness have had a negligible effect. A second example is that junctions facing in all four in-plane directions sometimes exhibit a distribution of critical currents that is direction-dependent. However, when all processing steps are made isotropic, state-of-the-art junction uniformity can be achieved as easily in a set of junctions facing in four directions as in a set facing just one way.

These results have led us to examine defects that are intrinsic to YBCO. The role of oxygen disorder in YBCO has been investigated in several types of edge junction experiments. Decreasing the number of oxygen vacancies by plasma oxidation or annealing in ozone simply scales  $I_c$  for all treated junctions by a constant factor as large as five. Experiments in which orthorhombic YBCO electrodes were replaced by doped YBCO compounds which were tetragonal have been inconclusive in determining the possible role that twinning in YBCO might have on parameter spreads. Junctions were fabricated to face in (110) in-plane directions instead of the standard (100) directions to minimize the effects attributable to twinning, but no improvement in junction uniformity was observed.

Finally, the fact that similar best-case critical current spreads are observed for different junction fabrication processes using the same base electrode materials suggests that microstructural defects in the base electrode or base electrode edge are limiting  $I_c$  spreads. Further improvements in materials quality and edge formation techniques are expected to lead to improved junction spreads.

### CONCLUSIONS

Josephson junctions based on YBCO are the fundamental building blocks for a variety of superconducting electronics applications operating at temperatures  $>50$  K. The properties of individual junctions fabricated in a variety of configurations are sufficiently close to ideal Josephson behavior to meet application requirements. However, integration of junctions into multilayer circuits and demands on reproducibility of junction parameters when higher junction counts are needed have narrowed development efforts to a few promising configurations. Most of the current HTS circuit fabrication effort in industrial laboratories is based on edge SNS junctions which have been used for the most sophisticated and extensible digital circuit

demonstrations. Further incremental improvements in the uniformity of these junctions to  $1-\sigma$   $I_c$  spreads less than 10% will permit medium-scale integrated circuit fabrication. A parallel effort, mainly by university researchers, is exploring higher-risk alternative junction configurations intended to circumvent some of the limitations to junction uniformity that may exist for edge junctions.

### BIBLIOGRAPHY

1. T. Van Duzer and C. W. Turner, *Principles of Superconductive Devices and Circuits*, New York: Elsevier, 1981.
2. D. Rubbes et al., The ac Josephson effect in constructions engraved in bulk  $\text{YBa}_2\text{Cu}_3\text{O}_7$  and its SQUID operation at 77 K, *Nature*, **334**: 151-153, 1988.
3. B. H. Moockly, D. K. Lattrop, and B. A. Huiteman, Electromigration study of oxygen disorder and grain boundary effects in YBCO thin films, *Phys. Rev. B*, **47**: 460, 1993.
4. D. Damas et al., Orientation dependence of grain-boundary critical currents in  $\text{YBa}_2\text{Cu}_3\text{O}_7$  bicrystals, *Phys. Rev. Lett.*, **61**: 210-222, 1988.
5. K. Char et al., Bi-epitaxial grain boundary junctions in  $\text{YBa}_2\text{Cu}_3\text{O}_7$ , *Appl. Phys. Lett.*, **59**: 733-735, 1991.
6. K. Herrmann et al., Characterization of  $\text{YBa}_2\text{Cu}_3\text{O}_7$  step edge junctions, *Supercond. Sci. Technol.*, **4**: 583-590, 1991.
7. J. Laine et al., Characteristics of high performance YBCO step-edge junctions, *Appl. Phys. Lett.*, **61**: 1128-1130, 1992.
8. M. J. Zuni et al., Focused ion beam high  $T_c$  superconductor dc SQUIDs, *Appl. Phys. Lett.*, **59**: 234-236, 1991.
9. R. B. Lombard et al., All high  $T_c$  edge junctions and SQUIDs, *Appl. Phys. Lett.*, **56**: 686-689, 1990.
10. B. D. Hunt et al., High Temperature Superconductor Josephson Weak Links, in S. L. Raider et al. (eds.), *Low Temperature Electronics and High Temperature Superconductivity*, Electrochemical Soc. Proc., vol. 93-22, Pennington, NJ: The Electrochemical Society, 1993, pp. 463-472.
11. B. B. Moockly and K. Char, Interface Engineered High  $T_c$  Josephson Junctions, in H. Koch and S. Karapeteds, (ISSE '97) *Int. Supercond. Electron. Conf. Extended Abstracts*, vol. 1, Bonn, Germany: PTB, 1997, pp. 8-10.
12. S. K. Tolpygo et al., High quality  $\text{YBa}_2\text{Cu}_3\text{O}_7$  Josephson junctions made by direct electron beam writing, *Appl. Phys. Lett.*, **69**: 1696-1698, 1996; A. J. Pausz et al., Electron beam damaged high  $T_c$  junctions stability, reproducibility and scaling laws, *IEEE Trans. Appl. Supercond.*, **5**: 2410-2413, 1995.
13. R. W. Simon et al., Progress Towards a YBCO Circuit Process, in R. McConnell and S. A. Wolf (eds.), *Science and Technology of Thin Film Superconductors II*, New York: Plenum, 1990, pp. 549-560.
14. B. H. Ono et al., High  $T_c$  superconductor-normal metal-superconductor Josephson microbridges with high-resistance normal metal links, *Appl. Phys. Lett.*, **58**: 1126-1128, 1991.
15. M. S. Dilorio et al., Manufacturable low-noise SQUIDs operating in liquid nitrogen, *Nature*, **354**: 513-515, 1991.
16. K. Char, L. Antognazza, and T. H. Geballe, Properties of  $\text{YBa}_2\text{Cu}_3\text{O}_7$ ,  $\text{YTiO}_2\text{Co}_2\text{FeCo}_{0.2}\text{O}_{7-x}$ ,  $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$  edge junctions, *Appl. Phys. Lett.*, **65**: 904-906, 1994.
17. B. D. Hunt et al., High-resistance HTS SNS edge junctions, *Appl. Supercond.*, 1996, in press.
18. M. A. J. Verhulst et al., Ramp-type junction parameter control by Ga doping of  $\text{PrBa}_2\text{Cu}_3\text{O}_7$  barriers, *Appl. Phys. Lett.*, **68**: 846-850, 1996.
19. C. T. Rogers et al., Fabrication of heteroepitaxial  $\text{YBa}_2\text{Cu}_3\text{O}_7/\text{Pt}$

- $\text{PrBa}_2\text{Cu}_3\text{O}_{7-x}$ ,  $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$  Josephson devices grown by laser deposition, *Appl. Phys. Lett.*, **55**: 2032-2034, 1989.
- 20 J. N. Eckstein, I. Buzdin, and G. F. Vershup, Atomic layer-by-layer engineering of high- $T_c$  materials and heterostructure devices, *MRS Bull.*, **19**(2): 44-50, 1994.
- 21 H. Sato, S. Goen, and H. Almb, Improvement of junction properties of  $\text{YBaCuO}/\text{PrBaCuO}/\text{YBaCuO}$  trilayer Josephson junctions, *IEEE Trans. Appl. Supercond.*, **7**: 2510-2513, 1997.
- 22 S. W. Goodyear et al., Vertical c-axis microbridge junctions in  $\text{YBa}_2\text{Cu}_3\text{O}_7/\text{PrBa}_2\text{Cu}_3\text{O}_7$  thin films, *IEEE Trans. Appl. Supercond.*, **6**: 3143-3146, 1996.
- 23 D. L. Miller, J. N. Prezhlyak, and J.-H. Kang, Margins and yields of SFQ circuits in HTS materials, *IEEE Trans. Appl. Supercond.*, **3**: 2728-2731, 1993.
- 24 P. G. DeGennes, *Rev. Mod. Phys.*, **36**: 225, 1964.
- 25 K. A. Delin and A. W. Kleinsasser, Stationary properties of high-critical temperature proximity effect Josephson junctions, *Supercond. Sci. Technol.*, **8**: 227-239, 1996.
- 26 M. G. Forrester, unpublished, 1997.
- 27 B. D. Hunt et al., High- $T_c$  SNS edge junctions and SQUIDs with integrated groundplanes, *Appl. Phys. Lett.*, **68**: 3605-3607, 1996.

BRIAN D. HUNT  
Centre de Recherches en Physique  
des Plasmas  
MARCUS C. FOURNIER  
Conductus, Inc.  
JOHN TALLARICH  
Northrop Grumman Science and  
Technology Center

## HTS PROCESSING: BULK, THIN FILM, AND WIRES

Superconductors are a class of materials possessing two unique properties: the complete loss of electrical resistivity below a transition temperature called the critical temperature ( $T_c$ ) and the expulsion of magnetic flux from the bulk of a sample (diamagnetism) in the superconducting state. The latter property is also known as the Meissner-Ochsenfeld effect or more commonly, the Meissner effect (1). At temperatures above  $T_c$ , these materials possess electrical resistivity like ordinary conductors, although their normal state properties are unusual in many aspects. This abrupt change from normal conductivity to superconductivity occurs at a thermodynamic phase transition determined not only by the temperature but also by the magnetic field at the surface of the material and by the current carried by the material. Several metals and metallic alloys exhibit superconductivity at temperatures below 22 K, and will be henceforth called low temperature superconductors (LTS). In 1950, superconductivity was explained as a quantum mechanical phenomenon by the London phenomenological theory (2). Later, the two-fluid phenomenological model explained the electronic structure of a superconductor as a mixture of superconducting and normal electrons, with the proportion of superconducting electrons ranging from zero at the onset of superconductivity to 100% at 0 K (3). In 1966, Bardeen, Cooper, and Schrieffer's (BCS) theory explained that superconductivity was the result of the formation of electron pairs of opposite spins (known as Cooper pairs), primarily owing to electron-phonon coupling (4). The BCS theory proved to be the most complete

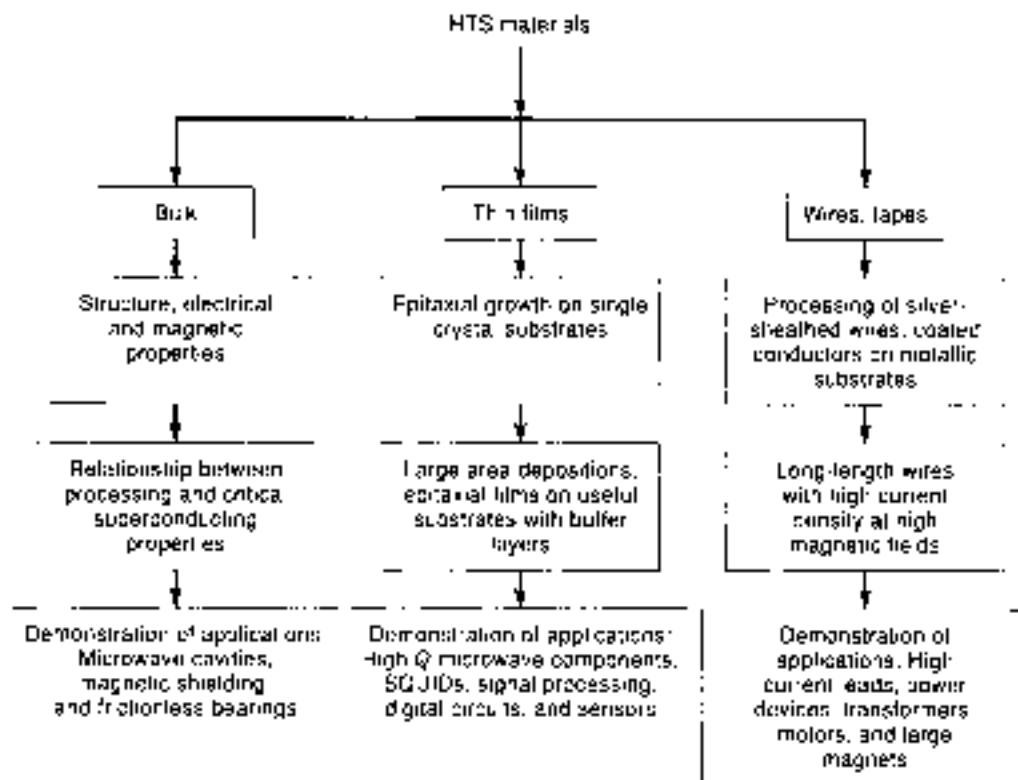
theory for explaining the superconducting state and the normal state of LTS materials. A major development in superconductors was the discovery by Josephson in 1963 that Cooper pairs show macroscopic phase coherence, and that such pairs can tunnel through a thin insulating layer sandwiched between two superconducting layers (the superconductor-insulator-superconductor (NIS) junction known as the Josephson junction (5). This effect, called the Josephson effect (5), caused a flurry of activity in the fields of high-speed computer logic and memory circuits in the 1960s and 1970s, since it can be used to make high-speed low-power switching devices. Problems were encountered in the mass fabrication of Josephson junctions for complex systems such as digital computers. Although the applications of LTS for electrical applications and electronics were demonstrated, the cost of cooling was too high for the commercial development of LTS.

The era of high-temperature superconductors (HTS) began in 1986 when two IBM Zurich researchers, K. A. Muller and J. G. Bednorz, reported the occurrence of superconductivity in a lanthanum barium copper oxide (LaBaCuO) at 30 K (6). Soon after, M. K. Wu, P. W. Chu, and their collaborators at the University of Alabama and University of Houston (7), respectively, announced the discovery of 90 K superconductivity. Since these two historic discoveries, there has been substantial progress in HTS technology. Several new families of cuprates including  $\text{BaSrCaCuO}$  (8),  $\text{TlCaBaCuO}$  (9), and  $\text{HgCaBaCuO}$  (10) have been found to be superconducting above 90 K. These discoveries make feasible electrical and electronics applications at temperatures above the boiling point of liquid nitrogen (77 K). Cuprate superconductors with a  $T_c$  value higher than 30 K have been classified as high-temperature superconductors. The obvious advantage of using liquid nitrogen rather than liquid helium for cooling is its higher heat of vaporization, which not only simplifies the design of cryostats but also the cost of cooling. Furthermore, liquid nitrogen (at \$0.25/liter) is more than an order of magnitude cheaper than liquid helium (at \$5/liter). Progress made in cryocoolers has made feasible HTS applications in electrical wires, magnets, and electronics. The excitement and challenges posed by these HTS materials have touched multiple disciplines, such as physics, chemistry, material science, and electrical engineering. Tremendous progress has been made in the application of HTS materials in such areas as Superconducting Quantum Interference Devices (SQUIDs), passive microwave devices, and long-length wires, as illustrated in the road map for the HTS technology, shown in Fig. 1. Better-quality materials emerging from refined processing methods have made it possible to separate the intrinsic properties of HTS from its extrinsic properties. The interrelationships of processing with structural, physical, electrical, and magnetic properties continues to be an area of intensive scientific research. In this article, we provide an overview of important high-temperature superconducting materials, their properties, and promising procedures for synthesizing bulk, thin film and wire forms of HTS conductors for engineering applications.

## PROPERTIES OF HIGH-TEMPERATURE SUPERCONDUCTORS

### Magnetic Properties

The critical temperature ( $T_c$ ) is defined as the temperature below which a superconductor possesses no dc electrical

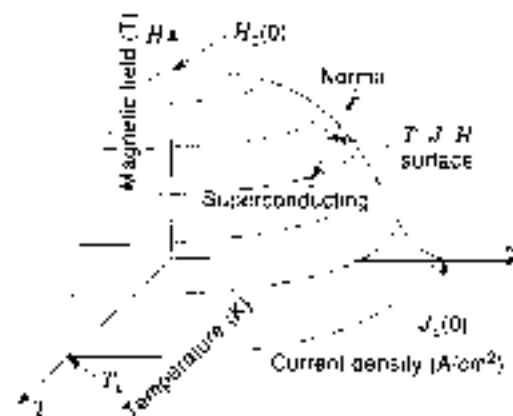


**Figure 1.** A generalized road map for the HTS technology. YBCO 123 compound has been the most studied material among the HTSs. Most applications have been demonstrated with the YBCO superconductor. Most applications in HTS wires have been demonstrated using BSCCO.

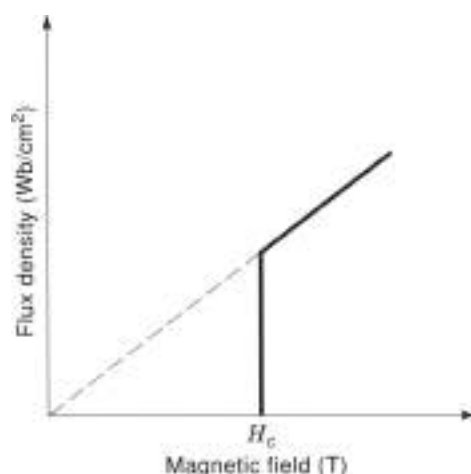
resistivity or the temperature below which a superconductor exhibits diamagnetic behavior. The useful range of operation of a superconductor is typically below  $0.6 T_c$ , as other important superconducting properties are enhanced below this temperature. Critical current density ( $J_c$ ) is one of the most important superconducting properties for engineering applications. It is an estimate of maximum current density (current per cross-sectional area of the conductor) a superconductor can support before becoming a normal conductor. The critical field ( $H_c$ ) for a superconductor is the maximum magnetic field below which a superconductor exhibits diamagnetic behavior and above which the superconductivity is quenched. The  $T_c$ ,  $H_c$ , and  $J_c$  parameters define a point in three-dimensional space, and for superconductors these points span a volume as shown in Fig. 2.  $T_c$ ,  $H_c$ , and  $J_c$  values are relatively low in type I superconductors. Type II superconductors are generally suitable for most electrical and electronic applications because of higher  $T_c$ ,  $H_c$ , and  $J_c$  values (11). In type I superconductors, the flow of shielding current in the superconductor is restricted to a thin layer from the surface, called the penetration depth ( $\lambda$ ), when the magnetic field is below  $H_c$ . The penetration depth is very small, near 0 K, and increases dramatically as the temperature approaches  $T_c$ . The penetration depth at 0 K ranges from 100 Å to 1500 Å for type I materials (11). Above the critical field, the magnetic field completely penetrates a type I superconductor, quenching superconductivity, as shown in Fig. 3. Figure 3 also illustrates the  $T$ -dependence of resistivity in a superconductor at  $T > T_c$ .

In type II superconductors,  $H_{c1}$  represents the lower critical field above which magnetic flux penetrates a superconductor to form a mixed state in which superconducting and normal

electrons coexist. When  $H > H_{c2}$ , the upper critical field, superconductivity is largely confined to the surface of the material. In the mixed state, magnetic flux penetrates through small tubular regions on the order of the coherence length ( $\xi$ ) (a length scale that characterizes superconducting electron pair coupling), called vortices (or flux tubes), with each vortex containing one quantum of flux,  $\phi_0$  (12). Abrikosov, in his study of type II superconductors determined that  $\phi_0 = hc/2e$ , where  $h$  is Planck's constant and  $e$  is the electronic charge (12). The vortices form a periodic lattice called the Abrikosov vortex lattice. The resistivity of a superconductor may be vanishing in the mixed state, provided the vortices are pinned or trapped.

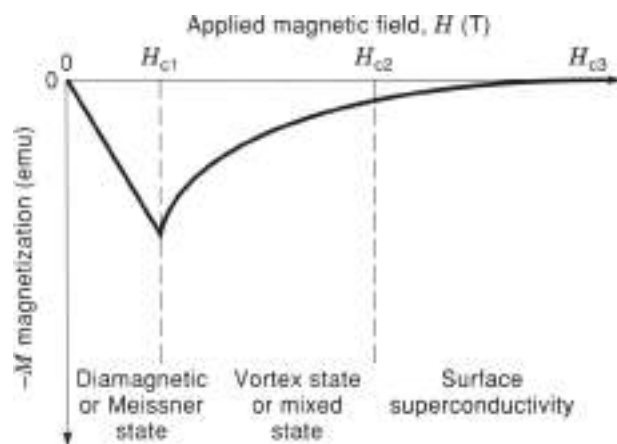


**Figure 2.** Temperature-current magnetic field ( $T$ - $J$ - $H$ ) three-dimensional surface defining the operating limits for a superconductor.

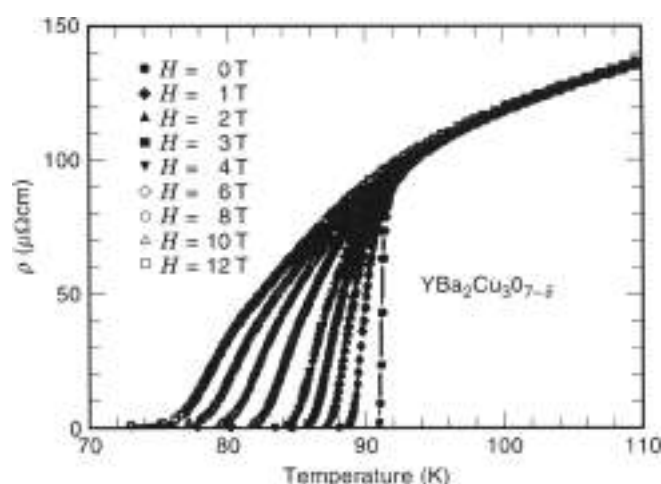


**Figure 3.** A type I superconductor's magnetic flux density versus applied magnetic field. A type I superconductor will not have any flux enclosed in the bulk below the critical temperature. Above the critical field, the applied field completely permeates in the bulk of the material.

As the applied magnetic field ( $H_a$ ) approaches  $H_{c2}$ , the number of vortices increases until there can no longer be any more addition of vortices, at which point the material becomes a normal conductor. Figure 4 shows the magnetic properties of type II superconductors. In the mixed state, each vortex resides in a normal region, which is separated by superconducting regions. The vortices experience three different types of forces: one is the Lorentz force due to the flow of external current, experienced in the direction of and proportional to the vector product of the current and the vortex field. Second is the force of repulsion from other vortices, and third is the pinning force from metallurgical defects. The Lorentz force causes motion of vortices (also called the flux flow). The vortex motion produces an opposing electric field to the flow of current, essentially contributing to ohmic losses. However, vortices can become trapped or pinned (called flux pinning) at metallurgical defects, in secondary phases or at impurity sites. Owing to the repulsive forces between each vortex, pinning a few vortices may lead to a frozen vortex lattice, or



**Figure 4.** The magnetic behavior of a type II superconductor, showing the Meissner state behavior below  $H_{c1}$ , mixed state behavior between  $H_{c1}$  and  $H_{c2}$ , and the surface superconductivity between  $H_{c2}$  and  $H_{c3}$ .



**Figure 5.** Temperature dependence of resistivity for an epitaxial c-axis-oriented YBCO thin film at different magnetic fields, with the field parallel to the c-axis. Courtesy of Ref. 13.

a lossless state, true only for direct currents and low-frequency alternating currents. The frozen vortex lattice occurs only below a critical field called the irreversibility field ( $H_{ir}$ ). Some of the trapped vortices could remain in the superconductor, contributing to a hysteretic behavior when an alternating current is applied, which produces further dissipation of energy. The trapped flux is analogous to remanent flux in ferromagnetic materials. While pinning centers tend to prevent the movement of vortices, there is a tendency for vortices to jump over the pinning defects. This phenomenon is called flux creep. At  $H_a > H_{ir}$ , vortices will move, causing additional energy dissipation. Resistive transitions in a superconducting sample in applied magnetic fields are shown in Fig. 5. Transitions at higher fields clearly show additional ohmic losses (13). Although the upper critical field is generally higher in type II materials, the limiting field is the irreversibility field, which is an order of magnitude lower than  $H_{c2}$ . Owing to the complex nature of the cuprate ceramic superconductors, and operation at higher temperatures, the ac losses in HTS materials are generally higher than in LTS materials (14, 15).

### Structural Properties

The presence of one or more copper oxide ( $\text{CuO}_2$ ) planes in the unit cell is a common feature of all HTS materials, also referred as cuprate superconductors. The most popular cuprate materials are  $\text{Y}_1\text{Ba}_2\text{Cu}_3\text{O}_8$  (henceforth referred to as YBCO),  $\text{Bi}_2\text{Sr}_2\text{Ca}_{n-1}\text{Cu}_n\text{O}_{2n+2}$  (where  $n = 2, 3$ ) BSCCO (henceforth referred to as Bi2212, Bi2223 for  $n = 2$ , and  $n = 3$ , respectively),  $\text{Tl}_2\text{Ba}_2\text{Ca}_{n-1}\text{Cu}_n\text{O}_{2n+4}$  (henceforth referred to as Tl2201, Tl2212, and Tl2223), and  $\text{HgBa}_2\text{Ca}_{n-1}\text{Cu}_n\text{O}_{2n+2}$  (where  $n = 1, 2, 3$ ). Table 1 lists the well developed cuprate superconductors, their superconducting properties, and important applications demonstrated to date. In YBCO, there are two square planar  $\text{CuO}_2$  planes stacked in the c-direction, separated by an intercalating layer of barium and copper atoms and a variable number of oxygen atoms. The conventional wisdom is that the  $\text{CuO}_2$  planes are the conduction channels of superconductivity, whereas the intercalating layers provide

**Table 1. Promising HTS Materials, Properties, and Applications**

| Material   | $T_c$     | $J_c(0)$ and $J_c(H)$<br>[A/cm <sup>2</sup> ]  | Applications                                    |
|--|-----------|--|---|
| Bulk YBa <sub>2</sub> Cu <sub>3</sub> O <sub>7-x</sub> (123) YBCO (melt grown)                   | 90 K      | 10 <sup>6</sup> at 77 K ZF<br>10 <sup>6</sup> at 77 K 10T  | HTS cavities<br>Bearings, EM shield             |
| Thin film YBCO on crystalline substrates   | 90 K      | 10 <sup>6</sup> at 77 K ZF<br>10 <sup>6</sup> at 77 K 5T   | SQUIDs, microwave<br>Electronics                |
| T <sub>1</sub> Ba <sub>2</sub> CaCu <sub>3</sub> O <sub>7-x</sub> (2121) thin film               | 87±100 K  | 2 × 10 <sup>6</sup> at 77 K ZF<br>1 × 10 <sup>6</sup> at 77 K 0.1T   | Microwave<br>Electronics                        |
| T <sub>2</sub> Ba <sub>2</sub> Ca <sub>2</sub> Cu <sub>3</sub> O <sub>7-x</sub> (2223) thin film | 117±123 K | 2 × 10 <sup>6</sup> at 77 K ZF<br>5 × 10 <sup>6</sup> at 77 K 1T   | Microwave<br>Electronics                        |
| Bi <sub>2</sub> SrCa <sub>2</sub> Cu <sub>3</sub> O <sub>7-x</sub> wires/tapes                   | 80 K      | 10 <sup>6</sup> at 77 K ZF<br>10 <sup>6</sup> at 77 K 1T<br>10 <sup>6</sup> at 4 K ZF<br>10 <sup>6</sup> at 4 K 1T | Low field HTS wires for<br>20±30 K applications |
| Bi <sub>2</sub> Sr <sub>2</sub> Ca <sub>2</sub> Cu <sub>3</sub> O <sub>7-x</sub> wires/tapes     | 110 K     | 5 × 10 <sup>6</sup> at 77 K ZF<br>5 × 10 <sup>6</sup> at 77 K 1T   | Magnets, current leads,<br>SMES                 |

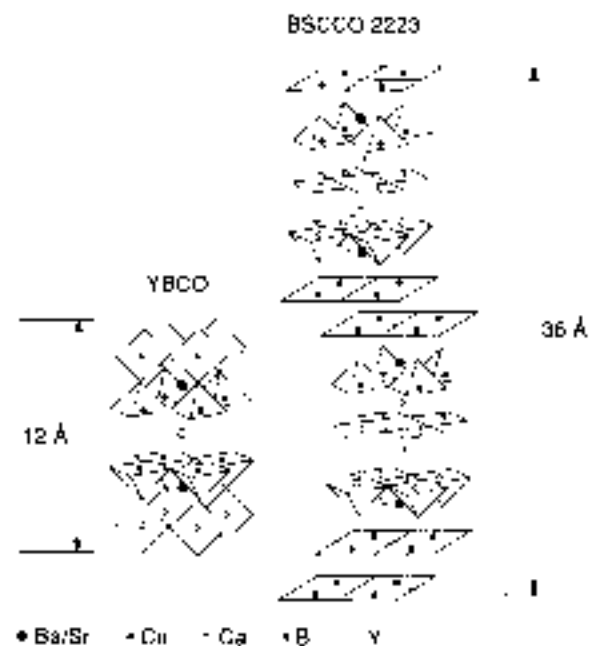
carriers or act as charge reservoirs necessary for superconductivity, although this view is not shared universally (16). The charge density, the number of superconducting charge carriers per unit volume, is determined by the overall chemistry of the system and by the charge transfer between the CuO<sub>2</sub> planes and the CuO chains. The charge density in a HTS material (10<sup>21</sup>/cm<sup>3</sup>) is two orders of magnitude lower than conventional LTS (10<sup>23</sup>/cm<sup>3</sup>). Remarkably, the oxygen content in the system changes the oxidation states of the copper chain atoms, which, in turn, affects their ability for charge transfer, charge density, and superconducting properties. Depending on the oxygen content, the YBCO 123 material could have a nonsuperconducting tetragonal ( $a = b \neq c$ ) phase or a 90 K superconducting orthorhombic ( $a \neq b \neq c$ ) phase. When fully oxygenated, YBCO possesses an orthorhombic unit cell with typical dimensions of  $a = 3.86$  Å,  $b = 3.58$  Å, and  $c = 12.0$  Å, and a  $T_c = 90$  K. Figure 6 shows the crystal structure of YBCO 123 and BSCCO 2223 superconductors, showing the conduction layers and the binding layers in each case.

BSCCO superconductor contains a weakly bonded double BiO layer that separates the CuO<sub>2</sub> planes. The Bi2223 structure is part of a family of several other HTS compounds in which Bi is replaced by Tl or Hg (with different oxygen coordination) and partially by lead. In some cases, the double layer of these metal-oxide layers can be reduced to a single layer, yielding another family of superconductors such as 1212 or 1223 [e.g., (Tl<sub>1-x</sub>Pb<sub>x</sub>)Ba<sub>2</sub>(As<sub>2-x</sub>Bi<sub>x</sub>)Cu<sub>3</sub>O<sub>7-δ</sub>, henceforth is referred as Tl1223]. The single-layer compounds offer strong flux pinning and low intrinsic defect structure compared to the double-layer compounds. The presence of a weakly bonded BiO layer is crucial for their superconducting properties in the Bi2212 and 2223 systems. The mechanical properties of Bi2212 and 2223 are micaceous, mica- or clay-like, and they have highly anisotropic growth rates along the  $ab$ -plane and  $c$ -direction. The latter is important to process long-length wires and to enhance the electromagnetic connectivity of the grains, flux making high transport current densities over long lengths possible (17). This is mainly due to the weak interlayer bonding of the BiO layer with the CuO<sub>2</sub> planes. On the other hand, the weak interlayer bonding of the BiO layer also leads to intermixing of Bi2212 and Bi2223 phases. In spite of this weakness, better grain connectivity and micaceous crystalline morphology are attractive for developing long-length HTS wires. Tl2212 and 2223

compounds have a structure similar to the Bi2212 and 2223 compounds with (TlO) double layers replacing the BiO layers.

### Other Important Properties

All the popular HTS materials possess a fundamental limitation, crystalline anisotropy, i.e., they possess different structural and electrical properties in different directions. Superconducting properties such as critical current density ( $J_c$ ) and critical magnetic field ( $H_c$ ) along the  $ab$ -planes ( $x$  or  $y$ ) are superior to those along the  $c$ -axis ( $z$ ). A major challenge for researchers has been to develop textured samples to take advantage of high  $J_c$  in the  $ab$ -planes. The HTS materials also exhibit higher penetration depths compared with LTS materials. The penetration depth is an important parameter for high-frequency applications of superconductors. It is defined as the depth through which a magnetic field penetrates into a superconducting sample and decays to 1/e of the field at the surface. The



**Figure 5.** The crystal structure of YBCO 123 and BSCCO 2223 compounds, showing the conducting layers, and the binding layers.

magnetic field decays in the form of  $H = H_0 e^{-x/\lambda}$ , where  $H_0$  is the field at the surface,  $x$  is the depth through the sample, and  $\lambda$  is the penetration depth of the superconductor, analogous to the skin depth in conventional electrical conductors. Penetration depth ( $\lambda$ ) increases with temperature. Penetration depth is a frequency independent parameter in contrast to the frequency-dependent skin-depth of normal conductors. This means that little or no dispersion will be introduced in superconducting components, and that it will be negligible up to frequencies as high as tens of gigahertz, in contrast to dispersion present in normal metals. Furthermore, lower losses in superconductors lead to a reduction in physical size, and this feature represents another advantage for HTS thin-film based circuits. Compact delay lines, filters, and resonators are possible with a high-quality factor ( $Q$ ) due to low-conductor losses (see Superconducting filters and passive components). The challenge in processing involves developing HTS materials with smooth-surface morphology to minimize high-frequency ac conductor losses. The extremely short coherence length in HTS materials ( $\sim 30$  Å along the  $ab$ -plane) increases the difficulty of making Josephson junctions (see TUNNELING AND JOSEPHSON JUNCTIONS).

#### PROCESSING OF HIGH-TEMPERATURE SUPERCONDUCTORS

Although  $T_c$  and  $H_c$  values of a superconductor are generally intrinsic properties of a specific material,  $J_c$  and  $\lambda$  values, on the other hand, are a function of sample microstructure and can vary by several orders of magnitude owing to various processing techniques. A high  $J_c$  in large magnetic fields is desirable for electrical applications of HTS. In general,  $J_c$  greater than  $10^5$  A/cm<sup>2</sup> at 77 K is required for wires, tapes, and magnets for high magnetic field applications (17). Although high  $J_c$ s in HTS are achieved in bulk, thin films, and wires, several processing-related problems need to be addressed. These problems relate to grain boundaries, metallurgical defects, flux pinning mechanisms, and flux creep. Grain boundaries are interfacial regions between adjacent grains and could contain impurity phases, normally conducting or insulating, extending beyond the coherence length of these superconductors. These grain boundaries create superconductor-normal conductor or insulator-superconductor Josephson junctions, the critical currents of which are much lower than in a homogeneous superconductor, and hence they are called as weak links. The grain boundaries are particularly a problem in polycrystalline bulk superconductors because they affect the bulk critical current density. Grain boundaries also contribute to surface resistance of a superconductor at high frequencies because they increase the residual ac losses. Processing of HTS materials to improve  $J_c$  requires ways to minimize the grain-boundary effects and enhance the flux pinning sites in the superconductors. Addition of a small percentage of silver into HTS materials improves the critical currents of grain boundaries. Also, texturing of a bulk superconductor to align grains along  $ab$  planes reduces the detrimental effects of grain boundaries. It has been determined that the presence of low-angle grain boundaries (i.e., adjacent grains interfacing at a low angle ( $<10^\circ$ )), is not deleterious to  $J_c$  (18). However, the elimination of high-angle grain boundaries in HTS materials has been a difficult challenge to overcome. Recent developments in YBCO-coated conductors show promise in reducing the presence of high-angle grain bound-

aries, discussed later. To date, YBCO is the most successful candidate for enhancements in flux pinning (19). Enhancement of flux pinning and suppression of flux creep have been achieved by the introduction of controlled amounts of defects, such as coexisting normal secondary phases by alloying, or by neutron or proton irradiation to define normal regions along the particle trajectory that give rise to higher  $J_c$ s and higher  $H_c$ s (20).

Tremendous progress has been made in the synthesis of bulk single-phase HTS conductors, silver-sheathed wires, tapes, and thin-film-coated conductors in the past decade as shown in Table I. Bulk conductors with large volume fraction of superconductivity are being routinely synthesized in YBCO-, Tl-, and Bi-based HTS materials. Improved powder synthesis, understanding of reaction chemistry, and phase diagrams have led to increased phase purity and improved flux pinning (resulting in higher current density at self field). Solid-state reaction and melt texturing are the two widely used processing methods for bulk HTS.

#### PROCESSING OF BULK SUPERCONDUCTORS

There are three important process considerations for the choice of bulk superconductors for applications such as magnets, current leads, and superconductor magnetic energy storage systems: (1) growth of large-area, single-phase bulk conductors, (2) ease of introducing strong flux pinning centers in the bulk, to enhance the critical currents, and (3) long-term stability (i.e., ability to withstand thermal cycling to modest temperatures) and aging under ambient conditions. Among the various bulk superconductors, YBCO appears to be the most plausible for electrical applications since the details of the crystal structure have been reliably established, and it is easy to synthesize single-phase specimens. Moreover, in YBCO, introduction of flux pinning centers to enhance  $J_c$  has been demonstrated, although this is not the case in Bi, Tl, and Hg superconductors. Further, melt-processing growth of YBCO superconductors provides grain-alignment resulting in  $J_c$  as high as  $10^5$  A/cm<sup>2</sup> at self field and 77 K (21). In the case of YBCO, the melt-powder-melt growth (MPMG) processing has significantly improved performance of bulk materials. The thallates are also attractive as  $T_c$  as high as 125 K,  $J_c$  in bulk greater than  $10^5$  A/cm<sup>2</sup> at 77 K and zero field (22) have been achieved. An advantage with the thallates is that multiple phases, especially Tl2223 and Tl2212 phases, can coexist, without degrading the superconducting properties in bulk conductors, thin films, and wires. The magnetic field dependence of  $J_c$  is comparable to the best YBCO superconductors. A disadvantage of thallates is the toxicity of Tl, which needs very careful handling and processing. The volatility of Tl gives rise to its losses during high-temperature synthesis. Most of the thallates contain structural defects, which affect the superconducting properties negatively. The thermodynamics of Tl-based compounds is quite different from other HTS materials due to the volatility of Tl (23).

At the present time, the synthesis of mercury-based HTS has unique problems mainly due to the volatility of mercury compounds. Since Hg(0) decomposes around 500 °C, all processing must be done in a sealed environment in the production of bulk superconductors. Also, the Bi compounds are not attractive for bulk conductors primarily due to the polycrystalline nature of the grains and weak flux pinning above 30 K (19).

Melt-texturing Bi compounds is not as simple as the YBCO. Bismuth compounds are generally thought to be attractive for wire applications owing to their micaceous morphology. In the following section, we provide processing details on YBCO, Tl2212, and Tl1223 bulk superconductors.

### Processing YBCO 123 Bulk Superconductors

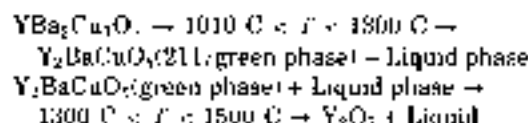
**Conventional Solid-State Reaction Method.** The conventional means of producing bulk HTS materials is the solid-state reaction method, or calcination. Carbonates, or oxides of Ba, Ca, Cu, and Y as starting materials, are mixed and ground to the desired size and then heat-treated in an alumina crucible at temperatures as high as 950 °C for 24 h. After cooling to room temperature, the reacted sample is finely ground and mixed again and heat-treated for an additional 24 h; this step is repeated to promote homogeneity. The samples are then pelletized in desired shapes using a hydraulic press at a typical pressure of 1500 psi. The pellets are placed in a crucible made of gold or platinum or alumina, introduced into a furnace preheated to 900 °C, and sintered typically in air for 24 h. Oxygen is introduced into the furnace at around 500 °C, during slow cooling of the samples to room temperature. This gives rise to the growth of the orthorhombic superconducting phase. Although the technique is fast, there are undesirable phases present in samples prepared this way, typically BaCuO<sub>2</sub>, CuO, and possibly some contamination from the heat-treatment process. In spite of these limitations, it is possible to get  $T_c$  of 90 K and current densities of 10<sup>4</sup> A/cm<sup>2</sup> at zero field at 77 K. The polycrystalline nature of grains in such samples results in poor field dependence of  $J_c$ , with  $J_c$  decreasing by three orders of magnitude in a magnetic field of 0.1 T.

### Texturing of YBaCuO 123 Superconductors

Texturing of YBCO can be achieved using mechanical, magnetic, and melt-growth techniques. The mechanical and magnetic means produce only a marginal improvement in superconducting properties, mainly owing to the presence of large-angle grain boundaries, whereas melt growth techniques reduce the large-angle grain boundaries and improve the critical currents significantly. The melt-texturing techniques are the most suitable for enhancement of  $J_c$ . Among the melt-texturing methods, melt texture growth (MTG) (21), quench and melt growth (QMG) (24), and melt powder melt growth (23) are techniques developed for nondirectional solidification. In these techniques, large YBCO crystals grow stacked along the *c*-axis, with improved connectivity between grains. Directional solidification is also possible by providing the right thermal gradient during the melt processing, either by the movement of the furnace or the sample. Directional solidification methods align the grains along the *ab*-planes, the direction of current flow, thus improving the current-carrying capability. It is also possible to align the grains along a preferred orientation using seeded melt growth (25) and liquid phase removal methods (26). In the seeded melt growth technique, a directionally oriented single crystal Sm123 or Nd123 is used as a seed material to grow YBCO, oriented along the same preferred direction.

The melt texture growth of YBCO 123 involves melting the YBCO above its peritectic temperature, approximately 1010 °C, in air. When heated above this temperature, the 123 compound undergoes maximum melting to form a solid and a liquid

phase according to the following reactions (27):



On subsequent slow cooling, large oriented domains of 123 phase with 211 (green phase) inclusions are formed. In such samples, grains grow preferentially along *ab*-planes, are stacked along the *c*-direction, and are coupled by low-angle grain boundaries. This microstructure is in contrast to small randomly oriented polycrystalline grains obtained by solid-state reaction. Quenching studies performed on MTC samples show the presence of faceted growth interfaces where the 211 volume fraction and particle size decrease abruptly upon being included into the growing 123 grains. The 211 phase appears to provide Y into the growing phase of 123. The MPMG process is the most commonly used melt growth process designed for smaller sizes of the 211 precipitates, and homogeneous distribution of the precipitates. This technique involves reaction of Y<sub>2</sub>O<sub>3</sub> with a liquid phase to form the 211 phase. The nonuniform distribution is handled by taking the powder and crushing and remixing the melt quenched samples. This powder is pressed into a pellet and subjected to the melt growth. It was found that the samples should not be held above the peritectic temperature for too long because the 211 phase will grow into coarse grains. After a short time above the peritectic temperature, it is cooled to just below the peritectic temperature and then slow-cooled in flowing oxygen to room temperature. The temperature profiles for MTC and MPMG processing techniques are shown in Fig. 7. The critical temperatures are identified for both processes. The evolution of the microstructure during the MPMG process is shown in Fig. 8 (27). The figure illustrates the formation of 211 inclusions in to the liquid phase.

The seeded melt growth is the most promising for large area domains of YBCO (27). In this technique, a single crystal seed of Sm123 or Nd123 is embedded in a presintered Y123 bar, and the whole assembly is introduced into a modified Bridgman furnace for 15 min to 30 min and then withdrawn at a very slow rate of 1 mm/h. As the growth front of Y123 passes by the seed, seeded growth occurs at the interface. Because of the favorable alignment along *ab*-planes, large current densities, as high as 46,000 A/cm<sup>2</sup> at 77 K and zero field have been demonstrated. In the melt growth techniques, the hold time and cooling rate have a strong influence on the microstructure of the samples, determining the grain size and nature of the grain boundaries. The hold time determines the amount of liquid phase formed for the peritectic recombination. The cooling rate determines the time available for the recombination of the liquid (25).

### Processing of Tl(BCCO) 2212 and 1223 Superconductors

In general, processing of thallates with high superconducting volume fraction ( $V_{sc}$ ) has been difficult, primarily because of the volatility of Tl. Processing of Tl based HTS can be done either using an open system, or a closed system. In the open system, a sample is heat treated in an oxygen atmosphere or air at temperatures between 850 °C and 900 °C, for 10 min to 15 min. In a closed system, samples are sealed in a gold tube and heat-treated for longer times of 24 h to 48 h (28). The open system is susceptible to loss of Tl and, hence, requires additional



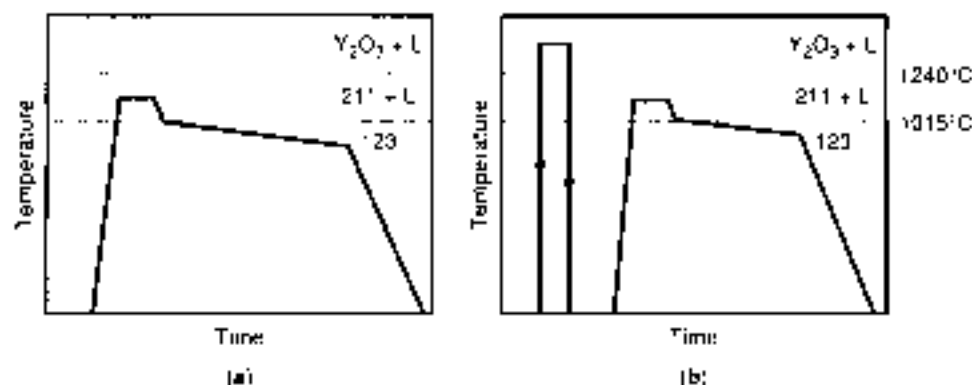


Figure 7. Temperature profiles for (a) the multistage growth and (b) the melt process melt growth. Courtesy of Ref. 27.

compensation for Tl. The closed system has the disadvantage of sealing the samples in a container or a foil, resulting in reaction between the Tl-based compound and the container. In our work, we were able to synthesize samples with  $T_c$ s greater than 100 K for the Tl2212 phase and 115 K for the Tl1223 phase, with superconducting volume fractions greater than 50% in each case, by tuning the process parameters (29). The processing parameters that were optimized included the starting compositions of the compounds, sintering temperatures, and sintering times. The starting compositions are critical mainly because of the loss of Tl during heat treatments. In the 1223 superconductor the Hitechi composition of  $Tl_{0.5}Pb_{0.5}Sr_{1.0}Ba_{0.4}Ca_2Cu_2O_7$  has been a popular starting composition for high quality 1223 superconductor (30). In our study, we examined the role of different Ba concentrations and excess Tl and Pb doping concentrations to obtain a single phase 1223 superconductor in an open system (29).

Tl2212 samples were synthesized starting from  $BaCuO_2$ ,  $CaO$  (99.999%), and  $Tl_2O_3$  (99.999%) compounds.  $BaCuO_2$  was prepared by mixing  $BaO_3$  (99%) and  $CuO$  (99.99%) in equal molar portions and heat treating at 900°C for 24 h in air. The latter

step was repeated after regrounding the pellet to promote homogeneity.  $BaCuO_2$ ,  $CaO$ , and  $Tl_2O_3$  were mixed in a molar ratio of 2:1:1.07, pelletized, and reacted in an enclosed Pt crucible. Our results indicate that samples sintered at 550°C for 10 min followed by a second heat treatment at 860°C for 10 min displayed optimum superconducting properties, which included a sharp  $T_c$  of 97 K, with a superconducting volume fraction as high as 54%. The porosity of the samples appears to be the limiting factor for the  $V_{sc}$ .

The first step performed in the synthesis of (Tl,Pb)(Sr, Ba)  $CaCuO_2$  1223 superconductor, was the preparation of a  $SrO$ - $CuO$  precursor, starting from 99.999% pure  $CaO$  and  $CuO$  and 99% pure  $SrO$ . Appropriate molar portions of  $SrO$ ,  $CaO$ , and  $CuO$  were mixed, heated to 900°C for 24 h, cooled, and ground into a fine powder, then the same process was repeated to obtain a homogeneous precursor compound. For the Ba-doped samples, appropriate amounts of 99.99% pure  $BaCO_3$  were alloyed with  $SrO$  in the precursor. The second step was to add appropriate amounts of 99.999% pure  $PbO$  and  $Tl_2O_3$  to the precursor. Pellets of approximately 500 mg were made and heat-treated in an enclosed Pt crucible, at a temperature of 870°C for 10 min, followed by a slow cooling.

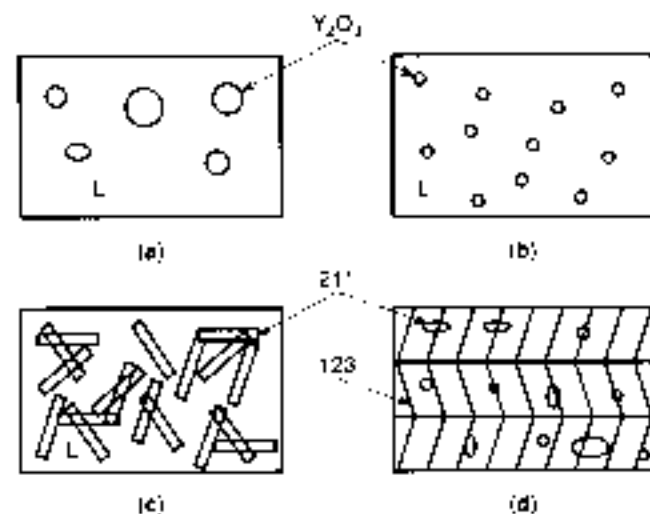


Figure 8. Schematic representation of microstructure evolution in the MPMO growth of YBCO: (a) after quenching—yttria and solidified liquid, (b) during heating to the 211 + L region—yttria particles redissolving, (c) in the 211 + L region, and (d) final microstructure showing the 123 and 211 phases. Courtesy of Ref. 27.

## CHARACTERIZATION OF BULK SUPERCONDUCTORS

### Structural Characterization

The primary characterization of any form of a superconductor consists in establishing the structure by X-ray diffraction (XRD). X-rays from a source are collimated to strike the sample at an incident angle and Bragg scattered from the sample. For X rays, atomic planes of the material serve as a diffraction grating, causing the monochromatic X-ray beam to diffract and form a pattern. From an XRD pattern, one can establish the various phases populated and, in principle, the crystal structure including lattice parameters of each phase. In practice, since the  $c$ -axis lattice parameter depends on the oxygen content, XRD patterns can be used effectively to correlate structural results with superconducting parameters. If there is more than one phase present in a sample, the XRD patterns can be used to obtain quantitatively the concentration of various phases present in a sample. Since bulk materials synthesized by solid-state reaction are polycrystalline, the XRD pattern from such a sample has peaks corresponding to scattering from different

crystallographic planes such as (100), (110), and (001). Powder XRD is an effective means to establish the major phases present in a sample.

### Magnetic Properties

Vibrating sample magnetometer (VSM) is a widely used technique for the study of all types of magnetic materials as it permits a measurement of the dc magnetization of a sample. In HTS materials, VSM permits establishing the  $T$  dependence of the diamagnetic susceptibility in both a zero-field-cooled (ZFC) and a field-cooled (FC) configuration. In the VSM technique, a sample is placed in a uniform magnetic field and vibrated at a low frequency so as to induce a voltage signal in a pickup coil proportional to the magnetic moment of the sample. The magnetic moment is proportional to the product of the sample susceptibility and the applied field (30). Measurements are typically performed at several temperatures to obtain the magnetic and superconducting properties. A plot of the magnetization versus applied magnetic field can yield the magnetization  $J_c$  using the Bean's critical state model. For a cylindrical sample, the magnetization  $J_c$  is approximately  $30 \Delta M/d$  where  $\Delta M$  represents the width of the hysteresis loop at a particular applied field  $H$ , and  $d$  the effective thickness of the sample normal to the applied field (31). This model typically results in an overestimate of the actual  $J_c$ . A more direct method to determine the transport  $J_c$  is by using the four-point probe method, in which a  $1 \mu\text{V}/\text{cm}$  electric field criterion is used to establish the transport  $J_c$ . Other magnetization measurement techniques include ac and SQUID magnetometers. In the ac magnetometer, a sinusoidally oscillating field is used. The primary coil is surrounded by a set of oppositely wound detector coils. Introducing a magnetic sample into the center of one or the other secondary detector coil leads to a voltage imbalance in the secondary circuit (32). The resulting flux change leads to a measurable induced emf, which is proportional to the sample susceptibility. Magnetic moment sensitivities as low as  $10^{-6}$  emu are possible using a commercial ac susceptometer. Enhanced sensitivity can be obtained using a SQUID magnetometer, as SQUIDs are used in the detector circuitry. There are a large number of tools available for researchers to correlate the physical properties with the processing variables. Table 2 lists characterization tools used by HTS researchers, the outcomes from each tool, and corresponding observable.

### Structural, Electrical, and Magnetic Properties of Bulk Superconductors

#### YBCO (123) Superconductors

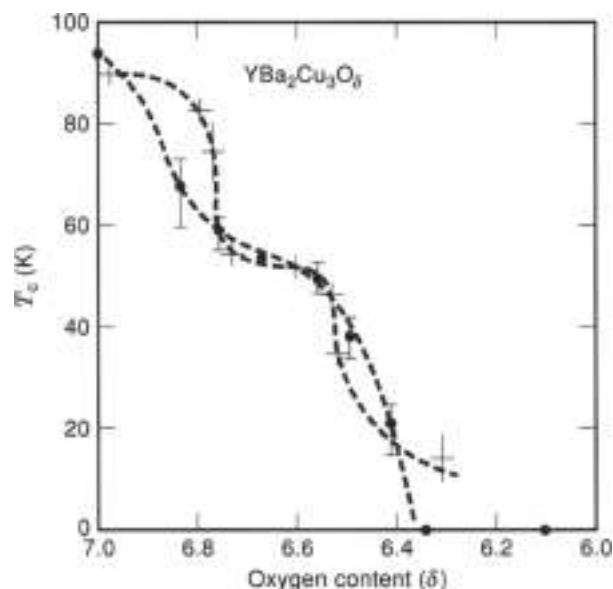
The  $T_c$  of YBCO is near 90 K, provided the oxygen content in YBCO is optimal. YBCO, in particular, has remarkable  $T_c$  sensitivity to oxygen stoichiometry as displayed in Fig. 9 (33). Specifically, one finds that at  $\delta = 7.0$  in the  $\text{Y}_2\text{Ba}_2\text{Cu}_3\text{O}_x$  superconductor,  $T_c = 90$  K. For  $\delta < 6.2$ ,  $T_c = 0$  K. The plot of Fig. 9 also reveals that in the range  $6.6 < \delta < 6.8$ ,  $T_c = 60$  K, and displays a plateau, a feature that is identified with the existence of a superstructure in oxygen bonding. Figure 9 also reveals that it is not only the value of  $\delta$ , but also the distribution of oxygen in the structure that controls the value of  $T_c$ . For YBCO, it is well established from neutron Bragg diffraction measurements that oxygen in the  $\text{CuO}_2$  planes remains intact, whereas oxygen in the chains ( $\text{Cu}_1\text{O}_2$ ) diffuses at modest temperatures ( $T < 300$  C) and effuses out of the structure at  $T > 500$  C. An orthorhombic

**Table 2. Basic Materials Characterization Tools Used in Research in High- $T_c$  Superconductors**

| Method   | Observable   | Outcomes  |
|--|--|---|
| Four Point Probe, ac/DC/Field and applied Magnetic Field | $T_c$ , Transport $J_c$  | Resistivity above and below $T_c$ , presence of flux creep                          |
| DC magnetometer  | Diamagnetic susceptibility, $T_c$ , magnetization $J_c$        | Superconducting volume fraction   |
| AC magnetometer  | $T_c$ , real and imag. susceptibility ( $\chi'$ and $\chi''$ ) | Superconducting volume fraction, weak link dissipation                              |
| SQUID magnetometer                                       | $T_c$ , diamagnetic susceptibility, Josephson tunneling        | High-resolution diamagnetic susceptibility, symmetry state of pairing wave function |
| Microwave conductivity                                   | Surface resistance   | Penetration depth at 0 K, conductor losses  |
| X-ray diffraction  | Lattice parameters, symmetry                                   | Crystal structure, presence of secondary phases, grain alignment                    |
| EXAFS  | Local co-ordination of atoms                                   | Local atomic scale structure, bonds   |
| Mossbauer spectroscopy                                   | $\delta$ , local fields, $A$ factor                            | Chemically specific local structure, local vibrational density of states            |
| Inelastic neutron scattering                             | $\rho(\omega)$ : vibrational density                           | Total vibrational density of states   |
| Raman scattering   | Phonon and phonon line shape                                   | Isotope effect, electron-phonon coupling  |
| Electron energy loss spectroscopy (EELS)                 | Local chemistry  | Oxygen hole distribution  |
| Thermogravimetric analysis (TGA)                         | Sample weight loss   | Oxygen stability  |
| Calorimetry  | Spectroscopic heat capacity                                    | Nature of pairing state   |
| Auger electron spectroscopy                              | Surface chemistry, electron density of states                  | Electronic structure near the Fermi-level and origin of pairing                     |

to tetragonal phase transformation occurs near an oxygen content of  $\delta = 6.5$ . Oxygen can be reversibly desorbed and absorbed from YBCO either by heating in a vacuum or an oxygen ambient, respectively. Plasma oxidation has been demonstrated as an attractive means to oxygenate HTS at rather modest temperatures for microelectronic applications. At a local level, it appears that the  $T_c$  variation can be traced to the length of the  $\text{Cu}(1)-\text{O}(4)$  bond that displays a threshold behavior as shown in Fig. 10 (33). Superconductivity in YBCO commences when the  $\text{Cu}(1)-\text{O}(4)$  bond length acquires a value of 1.82 Å or larger. Bond lengths less than 1.82 Å, apparently couple the pyramidal oxygen O(4) to the  $\text{Cu}(1)$  chain cations, inhibiting charge transfer from the chains to the  $\text{CuO}_2$  planes. Oxygenation of YBCO to a value of  $\delta \approx 6.5$  gives a  $\text{Cu}(1)-\text{O}(4)$  bond length close to the threshold value.

The melt-textured samples have higher  $J_c$  compared with conventional solid-state-sintered samples. Figure 11 shows the enhanced  $J_c$  typical of melt-processed samples (14, 15). For comparison,  $J_c$  of an epitaxial YBCO thin film is shown. The figure shows improvement in  $J_c$  for melt-textured samples compared with the conventionally sintered samples. Melt-processed

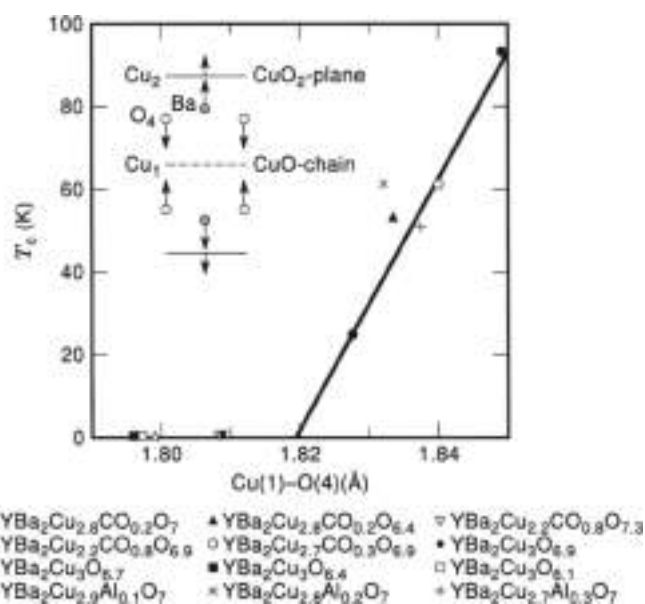


**Figure 9.** Critical temperature versus oxygen content for YBCO samples. The dots and crosses were measured on samples with oxygen removed by high-temperature quenching and low-temperature processing, respectively. Courtesy of Ref. 35.

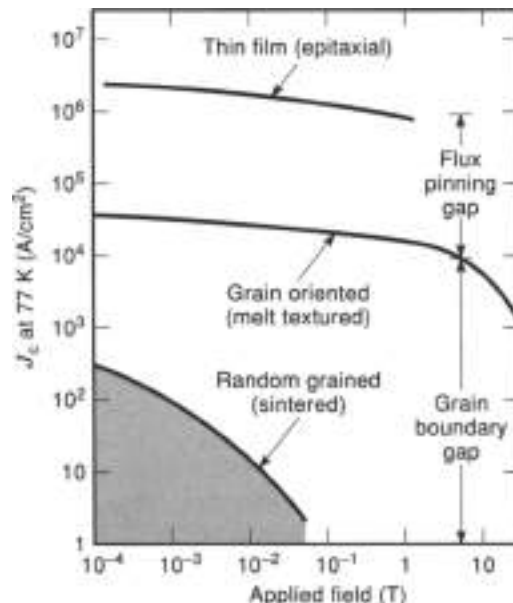
samples have at least an order of magnitude higher in field  $J_c$  and do not drop as rapidly at higher fields as the sintered samples.

**2212 superconductors.**

Synthesis of the Tl2212 superconductors was performed by reacting appropriate amounts of the  $Ba_2Ca_2Cu_3O_7$  precursor with

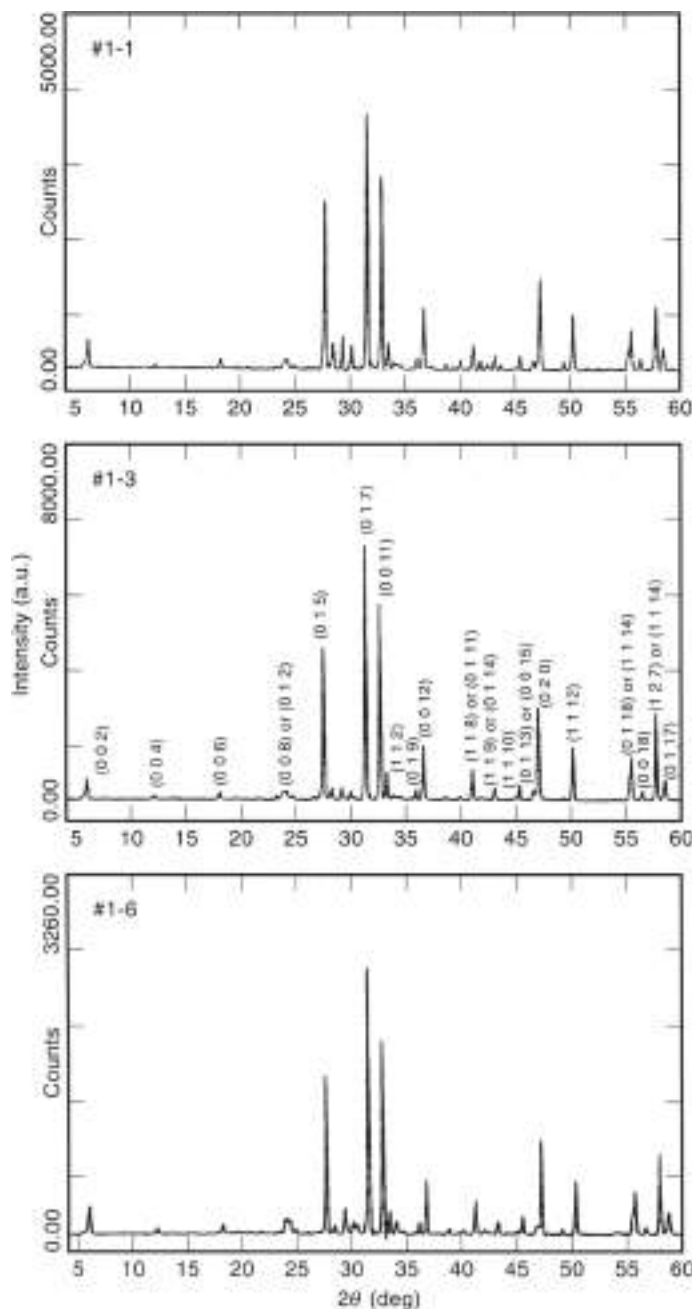


**Figure 10.** Critical temperature versus Cu(1)-O(4) (copper in the chain bridging oxygen) bond length in the 123 phase. Also shown are data for dopants introduced into the YBCO phase with similar structural changes. As shown in the inset, the  $CuO_2$ - $CuO_2$  plane distance increases, the O(4) moves toward the CuO chains, and the Ba moves toward the  $CuO_2$  planes. Courtesy of Ref. 33 and references therein.



**Figure 11.** Comparison of field dependence of critical current density at 77 K, for conventional solid-state-sintered YBCO, melt-textured YBCO, and the epitaxial YBCO thin films. The figure shows improvement in  $J_c$  as a function of processing. Courtesy of Refs. 14 and 15.

$Tl_2O_2$ . The results of three samples each processed differently are presented here. Sample 1.1 was processed at 850 C for 10 min. Sample 1.3 was processed at 850 C for 10 min followed by an additional heat treatment at 860 C for 10 min. Sample 1.6 was prepared by heat treatment at 850 C for 10 min followed by heat treatment at 860 C for 45 min. From XRD measurements, all samples showed the characteristic Bragg reflection peaks of the  $Tl_2Ba_2CaCu_3O_x$  phase. No other impurity phases were detected in the XRD scans. Figure 12 shows the XRD reflections for the three samples in which characteristic peaks of the Tl2212 phase are clearly visible. Figure 13 shows the magnetization hysteresis loops of the three samples recorded at 15 K. The magnetization scale for all samples was normalized to their mass. As seen in Fig. 13, the diamagnetism displayed by sample 1.3 is the largest among the samples investigated. This result has been reproduced in several samples. Figure 14 displays the temperature dependence of magnetic susceptibility. In this figure, the scans are labeled according to the second heat-treatment temperature and time. One finds that the second-step heat treatment at 860 C for 10 min not only increases the saturation diamagnetization at  $T = 15$  K but also significantly narrows the width of the superconducting transition, without affecting the  $T_c$  value. Our conclusion is that the second-step heat treatment produces more strain-free material, or chemical ordering of the phase. The underlying structural growth does not alter the bulk  $T_c$ , but do significantly enhances the diamagnetic response (Fig. 14), owing due to growth of a more strain-free material. The superconducting volume fractions for samples 1.1, 1.3, and 1.6 were 37%, 54%, and 27%, respectively. Clearly, the process conditions used for sample 1.3 appear to optimize growth of the superconducting phase. Further increase in the second-step heat treatment time, resulted in an increase of  $T_c$  to above 110 K. It has been shown by Sugase et al. [34], that prolonged heating of Tl2212

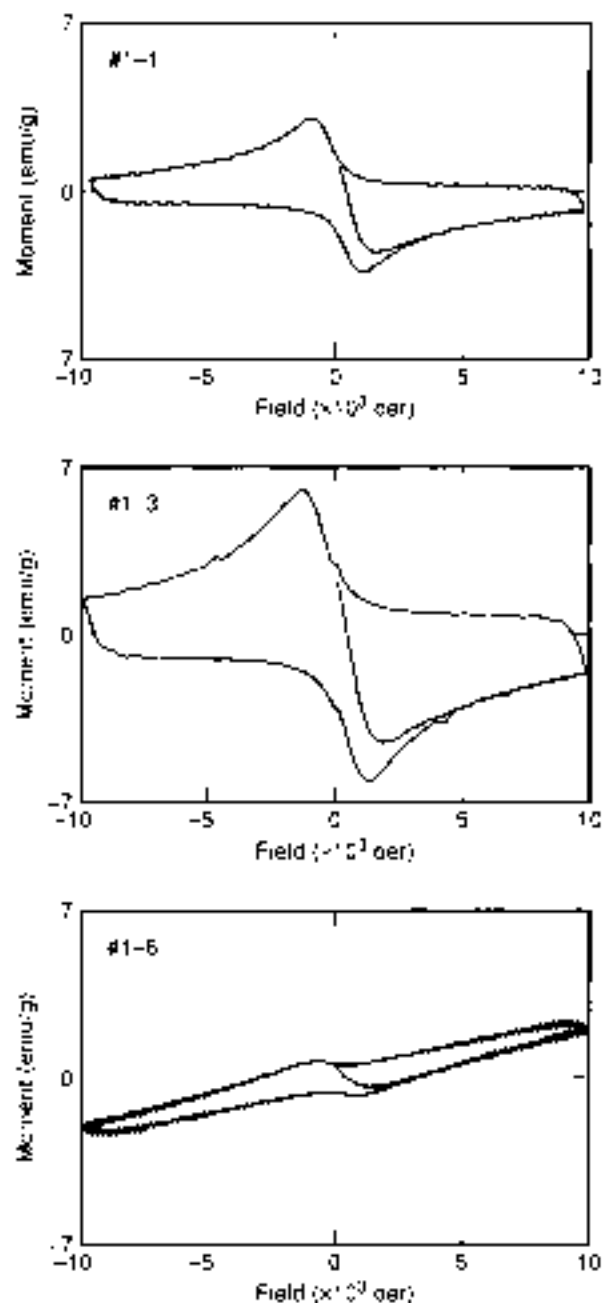


**Figure 12.** Powder X-ray diffraction spectra obtained three Tl2212-sintered samples processed at different conditions as outlined in the text.

sample in air can grow bulk 2223 samples. Depletion of  $Tl_2O_3$  from Tl2212 samples essentially leads to the transformation of the two-layered cuprate Tl2212 into the three-layered cuprate Tl2223 structure. Since the XRD scans provide no evidence of presence of the 2223 phase, it must be present as stacking faults, or planar defects, as documented by Raveau et al. (35) using high-resolution electron microscopy. Also, in Fig. 14, note that the sample with second-step 20 min heat treatment in a sealed gold tube displays much smaller diamagnetism than the samples heat-treated in a Pt crucible. This is evidence of a reaction between the thallates and the gold tube at 860°C in a closed system and suggests that gold should be avoided as a crucible for processing in a closed system of thallate superconductors.

#### Tl1223 Superconductors.

For the synthesis of 1223 compound, the starting composition of  $Tl_{2-x}Pb_xSr_2Ca_2Cu_3O_{10}$  (1223) stoichiometry upon thermal processing yielded samples with 1223 and 1212 phases that had a  $T_c$  of 107.5 K and a volume fraction of 66%. Replacement of some Sr by Ba in the starting composition  $Tl_{1.5}Pb_{1.5}Sr_{2-x}Ba_xCa_2Cu_3O_{10}$ , on the other hand, yielded samples that had the 1223 phase. Samples with a Ba content  $x < 0.4$  clearly were found to have both 1212 and 1223 phases, whereas samples with  $x > 0.4$  showed principally the 1223 superconductor. At  $x = 0.75$ , samples containing only the 1223 phase were obtained upon heat treatment at 870°C for 11 min, using starting materials in which 50% additional  $Tl_2O_3$  and



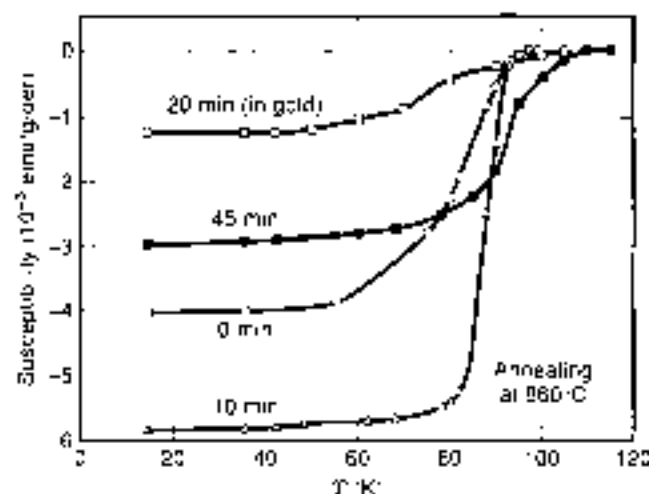
**Figure 13.** Diamagnetic response of the samples 1.1, 1.3, and 1.6 obtained using VSM at 16 K. Sample 1.3 shows the largest diamagnetic susceptibility among the three samples, indicating that the best treatment procedure used for sample 1.3 is close to the optimum.

PbO were compensated. Low-field ac susceptibility measurement of the sample shown in Fig. 15 reveals an onset  $T_c$  of 117 K with a superconducting volume fraction of about 80%. Note that the in-phase susceptibility does not indicate the presence of minority phases. Samples with a Ba content of  $x = 0.75$  showed the highest  $T_c$  of 117 K. Our results indicate that a minimum Ba content of  $x = 0.4$  is necessary to produce a majority 1223 phase, which appears to be consistent with the Hitachi composition. Our study also demonstrates that the highest  $T_c$  with the largest superconducting volume fraction are obtained at a Ba content of  $x = 0.75$ , in Tl and Pb over compensated samples.

#### Metal-Doping Effects in YBCO Bulk Superconductors

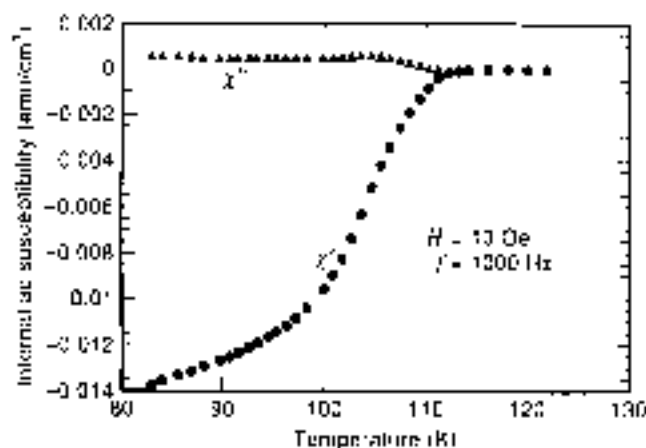
Doping transition metals of group III (Fe, Co, Ni) in YBCO exemplifies the trials and tribulations of working with high  $T_c$

superconductors. As material quality has improved, so have the challenges to understand the nature of the pairing state. One of the areas that has been intensively studied is the doping of Fe in YBCO. More than 1000 papers have been published in this area, with the main conclusion that the broken line curve in Fig. 16 displays the  $T_c(x)$  variation separating the superconducting phase ( $x < x_c$ ) from the antiferromagnetic insulators ( $x > x_c$ ) near  $x_c = 0.13$  (36). This is the typical result for samples prepared by standard solid-state reaction at ambient pressure of oxygen (henceforth referred to as APO). Recently, it has been demonstrated that the underlying phase diagram is actually an artifact of sample processing. When these samples undergo an additional sinter at 915°C at elevated pressures of oxygen (220 atm of pure oxygen),  $T_c(x)$  increases qualitatively, and all the samples studied up to  $x = 1/3$  are rendered superconducting with a  $T_c$  greater than 60 K. In particular, samples that were considered to be antiferromagnetic insulators ( $x >$



**Figure 14.** The temperature dependence of magnetic susceptibility obtained from VSM measurements. Sample 14 shows the sharpest transition among the samples studied.

(13) have been rendered superconducting, thus altering the presently accepted superconducting-antiferromagnetic phase diagram. Powder X-ray diffraction measurements have shown that the average structure of the samples synthesized at high pressures of oxygen (henceforth referred to as HPO) are also tetragonal as the usual samples synthesized at APO, except for a small excess of oxygen. The Fe local environments as revealed directly in Mossbauer spectroscopy measurements in the two types of samples (APO and HPO) at the same value of  $x$  are however, qualitatively different, suggesting that aspects of dopant-centered local structures control in a remarkable way the superconducting behavior in these cuprates. There are plausible reasons to suggest that the majority of Fe resides in the  $\text{CuO}_2$  chains, but that in APO samples at least  $\text{Fe}^{2+}$  does not replace  $\text{Cu}^{2+}$  but instead acquires a tetrahedral coordination as exemplified by  $\text{Ga}^{3+}$  cations in  $\text{Y}_2\text{Sr}_2\text{Ga}_2\text{Cu}_2\text{O}_7$  superconductors (17). In HPO samples,  $\text{Fe}^{2+}$  is oxidized to  $\text{Fe}^{3+}$  with localization of an additional oxygen atom in its near-neighbor coordination, thus qualitatively changing the local site symmetry. The Co doping effects in YBCO bear a similarity to those of Fe doping, although the effects are less dramatic. The HPO sintering of APO samples increases the  $T_c$  quantitatively in general



**Figure 16.** The ac susceptibility measurement of a 122d sample with  $x = 0.75$  obtained using an ac susceptometer.

at doping concentrations  $x < x_c = 0.20$ . At  $x > x_c$ , antiferromagnetism sets in, and even HPO synthesis is unable to transform such samples to superconductors. The Ni doping effects in YBCO are apparently unaffected by HPO synthesis as revealed by the results shown in Fig. 16. The  $T_c$  vs  $x$  variation of Ni-doped YBCO samples sintered at APO is similar to the variation in HPO samples. In contrast to Fe and Co doping, Ni doping in YBCO does not alter the lattice symmetry, which remains orthorhombic, probably because the dopant enters the structure substitutionally as  $\text{Ni}^{2+}$  replacing  $\text{Cu}^{2+}$ . The Fe and Co dopants enter the structure in higher charge states  $\text{Fe}^{3+}$  and  $\text{Co}^{3+}$  and  $\text{Co}^{4+}$  largely in planes and bring in additional oxygen to alter not only the lattice symmetry, which changes from orthorhombic to tetragonal, but also to create new local structures. The nature of these local structures continues to be a subject of current interest in large part because there appears to be a close correlation between the existence of specific structures and appearance of bulk superconductivity. These new results also demonstrate that Fe-doped YBCO samples synthesized at HPO are chemically more stable than those synthesized at APO. The higher chemical stability of HPO-synthesized YBCO samples may be attractive for various applications of bulk, thin films, and wires.

## PROCESSING OF HTS THIN FILMS

### *In-Situ* Versus *Ex-Situ* Processing

Thin films of HTS materials (see THIN FILMS) have been deposited on a wide variety of substrates using two different procedures: *in-situ* processing and *ex-situ* processing. Both processes require high-vacuum deposition systems. *In-situ* processing is a technique in which the samples are deposited and processed inside a vacuum chamber and may not require any postprocessing outside the chamber (i.e., the samples they will be superconductors when they are removed from the vacuum chamber). The *ex-situ* processing method, on the other hand, requires postprocessing of samples outside the chamber. A deposition process will yield an amorphous material if the substrates are not heated; therefore, additional heat-treatment becomes necessary to obtain superconducting samples. The *in-situ* procedure is preferred because it leads to thin films with uniformity, with good control over the growth process and stoichiometry. An *in-situ* deposition process usually uses a complex deposition system configured with a substrate heater assembly and process-monitoring capabilities. *Ex-situ* processing allows growth of multilayers with precise thickness control, essential for applications such as coated thin film conductors on metallic substrates with buffer layers. A drawback of the method is that it is not suitable for deposition on large area substrates. An *ex-situ* procedure is a simpler process because it is easier to optimize, and one can obtain reasonably good quality HTS thin films with postprocessing outside the vacuum chamber. The system requirements are less expensive compared to *in-situ* processing and are possibly applicable for thick or coated films of superconductors.

### Physical and Chemical Deposition Methods

Deposition of HTS thin films can be done using either a physical or a chemical deposition method. The physical deposition methods include techniques such as sputtering, electron beam

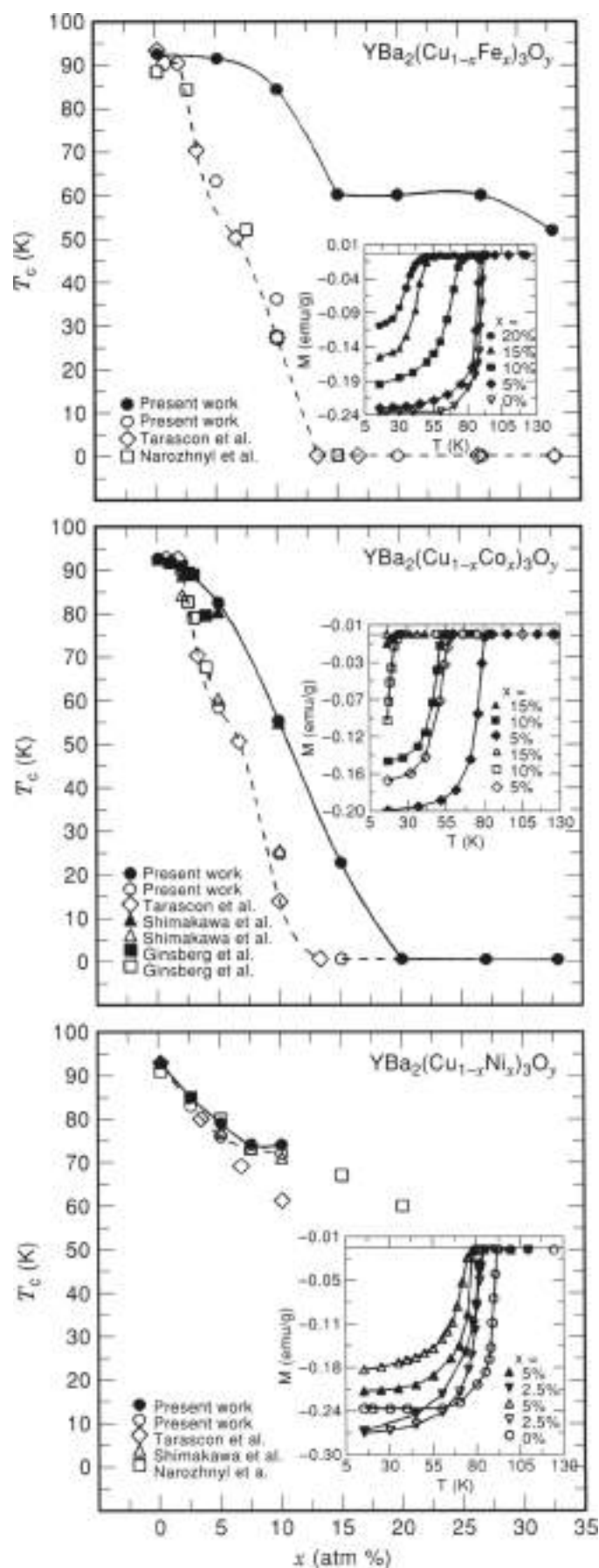


Figure 18.  $T_c$  plotted as a function of metal doping concentration  $x$  for  $M = \text{Fe}$  (top panel),  $M = \text{Co}$  (middle panel) and  $M = \text{Ni}$  (bottom panel) in YBCO samples synthesized at high  $P$  (filled symbols) and at ambient  $P$  (open symbols). Lines drawn through data points are guides to the eye. The  $T_c$  values determined either magnetically or resistively are representative of a vast literature. Magnetization measured in a 20-G field in the zero field cooling mode, are displayed in insets.

evaporation, and pulsed laser ablation or deposition. Chemical deposition methods include techniques such as metal-organic chemical vapor deposition (MOCVD) and metal-organic decomposition (MOD) (see THIN FILMS). Primarily, physical deposition techniques include some form of high-energy ion or electron bombardment of the target to release the material for deposition onto substrates. A chemical deposition method involves chemical vapors (transported using a gas such as hydrogen or nitrogen) reacting at the surface of heated substrates to form a thin film of desired composition. Both physical and chemical methods can be tailored for *in-situ* or *ex-situ* processing. Among the deposition methods, sputtering and pulsed laser deposition (PLD) are the most popular methods used, especially for the *in-situ* processes (see THIN FILMS).

#### Pulsed Laser Deposition

Pulsed laser deposition has been successfully used to deposit HTS thin films for *in-situ* processing; PLD is designed to replicate target stoichiometry in the deposited thin films (38). A typical PLD system consists of an ultra-high vacuum chamber containing fused quartz windows for *in-situ* spectroscopic investigations. Pulsed laser deposition is a vapor phase deposition process in which a focused pulsed laser beam of energy density (also called fluence) exceeding  $1 \text{ J/cm}^2$  strikes a sintered composite target material at an angle of  $45^\circ$ , vaporizes the target, and grows a thin film on a heated substrate placed nearby. A  $\text{CO}_2$  laser, a Nd:YAG laser, a XeCl, or a KrF excimer each have been successfully used for PLD growth of HTS. Multiple targets can be used with each target ablated for a required amount of time. In multitarget systems, a target is rotated and synchronized with an optical encoder for pulsing the laser. This helps reduce the focused laser beam's heating effects on the target. A water-cooled target holder is normally used for further heat removal. *In-situ* deposition of YBCO thin films is typically performed at an oxygen partial pressure in the 70 mTorr to 200 mTorr range, on substrates heated in the range of 600 to 800°C depending on the choice of substrates. The presence of oxygen is required for compositional control of the oxide superconductors. After deposition, an oxygen anneal of the deposited film is performed for several hours at a reduced temperature (400 to 500°C), with slow cooling, in a higher oxygen partial pressure of 500 mTorr to 700 mTorr to obtain superconductivity (38, 49).

Typical growth rates using a PLD system are in the range of a few angstroms per second. The growth parameters controlling the quality of thin films include the choice of the substrate, substrate temperature, background pressure of the vacuum system, laser energy density, frequency of the laser pulse sequencer, and oxygen partial pressure in the system. The HTS thin films prepared by PLD have excellent surface morphology, although one of the problems encountered has been the presence of particulates on the surface of such thin films. Understanding the chemistry of the laser interaction with sintered radicals of the target material is essential to control the growth quality of thin films over large areas and continues to be a subject of current research.

#### Sputtering of Thin Films

Sputtering is perhaps one of the more popular deposition methods used in the semiconductor industry for metals and dielectric materials. Broadly, there are two methods used: dc sputtering and RF sputtering. Both methods use ion plasma, typically

argon, because of its inertness. High-energy ions bombard a target material to release target atoms, which are deposited on a substrate. Substrates can be heated to a high temperature for *in-situ* processing of materials. A modified method in sputtering, called magnetron sputtering, uses a magnetic field at the target to confine the plasma, which leads to higher deposition rates at low working gas pressures. Sputtering systems can employ either a single target or multiple targets. A multitarget sputtering system gives precise control in multilayer depositions, because one can deposit a chosen target for a desired time, close a shutter and move on to the next target, essential for deposition of multilayers. Single-target systems are inexpensive and easier to use. Due to differing sputtering rates of various atoms, control of their relative content in a thin film is difficult. In general, sputtering is susceptible to effects such as negative ion bombardment, which are primarily due to oxygen ions bombarding and resputtering the deposited thin film. One way to minimize negative ion bombardment is to use an off-axis geometry in which a substrate is mounted transverse to the target assembly. Process optimization is difficult to achieve for deposition of complex oxides such as the HTS materials. Some of the process parameters that need to be optimized include the target stoichiometry, partial pressure of the inert gas, RF or dc power density, substrate-target distance, and the substrate temperature. High-quality YBCO, TBCO, and TPSCO superconducting thin films and multilayers have been deposited using off-axis magnetron sputtering (40, 41).

#### Metal-Organic Chemical Vapor Deposition

Metal-organic chemical vapor deposition is a proven technique for deposition of high-quality compound semiconductors. In principle, the advantages of MOCVD are the ease of large-area depositions, with high deposition rates and without a need for sophisticated high vacuum systems. However, the control of precise composition of the precursors is very difficult, and only small area films have been made to date. The MOCVD uses a deposition technique in which the required atomic constituents are transported in the form of vapors of metal-organic precursors by an inert transport gas such as argon or nitrogen, premixed with an oxidizing gas. Metal-organic precursors are placed in stainless-steel containers and heated to temperatures between 200 to 500°C to release vapors of the metal-organics. These vapors react with the heated substrate inside a reaction chamber. Mass flow controllers are used to control flow rates for each precursor. Mass flow rate, oxygen partial pressure and temperature of the substrate are some of the important process parameters. Typical metal-organic precursors for YBCO include metal  $\beta$ -diketonates such as  $\text{Y}(\text{C}_{11}\text{H}_{13}\text{O}_2)_3$ ,  $\text{Ba}(\text{C}_{11}\text{H}_{13}\text{O}_2)_2$ , and  $\text{Cu}(\text{C}_{11}\text{H}_{13}\text{O}_2)$  (42). A modification of this deposition technique is also currently used and is called the metal-organic decomposition (43). In this technique, metal carboxylates or acetates dissolved in an organic solvent are spin-coated on a substrate. The coated material is heat-treated at high temperatures to obtain the desired HTS phase. The technique is simpler to use because it does not require high vacuum equipment and is easily scalable for large area samples. The MOD does show promise for long-length wires. However, it is not pursued by many researchers due to the difficulty of producing films with only *c*-axis-oriented grains.

In our study, we have deposited *in-situ* films of YBCO on MgO substrates using the off-axis RF magnetron sputtering



process. Typical sputtering conditions for YBCO thin films are substrate temperature of 700 °C, sputtering gas pressure of 500 mTorr, oxygen partial pressure of 10 mTorr, RF power of 60 W, and a target dc bias of -30 V. We have also processed Tl2212 thin films by on-axis RF magnetron sputtering, and *ex-situ* processing. Tl2212 films were fabricated from both Tl2212- and Tl2223-sintered superconducting compounds at an RF power of 220 W, chamber pressure of 6 mTorr. Such thin films had to be postprocessed in an excess Tl<sub>2</sub>O partial pressure, provided in the form of sintered 2223 pellets in a platinum crucible used for sintering. In both cases, heat treatment times and temperatures were optimized to obtain smooth morphology, high phase purity, and superior electrical and microwave properties (44).

Sputtering of TlCaBaCuO thin films from a sintered powder target was performed using pure argon gas. The reason for not using oxygen and reactive sputtering is the volatility of Tl from the target. Tl readily combines with oxygen and forms Tl<sub>2</sub>O<sub>2</sub> which can be easily pumped out from the chamber. The target gets depleted of Tl much faster when oxygen is used as part of the reactive gases. The sputter-deposited thin films were postprocessed in two steps: first, sintering in air at 850 °C for 12 min to 15 min in an excess Tl<sub>2</sub>O partial pressure and, second, annealing in an oxygen flow of 500 sccm at 750 °C for 15 min to 30 min. An excess Tl partial pressure was maintained during the annealing process. Sintering was performed in a small box furnace in the free surface configuration (45). Thin films were placed on a sintered pellet of Tl2223 with the film side facing the free surface in a small covered platinum crucible. A second pellet was placed above the sample in a platinum wire mesh support. The pellets provided the excess Tl partial pressure in the crucible to minimize the loss of Tl from the thin film. Sintering of TlCaBaCuO thin film at 850 °C in air for 12 min to 15 min established the superconducting phase and morphology. At 850 °C, Tl<sub>2</sub>O<sub>2</sub> decomposes into a liquid Tl<sub>2</sub>O phase, which rapidly sinters with other constituents. After sintering, the crucible was removed from the furnace and cooled rapidly. Oxygen annealing of the air-sintered TlCaBaCuO thin film was carried out in the same configuration as used for sintering, with an oxygen flow of 500 sccm to 1000 sccm. During oxygen annealing, thin-film grains grow into large platelets. Oxygen annealing thus improves the electrical properties of the thin film by increasing its oxygen content.

Another promising method for the growth of high-quality Tl based HTS thin films is called the Tl<sub>2</sub>O<sub>2</sub> vapor process or the thallination process (46). In this process, thin films of the precursor materials of BaCaBaCuO or SrCaCuO are deposited using any of the thin-film deposition methods, and then the samples are heat-treated in a Tl<sub>2</sub>O<sub>2</sub> vapor pressure. Tl<sub>2</sub>O<sub>2</sub> vapor diffuses into the precursor and reacts to form the superconducting phases in the films. This technique is easier because the vacuum deposition chamber is not contaminated with the Tl<sub>2</sub>O<sub>2</sub> vapors.

#### Characterization of Processed Superconducting Thin Films

The XRD provides one of the direct characterizations of grown thin films because it gives information on the phase purity, their lattice constants, and their crystallographic orientation with respect to the film plane. Electrical transport measurements are performed for accurate measurement of transport current density at zero field as well as at finite magnetic fields,

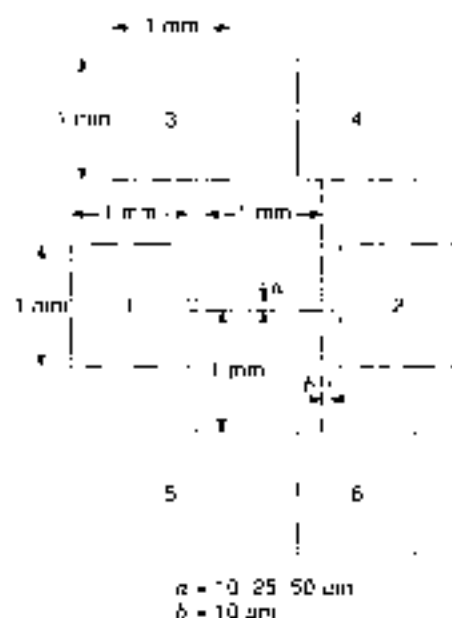


Figure 17. Geometry of four-probe test devices used for electrical transport measurements. A current source is connected between pads 1 and 2, and corresponding voltage measurements are taken across either pads 3 and 4 or pads 5 and 6.

The transport current measurements involve patterning a four-probe test structure on the HTS films. The four-probe test devices are designed with line widths of 10  $\mu\text{m}$  to 1 mm depending on the nature of the measurements.

The geometry of a typical test device is shown in Fig. 17. The test structure consists of two outer pads through which a constant current source is applied and a voltage is measured across the voltage sensing lines, which are 1 mm apart. The width of the voltage sensing should be less than the width of the line connecting pads 1 and 2 in order to approximate a point contact as closely as possible. The test devices are patterned on superconducting thin films using standard positive photoresist photolithography and a wet chemical etching technique using a 1:100 phosphoric acid:H<sub>2</sub>O solution. A pulsed current is applied between the outer pads, and the voltage across the inner sense lines is monitored. The electric field criterion of 1  $\mu\text{V}/\text{cm}$  is typically used (i.e., the current through the sample at which measurement of 1  $\mu\text{V}$  potential difference over a 1 cm spacing develops) gives an estimate of the superconductor's critical current. Also, the same test structure can be used to determine the  $T_c$  from temperature ( $T$ ) dependence of resistivity (47). In such a measurement, a constant dc current of 10  $\mu\text{A}$  is applied through the outer terminals, and the voltage across the sense lines is related to the resistance of the sample. Knowing the exact thickness of the superconducting thin film and the cross-section, the  $T$  dependence of resistivity plot can be obtained.

#### Structural, Electrical, and Magnetic Properties of Thin Films

High-quality YBCO and Tl2212 HTS thin films have been grown on various substrates such as YSZ, LaAlO<sub>3</sub>, GdAlO<sub>3</sub>, SrTiO<sub>3</sub>, SrTiO<sub>3</sub>, CaO, and MgO. Figure 18 shows an XRD scan of an annealed Tl2212 thin film on LAO substrate. The figure shows characteristic reflections of Tl2212, Tl2223 phases as well as those of the LAO substrate. From the prominent (0 0 1)

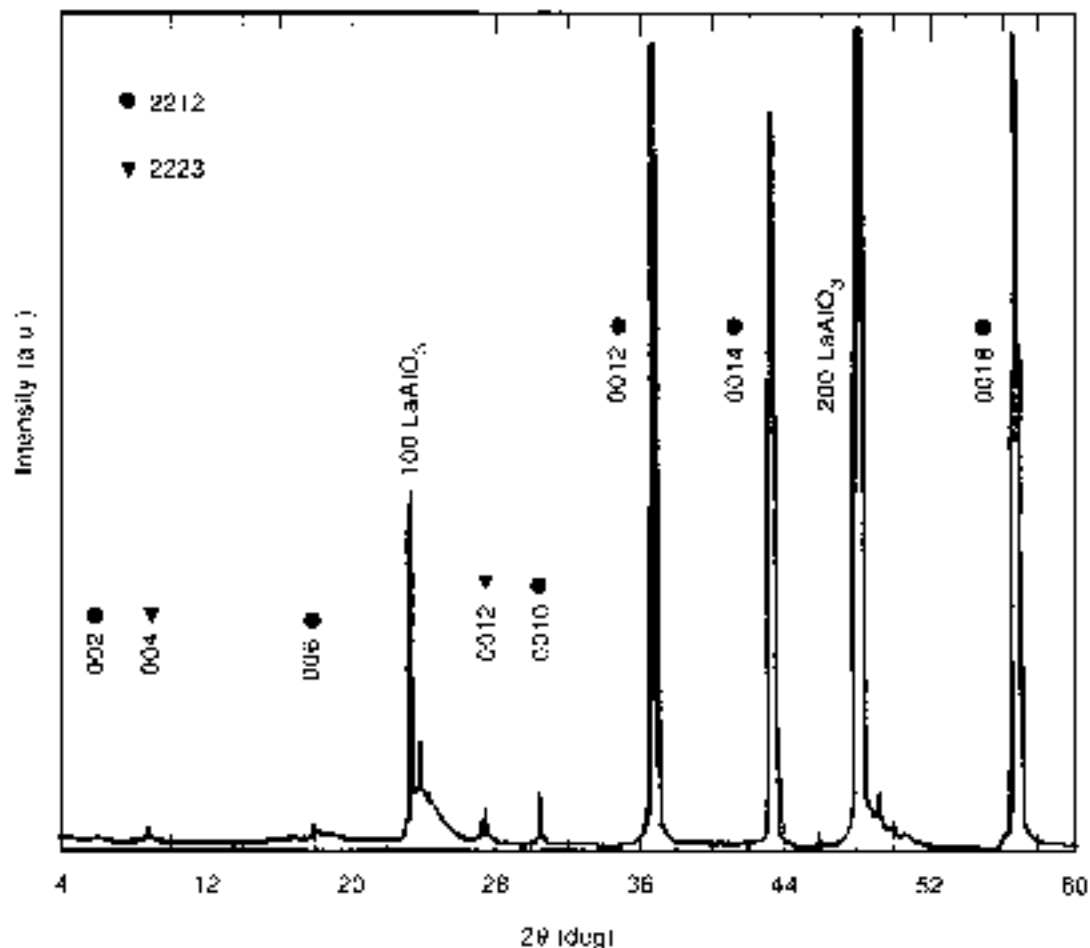


Figure 18. XRD reflections observed on a Tl2212 thin film on a LAO substrate. The reflections correspond to the Tl2212 phase, the Tl2223 phase, and the substrate.

reflections observed in the XRD spectrum, the highly *c*-axis-oriented growth is evident. The Tl2212 phase is the dominant one in the film, as determined from the intensities of the XRD peaks. The *c*-axis lattice constant calculated from the XRD scan is found to be 29.2 Å. The *a*-axis lattice constant of LaAlO<sub>3</sub> is 3.7801 Å. The *a*-axis lattice constant of Tl2212 phase is 3.8563 Å. Thus, the lattice mismatch is less than 2%.

Figure 19 shows the temperature dependence of resistivity for one of the Tl2212 thin films deposited on LAO, using a 50- $\mu$ m wide four-probe device with zero applied magnetic field. The zero-resistance  $T_c$  is approximately 100 K for the thin film. Figure 20 shows the typical zero field  $J_c$  variation with temperature obtained using a four-probe test device.  $J_c$  values at zero magnetic field as high as  $5 \times 10^4$  A/cm<sup>2</sup> at 77 K and approximately  $1 \times 10^6$  A/cm<sup>2</sup> at 60 K were obtained. The surface morphology of the films was essentially featureless and smooth, which is typical of high-quality films. In general, the  $J_c$  of Tl2212 and Tl1223 thin films are lower than the  $J_c$  of epitaxial YBCO thin films. The presence of grain boundary weak links and weak flux pinning in thallic thin films may be the main reasons for the lower  $J_c$ , compared to epitaxial *in-situ*-grown YBaCuO thin films. However, among the polycrystalline HTS, Tl2212 and Tl2223 thin films have shown superior electrical properties; hence, such films are very attractive for electronic applications (48). *In-situ* processing of Tl2212 and TlPh1223

superconducting thin films has not been highly successful owing to the complex chemistry of the thallic oxides. Tl2212 and Tl1223 thin films with very low surface resistance (at least two orders of magnitude below the oxygen-free copper conductor) at frequencies as high as 12 GHz and a current density greater than  $10^6$  A/cm<sup>2</sup> are routinely obtained (49).

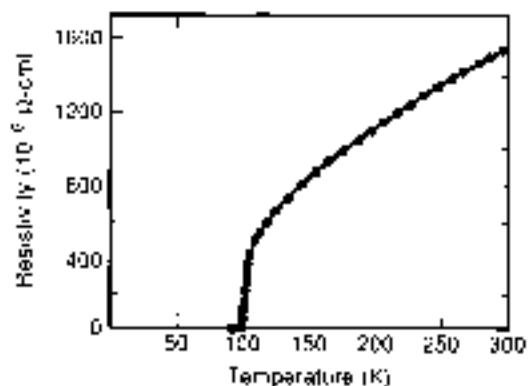
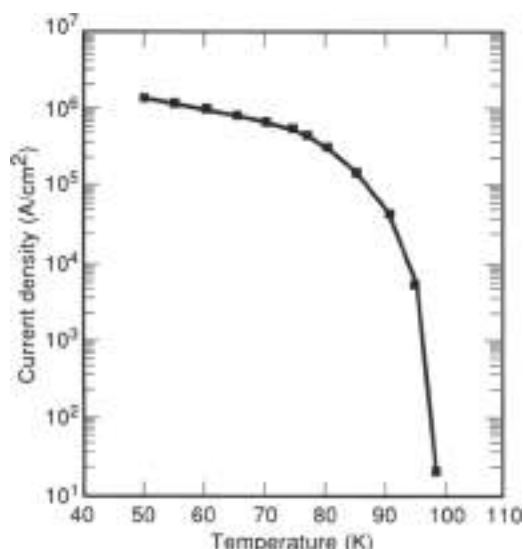


Figure 19. Temperature dependence of electrical resistivity for a 50- $\mu$ m wide Tl2212 thin film four-probe test device.



**Figure 20.**  $T_c$  pure temperature dependence of zero field current density for a postprocessed Tl2212 thin film. Measurements were performed on the four-probe test devices.

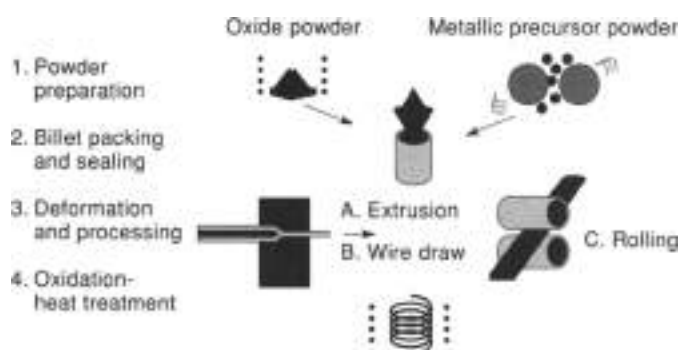
## PROCESSING OF SUPERCONDUCTING WIRES AND TAPES

### HTS Wires and Tapes, First Generation

The Bi2223 compound is perhaps the more attractive compound for processing conductors because its  $T_c$  (110 K) is higher than that of the Bi2212 compound (90 K). Surprisingly, the Bi2212 compound has better magnetic field dependence of  $J_c$  at 4 K, primarily due to better electromagnetic connectivity. At 77 K, the Bi2223 phase clearly has higher  $J_c$ s compared to the Bi2212 phase. The best BiSCCO HTS conductors have  $J_c$ s (KA/cm<sup>2</sup>) as high as 20 to 70 KA/cm<sup>2</sup> at zero field with  $J_c$  also depending on the length of the conductor (49). Oxide powder-in-tube (OPIT) has been a widely used technique for processing long lengths of BiSCCO wires.

### OPIT Method for BiSCCO 2212 and 2223 Superconducting Wires

Lead (Pb)-doped Bi2223 and Bi2212 are prime candidates for OPIT synthesis of long-length HTS wires. Of these two phases, Bi2212 is easier to synthesize. The reaction kinetics and thermodynamics for synthesis of the pure Bi2223 phase are quite complex and generally result in mixed phase samples, with the presence of Bi2223, Bi2212, and other secondary phases. Partial substitution of lead for bismuth is generally found to stabilize growth of the Bi2223 phase with a typical starting composition of Bi-Pb-Ca-Sr-Cu given by 1.3:0.4:2:2:3 for such growth (50). Figure 21 pictorially shows the four steps involved for the OPIT method of preparing HTS wires and tapes. In step 1, the precursor powder is prepared from a solid-state reaction (or calcination) of mixtures of Ba-, Cu-, and Pb-oxides and Sr- and Ca-carbonates in an alumina crucible at temperatures between 800 and 850 C for 48 to 60 h in air. Another approach for synthesis of the precursor material is to mechanically alloy the chemicals for higher packing density. In step 2, the precursor powder is filled inside a silver billet and then drawn into fine filaments less than 1 mm in diameter. Several tens to a hundred of these filaments are stacked in a silver tube,



**Figure 21.** Schematic diagram showing the general OPIT process used to produce BiSCCO tapes, developed by American Superconductor Corporation. Courtesy of Ref. 15.

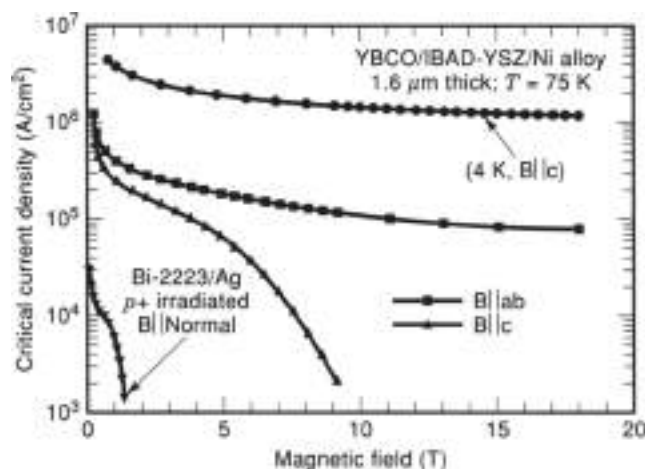
drawn, and then rolled into tapes 4 mm to 6 mm wide, and typically 0.2 mm thick in step 3. Finally, in step 4, OPIT tapes are processed in the presence of a partial melt, which aids in densification and grain growth. To create a partial melt, the sample is heated above 900 C for a short time, followed by a long anneal of approximately 100 h at 840 to 850 C. At present, such tapes can be reliably manufactured in several kilometer scale lengths. The main process parameters involved in the synthesis of wires thus include precursor material synthesis, mechanical processing, and heat treatment conditions. The precursor composition, particle size, and density are important factors in the quality of precursor material synthesis (51). Billet dimensions, filament restacking geometry, and mechanical properties of the tape are the important mechanical aspects for wires and tapes of HTS. Heating rate, sintering temperature, sintering time, and cooling rate are important parameters for heat treatment conditions that control electrical and mechanical properties of drawn wires. Heat treatment is performed after cutting wires to required lengths. Each filament is embedded in a silver matrix, which adds mechanical strength and flexibility. A threshold value for silver content is required to enhance electromagnetic grain connectivity, essential for long length conductors. At the same time, one needs to maximize the superconducting volume fraction so that the HTS rather than the silver present in the conductor provides the percolative path for charge carriers. Fill factor is one of the critical parameters in HTS wires, defined as the ratio of the volume of the superconductor core to the entire volume of the conductor. Typical fill factor in most HTS conductors to date is below 60% (51). Magneto-optic studies indicate that most of the current transport occurs at the interfacial region between silver and the bulk BiSCCO rather than in the bulk BiSCCO (52). The interfacial region with well-aligned grain structure appears to be the key to improve  $J_c$ s in these wires. The Bi2212 phase is attractive for low-temperature applications at 4 K because it is economical to produce long length multifilamentary conductors (52).

### HTS Wires and Tapes: Second Generation, YBCO-Coated Conductors

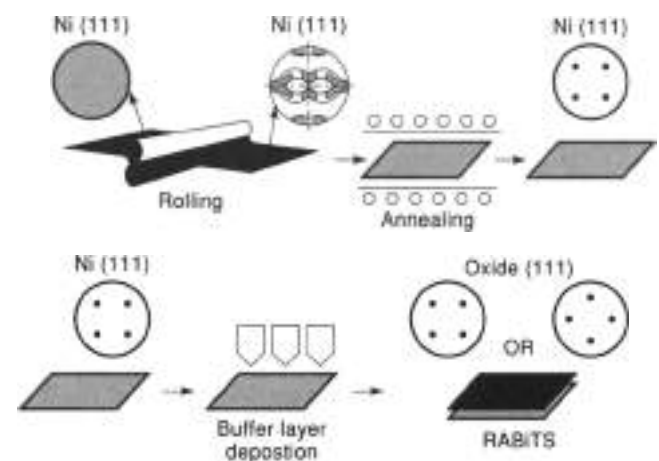
Primarily because of the higher  $J_c$ s demonstrated in biaxially textured thin films, YBCO has enormous advantage over BiSCCO for wire applications. One approach is to grow YBCO on a biaxially textured chemically compatible buffer layer. A processing technique proposed for biaxial texturing of chemically

compatible buffer layers on untextured substrates such as alloys of nickel is called the ion-beam-assisted deposition (IBAD) (53). Another approach proposed is the biaxial texturing of the substrates. It is well known that metals when deformed into wires develop alignment along preferred directions. Recent work at Oak Ridge National Laboratory has shown that biaxially textured (i.e., both in-plane and out-of-plane alignments) nickel can be produced over long lengths (54). Such substrates are called rolling assisted biaxially textured substrates (RABiTS) (54). An interesting example of biaxial texturing of metallic sheets already used in electrical engineering is in silicon-doped steel stampings used in the electric power industry, in which texturing results in the easy direction of magnetization parallel to the length of sheets, reducing hysteresis losses.

The IBAD process is used primarily for deposition of buffer layers such as yttrium-stabilized zirconia (YSZ), or cerium oxide ( $\text{CeO}_2$ ) on untextured nickel alloy tapes (53). The IBAD process is a thin-film deposition process developed by IBM for preferential growth of biaxially textured buffer layer material on untextured substrates. In this process, a rare gas ion beam is used to bombard the buffer layer along a preferred crystallographic direction as it is being deposited. The angle of incidence of the ion beam with respect to the plane of the substrate is the critical parameter for the preferential orientation of the buffer layer grown. Typical buffer layer of choice for IBAD has been YSZ, grown to a thickness of 200 nm to 500 nm. A thin layer of  $\text{CeO}_2$  is interposed between YSZ and YBCO for better lattice matching. Subsequent growth of postprocessed YBCO thin or thick films on the biaxially oriented buffer layers have yielded high  $J_c > 10^6 \text{ A/cm}^2$  at 75 K and in magnetic fields up to 5 T. Figure 22 shows the  $J_c$  versus magnetic field characteristics obtained for a short sample of a 1.6  $\mu\text{m}$  thick YBCO/IBAD-YSZ/Ni alloy tape, at 75 K (53). The figure also shows the magnetic field dependence of  $J_c$  for a proton-irradiated Bi-2223 OPIT wire for comparison. The figure clearly shows the vastly improved performance of the IBAD tape compared to the best OPIT wire. An advantage of the IBAD-buffered HTS tapes compared to OPIT wires is that the tapes can be wound with the HTS in compression, attractive for electromagnetic



**Figure 22.**  $J_c$  versus applied magnetic field for a 1.6  $\mu\text{m}$  IBAD tape as a function of magnitude and orientation with respect to the tape plane compared with a proton-irradiated OPIT Bi-2223/Ag wire. Courtesy of Ref. 14.



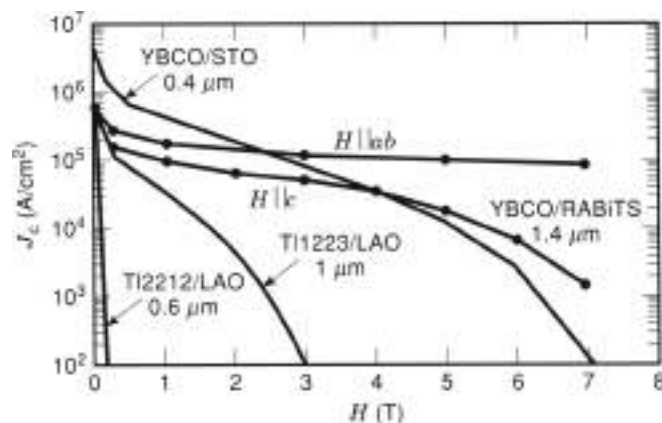
**Figure 23.** Schematic representation of the RABiTS process developed at Oak Ridge National Laboratory. Starting with a randomly oriented Ni (111) plate, rolling is used to produce a distinct copper-type rolling texture. This is followed by recrystallization to a cubic structure. Epitaxial metal and/or oxide buffer layer(s) are then deposited on the textured Ni. Courtesy of Ref. 54.

applications such as electric motors, and generators. Current efforts are underway to scale-up the process for long lengths necessary for commercialization.

The Rolling assisted biaxially textured substrates (RABiTS) process was developed by Oak Ridge National Laboratory for epitaxial growth of HTS thin films over long lengths. Figure 23 schematically shows the process for the development of RABiTS (54). Nickel is the primary candidate for the RABiTS because of its higher oxidation resistance compared with copper. Also, thermomechanical texturing of Ni gives a cube-textured substrate. A typical YBCO thin-film-coated conductor fabrication process involves thermomechanical processing of Ni substrate (125  $\mu\text{m}$ ), followed by PLD of buffer layers of  $\text{CeO}_2$  (0.2  $\mu\text{m}$ ), YSZ (0.2  $\mu\text{m}$ ), and finally 1  $\mu\text{m}$  thick YBCO film PLD deposited at  $\sim 750^\circ\text{C}$  in an oxygen pressure of 185 mTorr (54). After deposition, the film is cooled at 10  $^\circ\text{C}$  per min. in an oxygen pressure of 700 mTorr at 400  $^\circ\text{C}$ . All oxide buffer layers are grown at temperatures of 750  $^\circ\text{C}$ . RABiTS of lengths up to 100 feet (only the substrates with an HTS) have been processed already.  $J_c$  values of a short YBCO thin-film-coated conductor are shown in Fig. 24. For comparison, thin films of YBCO on LAO and thin films of Tl2212 and 1223 on LAO are also shown. Because PLD and sputtering techniques are not scalable, MOCVD may be the primary technique for long-length HTS coated conductors on RABiTS. The MOCVD-grown YBCO on RABiTS has yielded zero field  $J_c$  as high as 0.84  $\text{MA/cm}^2$  at 77 K.

## APPLICATIONS OF HIGH-TEMPERATURE SUPERCONDUCTORS

Bulk superconductors can be used in applications such as high  $Q$  cavities for microwave applications (see SUPERCONDUCTING MICROWAVE TECHNOLOGY), frictionless bearings, magnetic levitation (see SUPERCONDUCTING LEVITATION), and fault current limiters (see SUPERCONDUCTING FAULT CURRENT LIMITERS). Melt-textured as well as single crystal bulk YBCO conductors have been used as a microwave cavity with very high  $Q$  values greater than 10,000 at 77 K and at frequencies below 20 GHz. Progress in

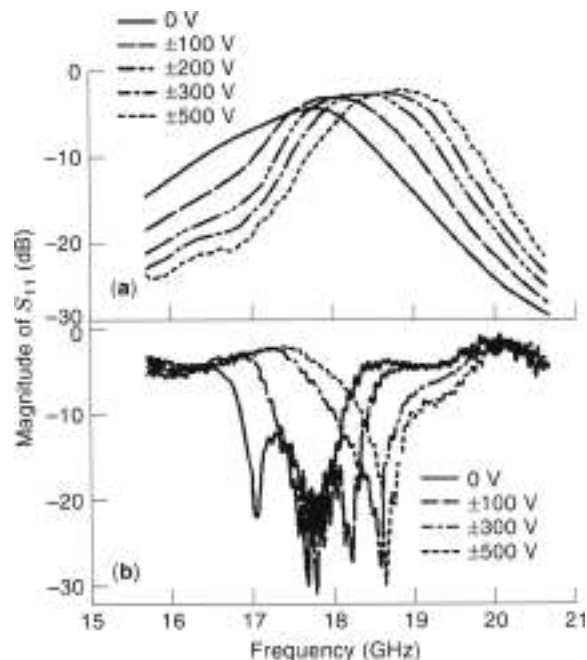


**Figure 24.** The critical current density versus magnetic field for YBCO thin film on a RABITS substrate of configuration ML3 corresponding to YSZ(140 nm)/NiO(100 nm)/Ni(100 nm) (25 wt. %). Data for both  $H \parallel c$  and  $H \parallel ab$  are shown. Also shown are data for YBCO thin film on STO, Tl2212, and Tl2223 thin films on LAO substrates. For  $H \parallel c$ , the high-field properties of YBCO thin films on RABITS are better than those on STO. Courtesy of Ref. 54.

bulk single-crystal growth could be attributed to newer crucible materials such as  $\text{BaZrO}_3$  (55). The use of  $\text{BaZrO}_3$  crucible precludes impurities such as Al, Zr in melts or crystals for the growth of 123 rare-earth-based HTS.

Thin film HTSs are most attractive for microwave applications because of their low surface resistance at frequencies up to 35 GHz (see SUPERCONDUCTING MICROWAVE TECHNOLOGY and SUPERCONDUCTING FILIBERS AND PASSIVE COMPONENTS). A large number of microwave components have been demonstrated using HTS technology, (e.g., high  $Q$  filters, resonators, phase shifters, delay lines). Epitaxial growth of HTS thin films on low-loss microwave substrates such as MgO and sapphire (with a  $\text{TiO}_2$  buffer layer) has been key for high  $Q$  microwave components. A promising new application in this area involves integration of HTS thin films and ferroelectric thin films for electrically tunable components (56). Tunable filters, resonators, and phase shifters have been successfully demonstrated. We have recently reported a two-pole tunable bandpass filter using a thin-film YBCO/STO/LAO multilayer microstrip configuration. The structure consists of a LAO substrate (254  $\mu\text{m}$  thick) on which an epitaxial STO thin film (0.3  $\mu\text{m}$  thick) was deposited by PLD, followed by a YBCO thin film (0.35  $\mu\text{m}$  thick) deposited by PLD. The two-pole filters were designed for a center frequency of 19 GHz and 4% bandwidth. Tunability is achieved through the nonlinear dc electric field dependence of relative dielectric constant of STO thin films. Figure 25 shows the swept frequency response of a tunable filter at 77 K, for bipolar bias voltages between 0 and  $\pm 500$  V (56). Note that the insertion and return losses are reduced with applied bias. Although high voltages are required, the power consumption is low owing to negligible currents through the high dielectric constant ferroelectric thin film.

Another area of electronics in which HTS thin films have a niche is in the use of HTS SQUIDS (see SQUIDS) as magnetic field sensors. A SQUID operates on the principle that the critical current through two Josephson junctions connected in a loop is a function of the magnetic flux threading the loop. Since each SQUID stores multiples of the basic flux quanta, SQUIDS are the most sensitive sensors of magnetic fields at



**Figure 25.** Electric field tunability of a tunable bandpass filter using a thin film YBCO/STO/LAO multilayered microstrip structure. The filter uses the electric field dependence of the nonlinear dielectric constant of the STO ferroelectric thin film.

levels of  $10^{-11}$  T at 77 K, thus making possible measurements of magnetic fields produced by current flow in a human brain, (about  $10^{-22}$  T) (see SQUIDS). The key to the HTS SQUID technology is the reproducible fabrication of grain boundary Josephson junctions. This technology has currently matured to the level of commercial usage in magnetic imaging for medical diagnostics such as a magnetencephalography (MEG), mapping the brain's magnetic activity. Another area of application for HTS thin films is an interconnect technology for semiconductor multichip modules. With the promise of high-quality YBCO thin films on various buffer layers such as  $\text{CeO}_2$  and YSZ, it is feasible to conceive of applications of HTS thin films on silicon substrates.

Applications such as nuclear magnetic resonance (NMR) require uniform magnetic fields over large volumes. Nuclear magnetic resonance requires a field level of 10 T to 20 T. The NMR technique uses RF and magnetic fields to study transitions between nuclear spin states. An RF coil surrounds a sample, placed inside a magnetic field. The magnetic field is swept through a small range with a separate unit. A secondary RF coil may be a superconducting coil picks up the resonance signal. Use of superconducting coils in magnet technology has made possible high magnetic fields with minimal energy dissipation. Use of superconducting lines for power transmission is attractive for energy efficiency. The current densities needed for the utility power lines are considerably lower than most other power applications. It is estimated that current densities below  $10^4$  A/cm<sup>2</sup> at 77 K are sufficient for utility power lines. The Bi-based superconductors are the most attractive for these applications. The design of underground cables cooled to liquid nitrogen temperatures is a challenging aspect because these superconductors are detrimentally affected by moderate mechanical stresses. Superconducting magnetic energy storage

**Table 3. Companies Producing HTS Materials**

American Superconductor Corporation  
Two Technology Drive  
Westborough, MA 01581  
(508) 830-4200

Conductor, Inc.  
959 W. Maude Avenue  
Sunnyvale, CA 94086  
(408) 523-9950

DuPont Superconductivity  
Experimental Station E3049/1110  
Wilmington DE 19880  
(302) 695-9230

Intermagetics General Corporation  
PO Box 461, 450 Old Niskayuna Road  
Latham, NY 12110-0461  
(518) 782-3122

Superconductive Components, Inc.  
1145 Chesapeake Avenue  
Columbus, OH 43212  
(614) 486-0261

Superconductor Technologies, Inc.  
400 Ward Drive  
Santa Barbara, CA 93111  
(805) 683-7546

(SMES) is an energy storage device wherein energy is stored in the magnetic field associated with circulating currents in superconductors. An SMES system consists of an ac-to-power converter to charge a large toroidal magnet. When needed, the energy stored in the magnetic field can be tapped and converted by ac power. Again HTS wires are needed for the large magnets used in such a system. Other applications attractive for HTS wires and tapes include power devices such as motors, generators and transformers. With the current technology, several prototypes of HTS wires and tapes have been demonstrated. American Superconductor Corporation had tested a 325 hp ac synchronous motor built using a BiSCCO 2223 composite tape. A 5 kJ HTS SMES magnet, a high-current pancake coil (100 A), and current development of 500 and 1000 hp motors indicate that HTSs have potential for large-scale power applications (49). Table 3 shows some of the HTS companies which provide superconducting materials and fabrication services.

## BIBLIOGRAPHY

- W. Meissner and R. Oschenfeld, Ein Neuer Effekt bei Eintritt der Supraleitfähigkeit, *Naturwiss.*, **21**: 787, 1933.
- F. London and H. London, The electromagnetic equations of the superconductor, *Proc. R. Soc. A*, **A149**: 71-86, 1935.
- F. London, *Superfluids, Microscopic Theory of Superconductivity*, vol. 1 New York: Wiley, 1950, pp. 2-4.
- J. Bardeen, L. N. Cooper, and J. Schrieffer, Theory of superconductivity, *Phys. Rev.*, **108**(2): 1175-1204, 1957.
- B. D. Josephson, Possible new effects in superconductive tunnelling, *Phys. Lett.*, **17**: 251-253, 1962.
- J. C. Bejanz and K. A. Muller, Possible HTS superconductivity in the Ba-Lu-Cu-O system, *Z. Phys. B*, **64**: 199-191, 1985.
- M. K. Wu et al., Superconductivity at 93 K in a new mixed phase Y-Ba-Cu-O compound system at ambient pressure, *Phys. Rev. Lett.*, **58**(9): 909-910, 1987.
- H. Maeda et al., A new HTS oxide superconductor without a rare earth element, *Jpn. J. Appl. Phys.*, **27**(2): L203-210, 1988.
- Z. Z. Sheng and A. M. Hermann, Bulk superconductivity at 126 K in the Tl-Ca-Ba-Cu-O system, *Nature*, **332**: 16160; 138-139, 1998.
- S. N. Peilini et al., Superconductivity at 94 K in HgBaCuO, *Nature*, London, **362**: 236, 1993.
- J. D. Das, *Engineer's Guide to High Temperature Superconductivity*, New York: Wiley, 1989.
- A. A. Abrikosov, On the magnetic properties of superconductors of the second group, *Soviet Phys. JETP*, **5**: 1174, 1957.
- H. P. Huebener et al., Thermoelectric and thermomagnetic effects in high temperature superconductors, *Proc. Materials Res. Soc. Symp.*, **1982**, pp. 13-22.
- E. M. Grant, Superconductivity and electric power: Promises, Promises . . . Past, present, and future, *IEEE Trans. Appl. Supercond.*, **7**: 112-136, 1997.
- T. P. Sheshan, *Introduction to High Temperature Superconductivity*, New York: Plenum, 1994.
- H. A. Blackstead and J. D. Dow, Implications of Abrikosov-Gorkov exchange scattering for theories of high temperature superconductivity, *Phys. Lett. A*, **206**: 107-110, 1996.
- H. C. LeFebvriere, Road to conductors of high temperature superconductors: 10 years do make a difference! *IEEE Trans. Appl. Supercond.*, **7**: 90-96, 1997.
- C. Blue and P. Bouchaud, *In-situ* preparation of superconducting YBCO 123 thin films by on-axis rf magnetron sputtering from a stoichiometric target, *Appl. Phys. Lett.*, **56**: 2036-2039, 1991.
- S. Tnaka, Reviews, prospects and concluding remarks: Materials needs for applications, *Physica C*, **232-237**: xxxi-xxxix, 1997.
- M. A. Kirk and H. W. Weber, Electron microscopy investigations of irradiation defects in the high T<sub>c</sub> superconductor YBCO, in *Studies of High Temperature Superconductors*, vol. 10 Commack, NY: Nova, 1992, p. 241.
- S. Jin et al., Melt textured growth of polycrystalline YBCO with high transport J<sub>c</sub> at 77 K, *Phys. Rev. B*, **37**: 7650-7653, 1988.
- S. S. P. Parkin et al., Bulk superconductivity at 125 K in Tl<sub>2</sub>Ca<sub>2</sub>Ba<sub>2</sub>Cu<sub>3</sub>O, *Phys. Rev. Lett.*, **60**(124): 2539-2542, 1988.
- M. Greenblatt et al., Chemistry and superconductivity of thallium based cuprates, *Studies of the HTS Superconductors*, vol. 4 Commack, NY: Nova, 1992.
- M. Murakami et al., Melt processing of bulk YBCO superconductors with high J<sub>c</sub>, *J. of Eng. Materials Tech.*, **114**: 138, 1992.
- M. Murakami et al., Melt processing of bulk high T<sub>c</sub> superconductors and their applications, *IEEE Trans. Magn.*, **27**: 1479, 1991.
- V. K. Toti et al., Processing of large YBCO domains for levitation applications by a NdBaCuO seeded melt growth technique, *J. Electron. Materials*, **23**: 1127, 1994.
- K. Salama and S. Sathyanarayana, Melt texturing of YBCO for high current applications, *Appl. Superconductivity*, **4**(10-11): 547-561, 1996.
- Z. Sheng, Tl-based high T<sub>c</sub> superconductors, *Mater. Sci. Forum*, **130-132**: 1-36, 1993.
- G. Subramanyam et al., Processing and physical properties of single phase TlBaCuCuO (2212) and (Tl-Pb)SrBaCuCuO (1224) superconductors, *Appl. Superconductivity*, **4**: 591-598, 1996.
- T. Kama et al., *Appl. Phys. Lett.*, **58**: 3186-3188, 1991.
- C. P. Bean, Magnetization of high field superconductors, *Rev. Mod. Phys.*, **36**: 31-39, 1964.
- W. M. Kauf, Magnetic susceptibility measurements: An important part of modern solid-state characterization, *Amer. Lab.*, **20**, February 1988.

33. L. H. Greene and D. G. Bagley, Oxygen stoichiometric effects and related atomic substitutions in the high  $T_c$  cuprates, in *Physical Properties of High Temperature Superconductors*, vol. II, Singapore: World Scientific, 1990, pp. 589-599.

34. R. Sugisu, et al., *Jpn. J. Appl. Phys.*, **27**, L2310, 1988.

35. B. Raveau et al., Layered thallium cuprates: Non-stoichiometry and superconductivity, *Solid State Ionics*, **39**: 49-62, 1990.

36. F. Shi et al., Metal (M) dopant centered local structures, high-pressure synthesis and bulk superconductivity in  $YBa_2Cu_3M_{1-x}O_{7-x}$  (M: Fe, Co, Ni), *J. Condens. Matter*, **9**: L307-L313, 1997.

37. A. Rykov, V. Caignaert, and B. Raveau, Quadrupole interactions and vibrational anisotropy of tetragonal Fe(II) in the 123 derivative  $LaSrCuGa_{1-x}Fe_xO_{7-x}$  ( $La - Y, Ho$ ), *J. Solid State Chem.*, **109**: 285, 1994.

38. J. Qian and C. Y. Yang, High  $T_c$  superconductors on buffered adatom: Materials properties and device applications, *Materials Sci. Eng. R14*: 157-202, 1995.

39. D. C. Payne and J. C. Dravman, (eds.), Laser ablation for materials synthesis, Materials Research Society (MRS) Fall Meeting, MRS Symp. Proc., Vol. 191, Boston, 1990.

40. D. S. Ginley, Thallium films for micro-electronic applications, in: A. M. Hermann, and Y. Yakhmi (ed.), *Thallium based superconducting compounds*, Singapore: World Scientific, 1994.

41. W. Holstein et al., Superconducting epitaxial Tl2212 films on sapphire with cerium oxide buffer layers, *Appl. Phys. Lett.*, **61**(16): 1949-1954, 1992.

42. J. Zhao and P. Norris, Metalorganic chemical vapor deposition of YBCO 123 thin films, *Mater. Sci. Forum.*, **180-182**: 243-254, 1994.

43. G. Braunstein et al., Process of formation and epitaxial alignment of SrTiO<sub>3</sub> thin films prepared by metal organic decomposition, *J. Appl. Phys.*, **78**: 8611-70, 1993.

44. G. Subramanyam et al., Studies on sputtered TlCaBaCuO high  $T_c$  superconducting thin films for microelectronic applications, *Materials Sci. Forum.*, **130-132**: 613-632, 1993.

45. D. S. Ginley et al., Morphology control and high critical currents in superconducting thin films in the Tl-Ca-Ba-Cu-O system, *Physica C*, **160**: 42-48, 1989.

46. B. Jubs et al., Preparation of high  $T_c$  Tl-Ba-Ca-Cu-O thin films by pulsed laser evaporation and Ti<sub>2</sub>O<sub>3</sub> vapor processing, *Appl. Phys. Lett.*, **64**: 1910-1912, 1989.

47. W. Felster, Accurate automated measurement of superconductor material resistance, *Rev. & Develop.*, **30**: 80-94, 1988.

48. DuPont Superconductivity, Technical Note on H based superconducting thin films, Wilmington, DE, 1996.

49. A. P. Malozemoff, Q. Li, and S. Fleshler, Progress in BSCCO-2223 tape technology, *Physica C*, **248-249**: 424-427, 1997.

50. U. Balachandran et al., Processing and fabrication of high  $T_c$  superconductors for electric power applications, *Appl. Superconductivity*, **5**(1-4): 197-191, 1997.

51. V. Selvamanickam et al., High temperature superconductors for electric power and high energy physics, *J. Metals*, **60**(10): 27-30, 1988.

52. S. Brady, The last days of BSCCO, *Superconductor and Cryoelectronics*, featured article on BSCCO wires, **11**: 26-30, Spring, 1995.

53. Y. Iijima et al., In plane aligned YBCO thin films deposited on polycrystalline perovskite substrates, *Appl. Phys. Lett.*, **60**: 769-771, 1992.

54. A. Goyal et al., Epitaxial superconductors on rolling-ascorted biaxially textured substrates (RBXTS): A route towards high critical current density wire, *Appl. Superconductivity*, **4**(11-12): 403-427, 1996.

55. A. Erb, E. Walker, and R. Flukiger, The use of BaZrO<sub>3</sub> crucibles in crystal growth of the high  $T_c$  superconductors: Progress in crystal growth as well as in sample quality, *Physica C*, **258**: 9-20, 1996.

56. E. A. Miranda et al., Tumble microwave components for Ku and K-band satellite communications, *Integrated Electronics*, **22**: 269-274, 1999.

GURU SUBRAMANYAM  
University of Dayton  
PHENIX BOULCHAND  
University of Guernsey

## HTS SUPERCONDUCTORS, PHYSICAL STRUCTURES, AND ROLE OF CONSTITUENTS

Since the discovery of the phenomenon of superconductivity in 1911 by H. Kamerlingh Onnes (1), superconductors have been a topic of great interest from fundamental as well as application points of view. With the discovery of high-temperature superconductors (HTSCs) in 1986 (2), the field of superconductivity achieved even greater interest. Since then HTSCs have moved into the realm of practical applications because they are mostly superconducting above the boiling point of liquid nitrogen (77 K), which is cheaper than and has a much greater cooling capacity than liquid helium. Therefore, for any application in which liquid nitrogen can replace liquid helium, the refrigeration cost for operating superconductive devices will be about 1000 times less.

The newly discovered HTSCs are copper-oxide based ceramics that remain superconducting near 100 K. This temperature regime does not seem high, but it is high enough for liquid nitrogen to cool most of HTSCs into the superconducting range. For example, the compound yttrium barium copper oxide (YBCO) has been found to be superconducting up to 92 K (3). Additional important high-temperature superconductors include bismuth strontium calcium copper oxide (BSCCO) (4), thallium barium calcium copper oxide (TBCCO) (5), and mercury barium calcium copper oxide (HBCCO) (6). HBCCO has the highest critical temperature  $T_c = 135$  K (7, 8). Table 1 presents the chemical formula and  $T_c$  values of each of these compounds. The HTSCs can be subdivided into three classes: three-dimensional compounds based on BaBiO<sub>3</sub>, layered copper-oxide compounds with hole conductivity based on La<sub>2</sub>CuO<sub>4</sub>, YBa<sub>2</sub>Cu<sub>3</sub>O<sub>7-x</sub>, and compounds of Bi and Tl, and compounds based on Nd<sub>2</sub>CuO<sub>4</sub> with electron conductivity. With increasing number of CuO<sub>2</sub> planes in copper-oxide superconductors, a certain increase in  $T_c$  is observed. Its values range from  $T_c = 36$  K in La<sub>2-x</sub>Sr<sub>x</sub>CuO<sub>4</sub> with a single CuO<sub>2</sub> plane to  $T_c = 125$  K in compounds of Tl-2223 with three CuO<sub>2</sub> planes.

Table 1. List of Common High- $T_c$  Superconductors

| Name                                   | Formula   | $T_c$ , K |
|--|---|-----------|
| Yttrium barium copper oxide            | YBaCuO  | 91        |
| Bismuth strontium calcium copper oxide | Bi <sub>2</sub> (Pb) <sub>2</sub> Sr <sub>2</sub> CaCu <sub>2</sub> O | 105       |
| Thallium barium calcium copper oxide   | TlBa <sub>2</sub> Ca <sub>2</sub> Cu <sub>2</sub> O                   | 115       |
| Mercury barium calcium copper oxide    | HgBa <sub>2</sub> Ca <sub>2</sub> Cu <sub>2</sub> O                   | 135       |

## COMMON FEATURES OF HIGH- $T_c$ SUPERCONDUCTORS

### Coherence Length and London Penetration Depth

In addition to the crystallographic aspects of the HTSC oxides, there are more fundamental parameters contributing to the anisotropic nature of these materials. Central to the theory of superconductivity is the Cooper electron pair discovered by Cooper in 1956 (9), when he showed that it is possible for electrons to have an attractive interaction leading to a lowering of energy of the system. One result of the pairing is that electric current can flow without resistance, giving rise to superconductivity. The critical temperature is related to the thermal energy required to destroy the pairing between the paired electrons. Earlier in 1950, Ginzberg and Landau showed that the spatial variation of the particles or superelectrons carrying the resistanceless current can be described by the coherence length  $\xi$ . The parameter  $\xi$  gives the length scale for the density variation of Cooper pairs, and simply put, the greater the length, the larger the hurdle the pairs can cross without breaking up. The breaking of pairs means the destruction of superconductivity, and so care must be taken to keep the material free of features that break these pairs. The other important length in superconductivity is the London penetration depth  $\lambda$ , which sets the scale for the decay of an external magnetic field as a function of depth in the superconducting material. There is a strong parallel between the skin depth of metals in response to high-frequency fields and the London depth in response to magnetic fields. In 1935, the London brothers first introduced the concept of penetration depth while following a phenomenological approach to superconductivity via Maxwell's equations for electromagnetism (10):

Ginzberg and Landau (GL) showed that the ratio of  $\kappa = \xi/\lambda$  determined the nature of the response of the superconductor to a magnetic field. For  $\kappa \leq 1/\sqrt{2}$ , type-I superconductivity exists, and there exists a critical field  $H_c$ , when superconductivity and diamagnetism are abruptly destroyed, while for  $\kappa \geq 1/\sqrt{2}$ , type-II superconductivity exists, and there exists a lower critical field  $H_{c1}$  above which materials are not perfectly diamagnetic and an upper critical field  $H_{c2}$  beyond which superconductivity and diamagnetism are completely destroyed. All the HTSC oxide superconductors are type II and are potentially useful because they can sustain high magnetic fields and remain superconducting.

### Anisotropy

In addition to the crystallographic anisotropy, there are a number of important properties that also show anisotropy such as electrical conductivity, thermal conductivity, magnetic field penetration, and current density. The large variation in electrical and thermal conductivity in directions parallel to and perpendicular to the  $c$  axis is a direct consequence of the conductivity of the  $\text{CuO}_2$  layers. Ratios of 100 and 17 for the electrical and thermal conductivities parallel to the  $a$ - $b$  plane and perpendicular to it have been measured for  $\text{YBCO}$  (11). These values are even higher in Bi- and Tl-based systems, indicating strong anisotropy parallel to and perpendicular to the  $\text{CuO}_2$  layers. Anisotropy in response to magnetic fields and current density is a result of the large variation in the coherence length with direction. For  $\text{YBCO}$ ,  $\xi$  is a few nanometers along the  $a$ - $b$  plane and an order of magnitude less perpendicular to it. This

produces a large variation in the magnetic field and current density it can support perpendicular to and parallel to the  $a$ - $b$  plane. Similar features are observed in the other HTSC systems.

## STRUCTURES OF HIGH- $T_c$ SUPERCONDUCTORS

The knowledge of structure is a prerequisite for understanding the properties of high- $T_c$  superconductors. The structures of high- $T_c$  superconductors can be easily analyzed by considering the structure of a perovskite. The original mineral called perovskite was  $\text{CaTiO}_3$  and was described in the early nineteenth century by Gustav Rose. In addition to the importance of perovskite structures in superconductivity, dielectric, piezoelectric, pyroelectric, and ferroelectric properties of compounds with perovskite structures have found widespread applications in memory devices, capacitors, tunable filters, microphones, loudspeakers, etc.

The chemical formula for a perovskite structure is  $\text{ABX}_3$ , where  $A$  and  $B$  are metal cations and  $X$  is a nonmetal anion.  $X$  is either oxygen (O) or a member of the halogen family, such as fluorine, bromine, or chlorine. The perovskite cell consists of a relatively large metallic atom ( $A$ ) at the corners (or center), smaller metallic atoms ( $B$ ) at the center (or corners), and the nonmetallic atoms ( $X$ ) at face centers (or midpoint between the corners). Figure 1 illustrates the position of the various atoms in a cube. The chemical formula of  $\text{ABX}_3$  is easily deduced from this figure if the eight unit cells sharing the corner  $A$  atom and the two unit cells sharing the  $X$  atoms are considered. The atom positions can be designated as  $A$ : (0, 0, 0) corner sites;  $X$ : (0, 1/2, 1/2), (1, 1/2, 1/2), (1/2, 1/2, 0), (1/2, 1/2, 1), (1/2, 0, 1/2), (1/2, 1, 1/2); edge centers;  $B$ : (1/2, 1/2, 1/2) body center.

Another representation of the perovskite structure that is extremely useful to the understanding of HTSC oxide structures is the layered or planar approach in which the structure is formed by stacking of layers in the  $z$  direction. In this approach, the molecular grouping comprising each layer is also shown mainly in order to stress the importance of the layered approach to an understanding of the properties of the HTSC oxides. The layers themselves usually do not have any significance and may not even exist as independent structures. In

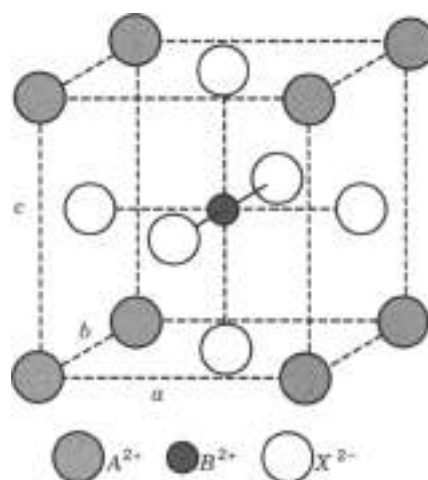


Figure 1. A schematic representation of a perovskite structure.



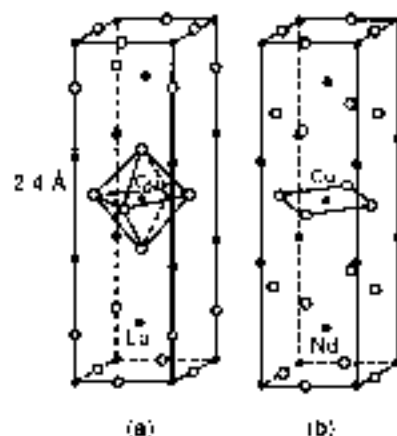
HTSC oxide terminology, the long axis is denoted by  $c$  (or  $a'$ ) as compared to the smaller  $a$  or  $b$  axis.

An ideal perovskite is cubic, like  $\text{CaTiO}_3$ . Its strong ionic bonding makes  $\text{CaTiO}_3$  electrically insulating and a hard mineral. It also shows three-dimensional isotropy in many properties due to the cubic symmetry. The crystallographic nature of anisotropy or isotropy does not necessarily determine the nature of fundamental properties of materials such as its electrical conductivity. For instance, the element gallium exists as an anisotropic hexagonal structure, but the momentum distribution of its electrons termed *momentum space* is isotropic, and similar features could occur for superconductors. In general, there are a number of variations from the ideal form of  $\text{CaTiO}_3$ . A typical example is the diversity in properties demonstrated by  $\text{BaTiO}_3$ . The crystal structure of  $\text{BaTiO}_3$  undergoes structural phase transitions as a function of the ambient temperature, changing its symmetry and response to external electric fields. Below 5 °C it has an orthorhombic structure, between 5 °C and 210 °C it is tetragonal, while above 210 °C it is cubic. In the tetragonal state, the structural distortion results in a permanent electric dipole moment, making it ferroelectric, while in the cubic phase the symmetry prevents the existence of a permanent dipole moment. In noncubic perovskite minerals the phenomenon of twinning is commonly observed. Twinning is most easily described by considering a structure obtained by interchanging the  $a$  and  $b$  axis in a tetragonal phase, resulting in regions that are mirror images of each other.

In addition to these common structural features, the role of the  $A$  and  $B$  cations is also crucial in the properties of perovskites. The replacement of the smaller  $\text{Ca}^{2+}$  with an ionic radius of 0.99 Å with the larger  $\text{Ba}^{2+}$  (1.34 Å) plays a central role in the tetragonal distortion leading to the ferroelectric property of  $\text{BaTiO}_3$ . The doping or substitution of  $A$  or  $B$  cations with other ions having radii within 15%, but with different valence, results in a charge imbalance that is compensated by the creation of holes or equivalently by vacancies in the  $X$  sites. This charge imbalance results in nonstoichiometric perovskite structures. All the HTSC oxide superconductors contain  $\text{Cu}^{2+}$  instead of  $\text{Ti}^{4+}$ , resulting in a  $\text{CuO}_2$  layer that plays a central role in the nature of electric current transport in these materials. Because of this feature all the HTSC oxides exhibit an almost uniform size of  $a$  and  $b$  axes (usually between 3.81 Å and 3.83 Å). In addition to these aspects, structural transformations, twinning, anisotropy, etc., are all key features in controlling the superconducting properties and need to be studied in detail to obtain a clearer picture of the behavior of these complex and potentially useful structures.

#### Structure of La-(Ba,Sr,Ca)-Cu-O and Nd(Ce)-Cu-O Compounds

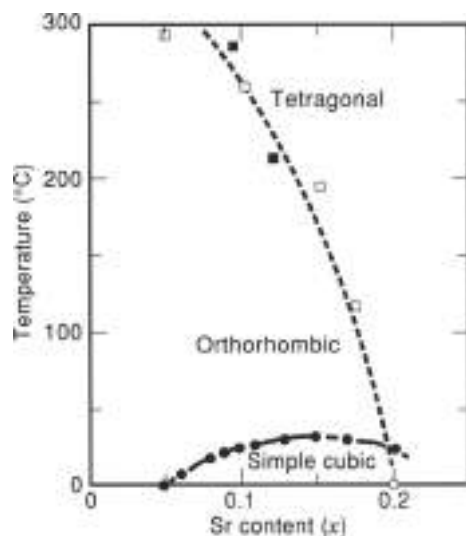
Superconductivity in the high- $T_c$  oxides was first discovered in the  $\text{La}_2\text{CuO}_4$  system when Bednorz and Müller (2) observed a sharp drop in the electric resistance in Ho-doped  $\text{La}_{2-x}\text{Ba}_x\text{CuO}_4$ . Subsequent observation of the Meissner effect confirmed that the compound was superconducting. The  $T_c$  was maximized at 30 K for a value of  $x = 0.15$ . The structure of the superconducting phase was identified as a derivative of the layered perovskite  $\text{K}_2\text{NiF}_4$  and is shown in Fig. 2. The  $\text{CuO}_2$  layers consist of octahedrally coordinated Cu, with the Cu-O bond length of 1.9 Å in the  $a$ - $b$  plane and 2.4 Å in the  $c$  direction. Also, the Cu atoms occur at (000) and (1/2, 1/2, 1/2) $\times$  lattice sites in the unit cell, similar to a body-centered structure. The substituted  $\text{Ba}^{2+}$



**Figure 2.** Comparison of structures of La- and Nd-based superconductors. (a) Structure of  $\text{La}_{2-x}\text{Ba}_x\text{CuO}_4$  with sixfold-coordinated Cu. This is a  $p$ -type superconductor. (b) Structure of  $\text{Nd}_2\text{CuO}_4$  with planar fourfold coordinated Cu. When some  $\text{Nd}^{3+}$  ions are replaced by  $\text{Ce}^{4+}$ , this material becomes an  $n$ -type superconductor.

ions occupy a fraction of the  $\text{La}^{3+}$  sites, and the difference in valence brings about an increase in the number of holes. Superconductors with hole conduction are called  $p$  type, a terminology similar to that followed in semiconductors. While substitution with atoms of different valence is a well-known technique to introduce excess holes or electrons (once again in parallel with semiconductors), oxidation or reduction to change the oxygen content also changes carrier concentration. Interstitial oxygen between the  $\text{LaO}$  sheets or missing oxygen from the lattice sites, that is, vacancies, can also lead to excess charge carriers. For example, the electrochemical oxidation of  $\text{La}_2\text{CuO}_4$  results in  $p$ -type superconductivity (12). Substitution of the smaller (1.13 Å)  $\text{Sr}^{2+}$  ions instead of  $\text{Ba}^{2+}$  (1.34 Å) results in a higher  $T_c$  (39 K) for  $x = 0.15$ . The increase in  $T_c$  by substitution of a smaller atom and also the fact that increasing pressure tends to increase  $T_c$  (i.e., in the  $\text{La}_{1-x}\text{Ba}_x\text{CuO}_4$  phase,  $T_c$  increases from 31 K to 36 K for a pressure increase from atmosphere to 1.2 GPa (13)) is a significant property for a fundamental understanding of high-temperature superconductivity and it also points the way to the synthesis of oxide superconductors with higher  $T_c$ .

The structure of compounds in the  $\text{Nd}_2\text{CuO}_4$  system is identical to the compounds in the La-based copper oxides. The only difference arises in the positions of the oxygen atoms, giving rise to an  $\text{O}_2$  layer instead of the  $\text{La-O}$  layer present in La-based structures. While this does not change the  $\text{CuO}_2$  layers, it results in a square planar fourfold copper coordination as seen in Fig. 2, which compares the structures of the La- and Nd-based superconductors. The most important feature of Nd-based materials derives from the fact that superconductivity in this system is by electroconduction or  $n$ -type rather than hole conduction as in La-based superconductors. The difference in conductivity between the two systems, which in the undoped state have almost identical structures and trivalent lanthanum or neodymium, is owing to the effect of doping the substituting La with divalent Sr and the Nd with tetravalent Ce. In one case it results in a decrease in the number of electrons ( $\text{La}^{3+} + \text{Ce}^{4+} \rightarrow \text{Sr}^{2+}$ ), leading to excess holes, while in the other it increases the number of electrons ( $\text{Nd}^{3+} + \text{Ce}^{4+} \rightarrow \text{Ce}^{4+}$ ), leading to  $n$ -type conductivity. The electroconducting compound  $\text{Nd}_{1-x}\text{Ce}_x\text{CuO}_4$  with  $T_c = 24$  K is the most studied material in this system (14).



**Figure 3.** Structural phase change in La-based copper oxides as function of Sr concentration. (Adapted from Ref. 15)

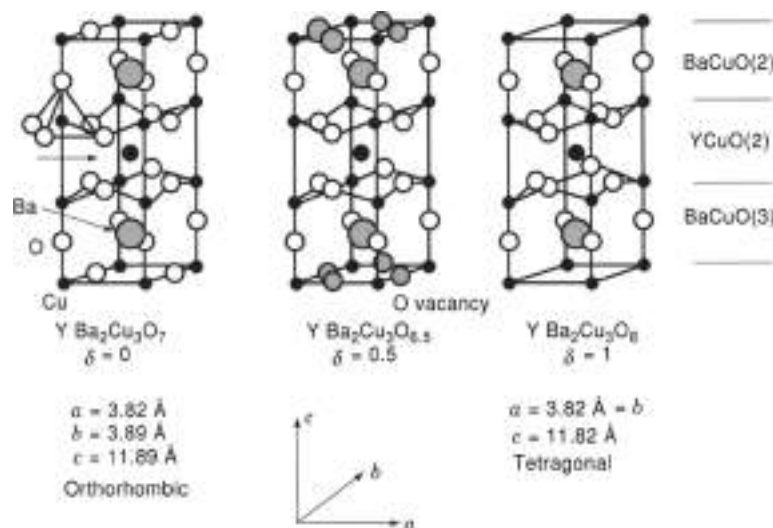
Another common feature observed among the HTSC oxide families is easily observed in  $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$ . Here, the lattice parameters and crystallographic symmetry depend on the value of  $x$ . For  $x = 0.2$  and  $B = \text{Sr}$ , the structure is tetragonal, while it is orthorhombic for  $x < 0.2$ . Figure 3 shows the structural phase diagram for the  $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$  system (15). Similar features can be found for  $B = \text{Ba}$ . One aspect of the structures in La-based compounds is that superconductivity is observed both in the orthorhombic and tetragonal phase, a feature not seen in the Y-Ba-Cu-O system, where only the orthorhombic phase superconducts.

#### Y-Ba-Cu-O Compounds

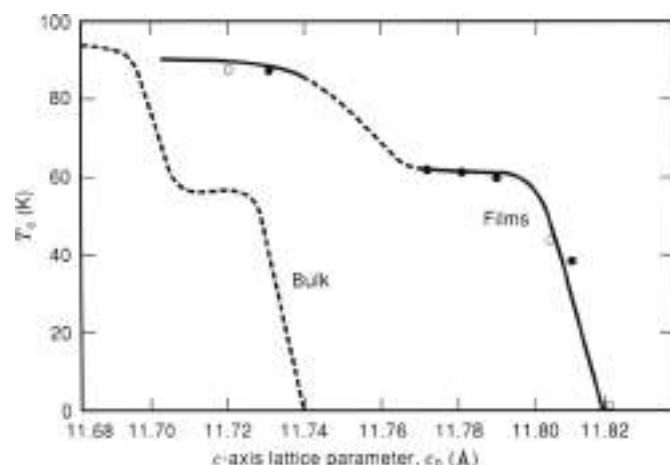
The increase in  $T_c$  observed by increasing pressure and by substitution of smaller atoms led Uhm and his co-workers (17) to replace  $\text{La}^{3+}$  ionic radius of 1.14 Å by the smaller  $\text{Y}^{3+}$  (0.94 Å), leading to the first superconductor with a transition temperature ( $T_c$ ) higher than the boiling point of liquid  $\text{N}_2$ .

(77 K). The compound synthesized was  $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$  with a  $T_c$  of 92 K.  $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$  is also commonly referred to in the literature as Y123 or generally as YBCO and is the most widely studied of all the HTSC oxides. Its superconducting properties, such as  $T_c$ , the maximum current that it can carry (measured by the current flowing per unit cross-section area of the superconducting material,  $i_c$ ), the critical current density ( $J_c$ ), and the maximum strength of magnetic field that it could expel via the Meissner effect ( $H_{c1}$ ), are all highly sensitive to the crystal structure and crystal quality of the YBCO material.

In terms of the perovskite structure,  $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$  is modeled by stacking three perovskite-type cells of type  $\text{BaCuO}_2$ ,  $\text{YCuO}_2$ , and  $\text{BaCuO}_2$  as shown in Fig. 4. A  $\text{BaO}$  plane separates two layers containing Cu and O. While one contains atoms in the ratio of 1:2, that is,  $\text{CuO}_2$ , the other contains a varying quantity of O, ranging from  $\text{CuO}$  to  $\text{Cu}$ . The two layers are commonly referred to as the  $\text{CuO}_2$  plane or  $\text{CuO}$  chain, respectively. The quantity of oxygen in the chains is the key to properties of  $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$  and is represented by the value of  $\delta$ , known as the oxygenation parameter. For  $\delta = 1$ , the chains do not contain O (Fig. 4), while for  $\delta = 0$ , the chains are fully oxygenated, giving a  $\text{CuO}$  layer. The  $\delta$  value also determines the crystal structure of YBCO, and Fig. 4 (left to right) shows the transition from an orthorhombic structure ( $a = 3.82$  Å,  $b = 3.89$  Å,  $c = 11.68$  Å) for  $\delta = 0$  to a tetragonal structure for  $\delta = 1$  ( $a = b = 3.86$  Å,  $c = 11.62$  Å). Significantly, in the tetragonal phase obtained for  $\delta \geq 0.6$ , the material is insulating and does not have a superconducting transition, while in the orthorhombic state obtained for  $\delta \geq 0.6$ , it is metallic and superconducting.  $T_c$  is seen to maximize at 92 K as  $\delta$  approaches zero.  $\delta$  can be measured either using the titration method or from the lattice parameter of YBCO obtained from X-ray diffraction (XRD) measurements using an empirical formula,  $\delta = (c - 11.68) / 0.1501$ , suggested by Tranquada et al. (16). Here,  $c$  is the value of the  $c$  parameter obtained from XRD measurements. Figure 5 shows the experimentally observed correlation between the  $c$ -axis lattice parameter and  $T_c$  for YBCO in both thin-film form (17) and bulk form (18). The higher values of the  $c$ -axis lattice parameter for YBCO thin film than those for bulk YBCO with the same  $T_c$  are attributed to the expansion of the unit cell under the influence of the substrate material.



**Figure 4.** Structure of  $\text{Y}_1\text{Ba}_2\text{Cu}_3\text{O}_{7-x}$ . The oxygen content ( $\delta$ ) varies from 1 (rightmost figure) to 0 (leftmost figure) changing the structure from a tetragonal insulating phase to the superconducting orthorhombic YBCO phase. The fivefold coordination of the Cu in the copper oxide plane is also shown. The chain oxygen content in the  $a-b$  plane varies with  $\delta$ . The direction of the  $\text{CuO}_2$  planes is also seen.



**Figure 5.** Experimental correlation between the  $c$ -axis lattice parameter and  $T_c$  for YBCO thin film form 141 and YBCO bulk form. (Adapted from Ref. 16.)

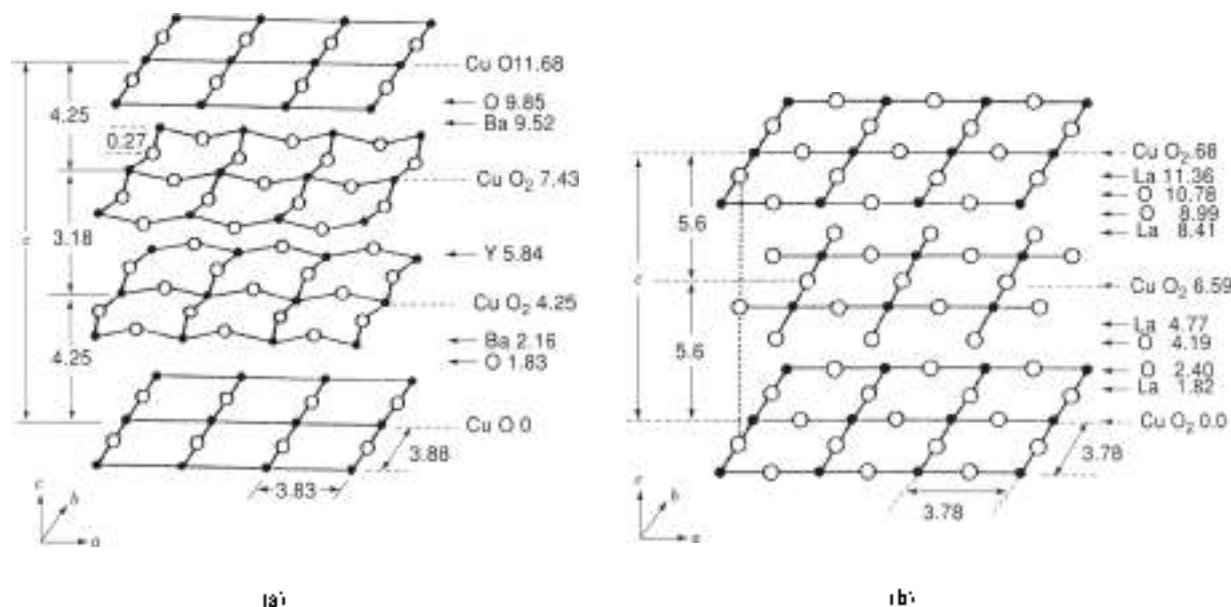
In many cases, it is easier to interpret the properties of the HTSC oxides by understanding the stacking sequence of the individual Cu-O layers making up the unit cell. But, unlike the case of true layered compounds such as graphite and mica, where the weak bonds between layers makes intercalation with foreign atoms easy, this is not the case with the HTSC oxides, and so the layers must not be interpreted as such. In the case of the Bi-based HTSC oxides, though, the binding between BiO sheets is weak and intercalation of iodine is possible. Figures 6(a) and 6(b) compare the layering sequences in YBCO and  $\text{La}_2\text{CuO}_4$ . The main difference seems to be in the form of the position of the individual Cu atoms. For the case of YBCO, all the Cu atoms are of one type; that is, edge type or all centered (depending upon the representation), while for the La-based structure, they occupy (100) and (1/2, 1/2, 1/2) po-

sitions. The Y layer acts as a spacer for the two  $\text{CuO}_2$  planes in YBCO. Also unlike the case of La-Sr-Cu-O, the atoms on the  $\text{CuO}_2$  layer in YBCO do not all lie on the same plane, that is, their  $z$  values differ slightly and so they are actually  $\text{CuO}_2$  sheets as compared with  $\text{CuO}_2$  planes. This slight distortion is visible in Figs. 4 and 6(a). Heating YBCO through 350 C transforms the orthorhombic structure to the tetragonal phase owing to the loss of oxygen. This displacing transition is the reason for twinning and is usually seen when the strain accompanying the phase change cannot be completely released, as in the case of crystalline thin films deposited on substrates. Thus, YBCO thin films undergo twinning when grown at high temperatures and subsequently cooled through 350 C in an oxygen ambient. Figure 7 shows a typical YBCO film showing twinning observed under an electron microscope.

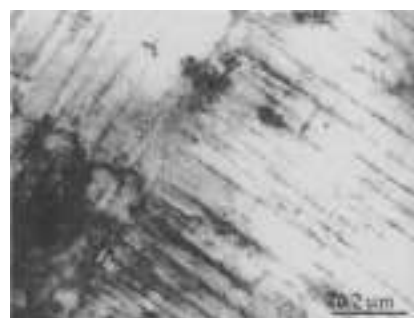
In addition to YBCO, the Y-Ba-Cu-O family also contains other phases of which  $\text{YBa}_2\text{Cu}_3\text{O}_7$  (Y124) has a  $T_c$  of ~90 K and  $\text{Y}_2\text{Ba}_4\text{Cu}_7\text{O}_{13}$  (Y247) has a  $T_c$  of ~55 K. These two phases are not observed under normal processing conditions and require extremely high oxygen pressures or the addition of alkali-metal compounds in air at normal pressures (19). The structure of Y124 is similar to that of YBCO with the addition of an extra CuO chain layer displaced with respect to the first by a lattice parameter of  $b/2$  where the  $b$  axis contains the oxygen of the chain. Y124 is orthorhombic with  $a = 3.84 \text{ \AA}$ ,  $b = 3.87 \text{ \AA}$ ,  $c = 27.2 \text{ \AA}$ . The structure of Y247 can be represented by the intergrowth of alternating layers of YBCO and Y124, resulting in an orthorhombic structure of  $a = 3.86 \text{ \AA}$ ,  $b = 3.87 \text{ \AA}$ ,  $c = 50.3 \text{ \AA}$ . As compared to YBCO, Y124 has a fixed oxygen stoichiometry and so does not undergo the displacive transformation resulting in the twinning morphology.

### Bi-, Tl-, and Hg-Based Compounds

In early 1985, two new superconducting systems based on bismuth and the toxic metal thallium with transition



**Figure 6.** Layered approach to the structure of copper-oxide superconductors. (a) The  $\text{Y}_2\text{Ba}_2\text{Cu}_3\text{O}_7$  structure with the individual planes and plane spacings in the  $c$  direction marked. (b) Similar approach for the La-based superconductors. (Adapted from C. D. Paele, H. A. Farach, and R. J. Creswick, in: *Su.*, San Diego, CA: Academic, 1985, pp. 182-191.)



**Figure 7.** Twinning in  $Y_1Ba_2Cu_3O_{7-x}$ . The figure shows a transmission electron microscope (TEM) image of a YBCO film grown on MgO by pulsed laser deposition. The structures correspond to the twin boundaries.

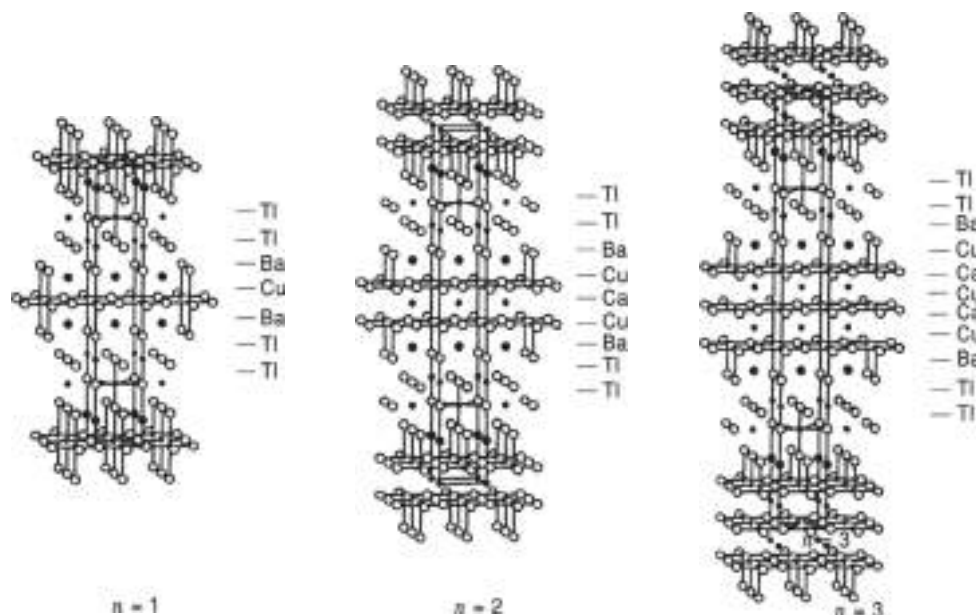
temperatures considerably higher than YBCO) were discovered (20, 21). These two systems have *a* and *b* lattice parameters similar to the Y- and La-based compounds, but with a much larger *c* parameter. The general formula for compounds occurring in this system is given by  $A_2B_2Ca_{n-1}Cu_nO_{2n+4}$ , where *A* = Bi or Tl, *B* = Sr or Ba, and *n* = 1 to 4. The compounds in these two systems have identical structures and differ only in atom positions, so it is convenient to deal with their structures together. The highest  $T_c$  recorded so far in these systems are 123 K for *n* = 3 in the Tl system, giving  $Tl_2Sr_2CaCu_3O_{10}$ , also referred to as Tl2223, and 110 K for the structurally similar Bi2223. The toxicity of the Tl oxides has prevented extensive characterization of these compounds as compared to the Bi structures. The crystal structures (Fig. 8) are similar to those of YBCO, but the CuO chain layers are instead replaced by a double layer of TlO. The absence of chains makes the Bi- and Tl-based oxides less susceptible to the stoichiometry of oxygen, though nonstoichiometry is commonly present in these systems due frequently to Ca:Sr or Cu depletion. The multiplicity of phases is a major problem in the synthesis of homogeneous material in these systems.

The Cu coordination is fivefold square pyramidal with O as for YBCO (Fig. 4). But for *n* = 3, the middle CuO layer contains Cu, which is planar fourfold coordinated. The layer in these systems consists of  $CuO_2$ -Ca-CuO<sub>2</sub> layers sandwiched between alkaline-earth-metal oxide B:Sr or Ba-O layers and by a double layer of A(Tl or Bi)-O. An interesting feature shown by these compounds is the variation of  $T_c$  with the number of CuO<sub>2</sub> planes, *n*. The  $T_c$  is found to increase with increasing *n*, maximizing at *n* = 3. While this confirms the widely accepted fact that the CuO<sub>2</sub> layers play a vital role in superconductivity, for example, the structures with *n* = 0 do not superconduct; increasing the number of layers indefinitely does not necessarily increase  $T_c$ .

In addition to these stoichiometric compounds, long sintering times can result in the evaporation of part of an A-O layer, resulting in a monolayer of Tl-O or Bi-O, giving nonstoichiometric oxides.  $T_c$ 's in the Tl-based compounds vary from 10 K to 121 K depending on the number of copper oxide planes. The family of systems given by the formula  $A_2B_2Ca_{n-1}Cu_nO_{2n+4}$  or equivalently by  $A_2B_2Ca_xCu_{n-1}O_{2x}$  (here *n* does not correspond to the number of CuO<sub>2</sub> layers) can be extended to other cations such as mercury (Hg), which forms a series of compounds (with A<sub>2</sub> replaced by Hg and B by Ba) with  $T_c$ 's ranging from 95 K to 133 K, again maximizing for three CuO<sub>2</sub> layers.

#### ROLE OF MICROSTRUCTURES IN HIGH- $T_c$ SUPERCONDUCTORS

After the initial excitement and directions towards large-scale technological applications of the HTSC oxides following their discovery in 1986, scientists have realized the difficulty in putting these ceramic oxides to use because they are hard, brittle, and difficult to draw into sheets or wires. The major emphasis in materials science and technology is now on developing the processing and fabrication of thin films for devices. Their novel properties, such as the ability to shield closed volumes from electromagnetic interferences via the Meissner effect, the low



**Figure 8.** Crystal structure of the Tl-Bi based superconductors. Changing the number of Cu atoms increases copper oxide layers (*n*). The structure of the Bi-based compounds is similar, with the Tl positions occupied by Bi. (Adapted from Ref. 21.)

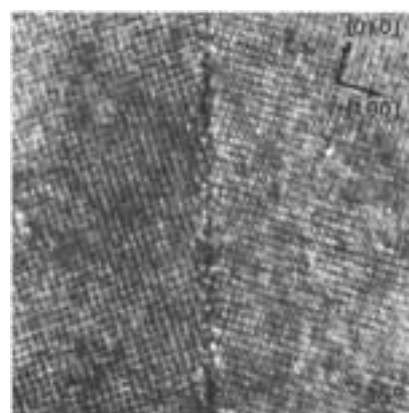
microwave surface resistance for use in filters for cellular base stations, the Josephson effect in superconducting quantum interference devices (SQUIDs), and the sharp transition in the use of baluns to detect infrared radiation are in stages of development and use. The exciting application is the use in microelectronics as interconnects—the reduction in Joule heating due to resistanceless current flow will increase the density of devices per chip and the efficiency of packaging. The improvement in  $I/C$  time constants, increasing the speed of devices, also makes it attractive to consider the integration between superconductors and semiconductors to operate at liquid-nitrogen temperatures. In fact, there are a number of transistors that improve performance when cooled to 77 K.

However, there are a number of problems to be solved, some, such as the magnitude of electric current that HTSC oxides carry and the erratic behavior of weak links in Josephson junctions across grain boundaries, are functions of the Cooper pair coherence length and material quality, while others, such as the degradation on exposure to moisture of YBCO and the formation of ohmic contacts to these oxides, can be solved by proper combination of materials and processing conditions. Thus the material quality or microstructure of thin films seems to be the most important aspect of thin-film study of HTSC superconductors. This leads us to the large influence of grain boundaries on the current-carrying ability of HTSC oxides. While the discussion here pertains directly to YBCO, its extension to other HTSC oxides seems highly likely owing to the central role played by coherence length. As mentioned earlier, the coherence length in oxide superconductors is on the order of a few nanometers, meaning that any imperfection in the material with an influence scaling across dimensions of the same magnitude will break Cooper pairs, mainly because they are formed owing to certain properties exhibited by the perfect crystal structure.

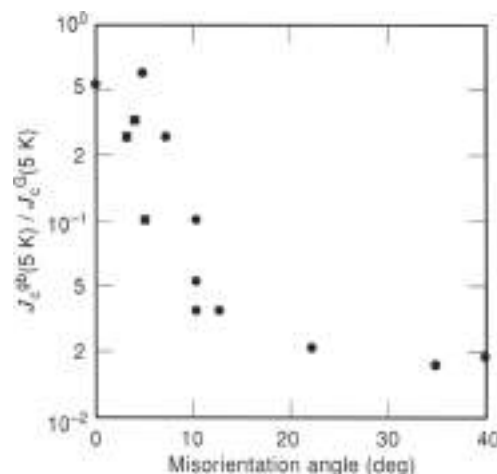
Figure 9 shows an electron micrograph of a grain boundary occurring in a YBCO film with a misorientation between the grains being quite large (25°). It is now well accepted that when the Cooper pairs strike the boundary at an angle, some pass through unbroken, while others break and become nonsuperconducting, resulting in a reduction in the current that emerges on the other side of the boundary. In HTSC terminology, the processes are referred to as pair tunneling and pair breaking, respectively. The grain boundary is said to constitute a *weak*

*link* as a result of this process. Actually, there are two subtle processes involved across boundaries in the HTSC oxides: the weak-link behavior is a result of pair tunneling across an insulating region separating the two superconducting regions or semiconductor-insulator-semiconductor tunneling, as it is referred to, while the other is the proximity effect, and occurs across a metallic or normal boundary, that is, a superconductor-normal conductor-superconductor structure. These two processes have characteristic current-voltage behavior and are easily discernible in classical superconductors where the coherence lengths are large, almost 10 nm to 100 nm. In HTSC oxides, these two processes cannot be differentiated because they occur on length scales of a few nanometers, the same as the coherence length.

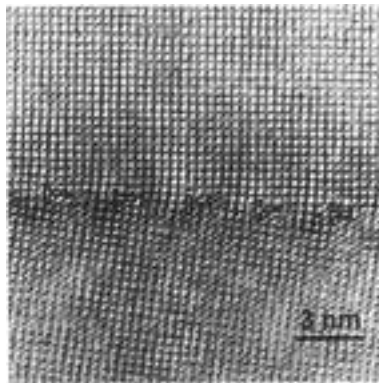
The significant role of grain boundaries has been emphasized by some lucid experimental demonstrations like those by researchers at IBM (22, 23), where they showed that not only does the critical current density ( $J_c$ ) of the material decrease due to the presence of the boundary, but the  $J_c$  also depends on the misorientation angle of the boundary. Figure 10 is a plot of the experimental result of  $J_c$  across grain boundaries in YBCO films grown by pulsed laser deposition (PLD) (24, 25). The plot clearly shows that increasing the misorientation decreases the  $J_c$ . This has important implications for devices. High-current-carrying films need boundaries with extremely small misorientations or better still no boundaries at all. To explain the reason for this behavior a number of researchers (21, 26, 27) have invoked the dislocation structure of low-angle grain boundaries first proposed by Read and Shockley in 1950 (28). Figure 11 demonstrates this idea in a low-angle boundary in a YBCO film grown by PLD. The almost regular separation of the dislocations marked in the figure generates the boundary misorientation, and in fact a plot of the dislocation separation versus misorientation (Fig. 12) shows a trend similar to that followed by the  $J_c$  in Fig. 10. This similarity was an irresistible attraction to model the current flow as a strong function of the grain-boundary features arising from the dislocation arrays. This includes the finite region along the boundary where the lattice was highly distorted and so superconductivity can be destroyed (29). This is usually observed near or around dislocations like that shown in Fig. 13.



**Figure 9.** Large-angle (25°) grain boundary in  $Y_{1.85}Ba_{0.15}Cu_3O_{7-x}$ . This figure shows a high-resolution TEM image of a typical high-angle pure (101) tilt grain boundary in YBCO.



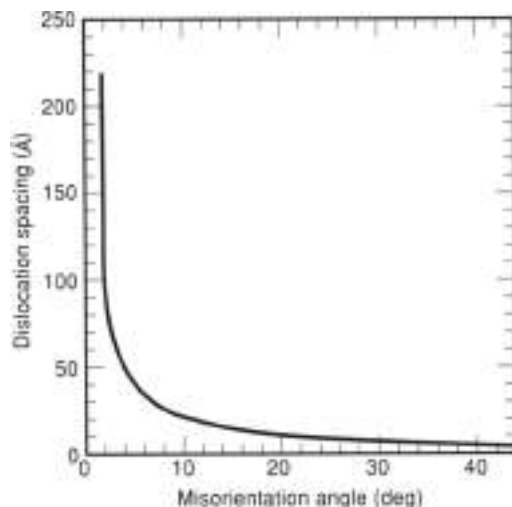
**Figure 10.** Experimentally observed variation of the critical current density ( $J_c$ ) with boundary misorientation (°). (Adapted Ref. 23.)



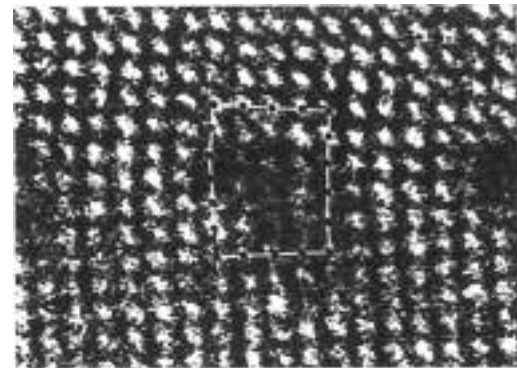
**Figure 11.** Low-angle TEM grain boundary in  $\text{YBaCuO}$ , showing the dislocation array. This figure shows a high-resolution TEM image of a low-angle pure (100) tilt boundary in a YBCO film grown by pulsed laser deposition. The boundary misorientation is generated by the equally spaced (100) edge dislocations.

The long-range strain fields are due to the dislocations (30), that distort the crystal structure and affect superconductivity by changing the order parameter (31). The order parameter is the complex quantum-mechanical wave function used to describe the density of Cooper pairs as a function of space, similar to a wave function describing the electron density in a metal, for example, the ability of Cooper pairs to tunnel across small distortions that are present along the boundary surface due to strains (32). This feature is a function of the boundary width, which describes the region along the boundary in which current flows via tunneling rather than normal current. This parameter is experimentally difficult to observe and various approximations such as the strength of the Burgers vector or via energy minimization (33) are used to calculate it. Another important influence of grain boundaries is the effect of nonstoichiometry commonly observed along the boundary (34), leading to excess charge and subsequently scattering of the Cooper pairs.

Jagannathan and Narayan have proposed an elegant model to represent the  $J_c$  across low-angle boundaries that



**Figure 12.** The variation of dislocation spacing with boundary misorientation ( $\theta$ ). The geometrically calculated variation in spacing parallels the experimental observation of  $J_c$  with misorientation shown in Fig. 10.



**Figure 13.** Dislocation core. This high-resolution TEM image shows the highly distorted structure of the YBCO lattice around a [100] edge dislocation.

incorporates all the features mentioned before. While the previous calculations can be done fairly simply, the actual nature of the grain boundaries in HTSC materials needs to be understood fully. It has been shown by various authors that there is a large variation in grain-boundary microstructure even in films prepared under identical conditions (35, 36). Recently, it has been seen that even boundaries with a similar misorientation angle along identical crystallographic directions can have different atomic scale structures (37, 38).

Compared with strong variation in the low-angle region of 0 to 10° observed in Fig. 10, the  $J_c$  in the high-angle region of 10° or more is uniformly low. An early explanation of the JHM group for this saturation effect was the overlapping of dislocations as the boundary misorientation increased. While the notion of the dislocation structure boundaries loses significance at higher angles, it is difficult to even model the current in the high-angle regime for the lack of a uniform description of the boundaries occurring there. Various representations such as the near-coincidence site lattice model (near CSL) or more recently the grain-boundary structure representation in terms of distinct structural units (38, 39) are being used to understand the properties of the high-angle boundaries. It has also been difficult to interpret exceptions to the low- $J_c$  in high-angle cases, especially for the 90° twist boundary, formed by rotating one unit cell of YBCO 90° with respect to an adjoining cell with the rotation axis being perpendicular to the  $c$  axis or [001] direction (40, 41). Here the absence of the weak link behavior is a strong suggestion for the fact that high  $J_c$ 's are associated with boundaries possessing low energy, also usually accompanied by high symmetry, which is true for the 90° twist boundary. In this light the predominance of asymmetry in grain boundaries in YBCO grown by PLD is an important feature and must be understood more clearly. All these factors lead one to realize that the exact nature of the effect of grain boundaries is far from being fully comprehended and detailed research must be pursued for a better understanding.

#### ROLE OF SUBSTITUENTS IN HIGH- $T_c$ SUPERCONDUCTORS

The simple copper oxide perovskites are insulators. By substituting for certain atoms in the unit cell, these materials are made to behave as metals and hence possibly as superconductors. The key characteristic of any superconductor is an energy

gap exactly at the Fermi level. Owing to this characteristic, Cooper pairs of electrons do not scatter off the lattice site and break. In the absence of scattering, the propagation of electrons encounters no resistance and the phenomenon of superconductivity is observed. The transition temperature depends strongly on the electronic density of states at the Fermi level. That parameter in turn, is strongly affected by doping of the initial ceramic with other atoms of a different valence to provide extra electrons (or too few electrons, called holes), which are then available for participation in the superconducting mechanism.

The practice of doping has a long history in the field of semiconductors, and chemists have developed great skills in modifying the properties of materials to allow practical electronic devices to be made. It is not surprising to see similar efforts applied to the HTSCs. A tremendous array of substitutions have been tried to improve the mechanical, magnetic, or current-carrying properties of these materials. The complexities of multicomponent phase diagrams prevent some attempted compounds from forming at all. Of those that can be made, either in chemical equilibrium or via a narrow path of kinetics, success is not guaranteed because it is hard to change the electronic properties precisely at the Fermi level. Moreover, the Fermi level is not one single number, but rather there is a Fermi surface in energy space, which is badly distorted from a simple spherical shape by the anisotropy present in the HTSC compounds. Consequently, there is considerable empiricism and guesswork present in any particular choice of chemical substitution in the HTSCs.

#### Role of Oxygen and Oxygen Vacancies

The role of oxygen in influencing the electronic properties of HTSCs is very crucial. Normally, an oxygen atom takes on two electrons from another atom. Therefore, if it is absent, two more electrons are free to go elsewhere in the crystal. This is how vacancies affect the charge balance in a crystal. The first thing that oxygen vacancies do is change the number of free carriers available in the crystal lattice, which in turn adjusts the Fermi level slightly. The density of states at the Fermi level ( $N_0$ ) is a key parameter of superconductors; the transition temperature  $T_c$  depends on  $N_0$  as  $\exp(-1/N_0V)$ , where  $V$  is the Cooper pairing potential. Small changes in  $N_0$  caused by oxygen vacancies can translate into substantial changes in  $T_c$ . It does not take much to change  $N_0$  appreciably. There are about  $8.5 \times 10^{11}$  vacancies/cm<sup>2</sup> in a double-layer compound such as BSCCO, which corresponds to having about 1% of the oxygen atoms missing.

Explaining the role of oxygen vacancies depends upon the charge-transfer model, by which the electrons normally in the  $\text{CuO}_2$  planes are transferred to sites elsewhere in the unit cell (42). In YBCO, the oxygen deficiency shows up especially in  $\text{CuO}$  chains, making their formula  $\text{CuO}_{1-x}$ , while the  $\text{CuO}_2$  layers remain chemically complete. A key experimental quantity in this model is the oxidation state of copper atoms in the  $\text{CuO}_2$  planes. Any deviation from 2.0 indicates that charge transfer is occurring. Because of the mild geometric distortion in the unit cell, the copper-oxygen bonds are stretched slightly in this plane. By measuring the bond lengths around that copper atom, a bond valence sum can be calculated, and this gives the oxidation state (43). As the number of oxygen vacancies varies, both this sum and the superconducting transition temperature  $T_c$  vary in the pattern shown in Fig. 14. This

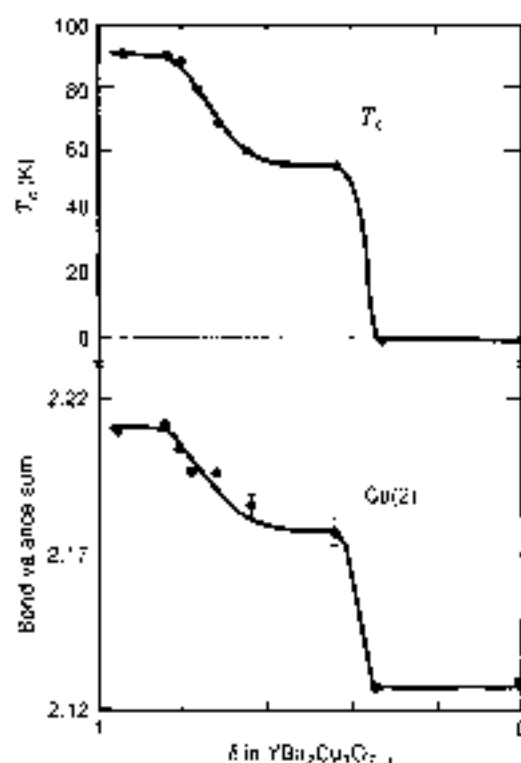


Figure 14. (a)  $T_c$  versus oxygen content in YBCO. (b) Bond valence sum around the copper atom in  $\text{CuO}_2$  layer. The purpose of showing the two figures together is to illustrate that  $T_c$  is proportional to the oxidation state of the copper atoms in the conductor layer.

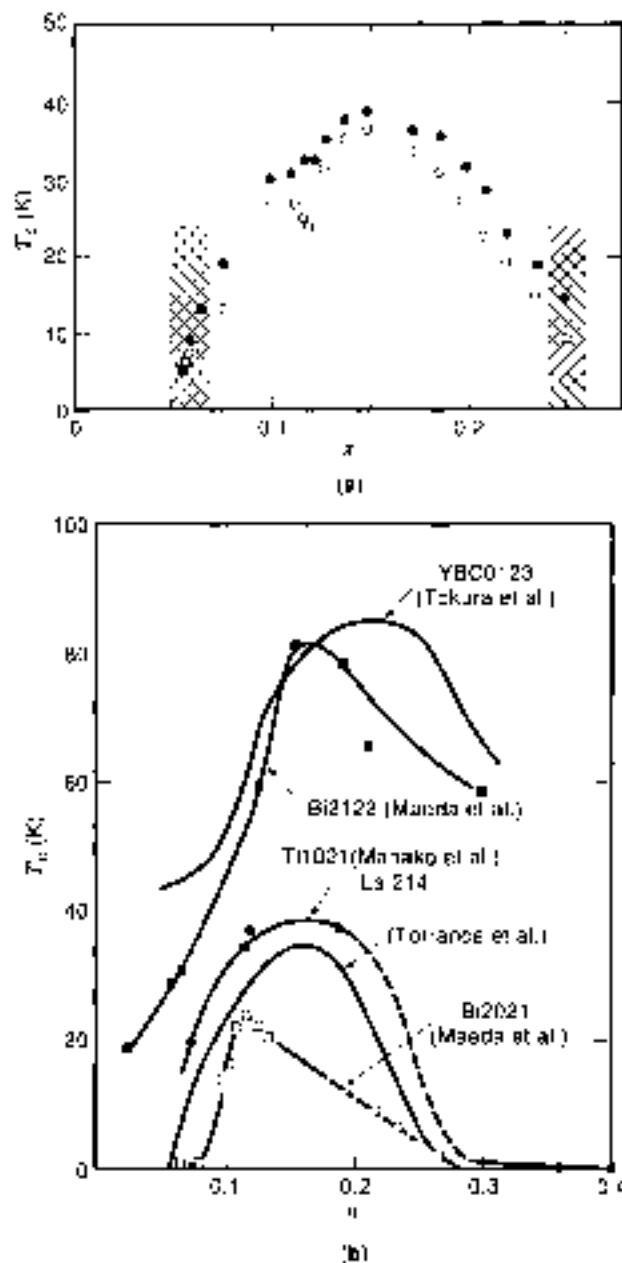
presents a very convincing argument for the charge-transfer model.

In TBCCO, there is competition between several different means of causing charge transfer. Those means may be missing metal atoms, extra metal atoms, missing oxygens, extra oxygens, etc. BSCCO exhibits all these conditions, plus the size mismatch between the normal dimensions of the  $\text{CuO}_2$  planes and the other planes in the stack is so large that atoms occasionally wind up in the wrong layer. The presence of many slightly different but similar structures in one crystal makes it difficult to interpret experimental data. Nevertheless, despite all the complexity, the average oxidation state for copper is 2.21, which tends to confirm the charge-transfer model.

It should also be noted that an excess of oxygen atoms also acts as a dopant, because it increases the number of sites where electrons can reside, which amounts to creating hole carriers in the crystal. The compound  $\text{La}_2\text{CuO}_4$  is normally an insulator, but when extra oxygen is forced in, changing it to  $\text{La}_2\text{CuO}_{4-x}$ , it becomes a superconductor. Thus, it appears that a small number of oxygen vacancies offer a way to make small changes in the carrier concentration. Oxygen vacancies are equivalent to the substitution of metal atoms in terms of their effect on the availability of charge carriers.

#### Role of Substituents in $\text{La-(Ba,Sr,Ca)-Cu-O}$ and $\text{Nd-Ce-Cu-O}$ Compounds

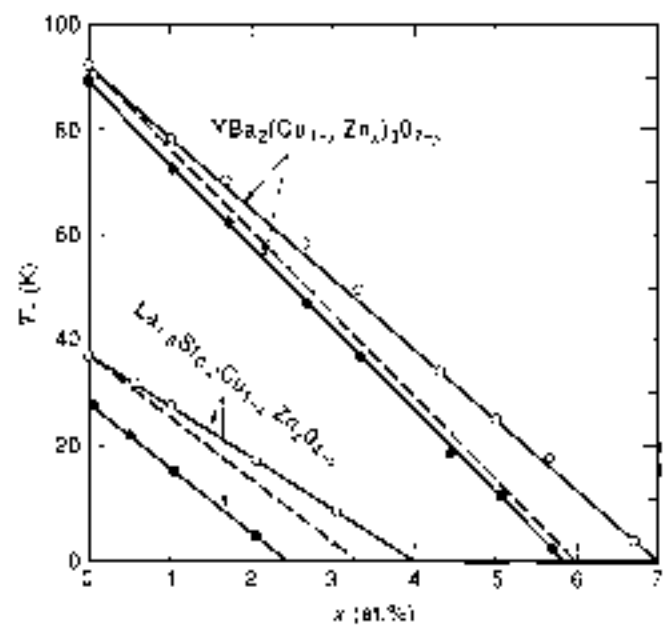
Substituting for  $\text{La}^{3+}$  by divalent ions  $M = \text{Ba}^{2+}, \text{Sr}^{2+}, \text{Ca}^{2+}$  leads to the appearance of metallic conductivity and for  $0.05 < x < 0.3$  superconductivity appears in  $\text{La}_{1-x}\text{M}_x\text{CuO}_{4-y}$  (LMCO) compounds.  $T_c/x$  is plotted in Fig. 15(a) where the black dots



**Figure 15.** (a)  $T_c$  as a function of Sr concentration in  $(\text{La}_{1-x}\text{Sr}_x)_2\text{CuO}_4$  (S.13), and the general dependence  $T_c(x)$  on the concentration of holes  $x$  in the  $\text{CuO}_2$  planes for some copper-oxide superconductors. (After Tokura, 1992.)

correspond to the appearance of the Meissner signal, and the white dots to its saturated value (44). For larger concentrations of Sr,  $x > 0.15$ , the oxygen vacancies can appear,  $y > 0$ ; their formation, however, can be suppressed by performing the annealing under high oxygen pressure. Metallic conductivity and superconductivity may also be obtained by increasing the content of oxygen above the stoichiometric value  $\text{La}_2\text{CuO}_{4+d}$ ,  $d > 0$ . In this case, oxygen occupies interstitial positions in layers of  $\text{La-O}$ . In general, for  $\text{Lu}_{2-x}\text{M}_x\text{CuO}_{4-d}$ , one observes the typical dependence  $T_c(x)$ , which is shown in Fig. 15(b), with  $x$  substituted by  $x - 2y$ , the number of free charge carriers (holes) per cell in the  $\text{CuO}_2$  planes.

Under the substitution of La by the trivalent rare-earth ions RE = Nd, Sm, or Gd, a smooth decrease of  $T_c$  with decreasing



**Figure 16.**  $T_c$  as a function of impurity  $x$  in a  $\text{La}_{1-x}\text{Sr}_{0.2}(\text{Cu}_{1-x}\text{Zn}_x)_2\text{O}_{4-y}$  compound. For comparison, the variation of  $T_c$  in  $\text{YBa}_2\text{Cu}_{1-x}\text{Zn}_x\text{O}_{7-y}$  is also shown as a function of  $x$ .

ionic radius in the series of these ions (45) occurs. At the same time, the value of magnetic moment of RE ion has no effect on  $T_c$ , which indicates that the Cooper pairs in the  $\text{CuO}_2$  plane are weakly coupled to the magnetic moments of ions in the La-O layers. An analogous situation is observed for electronic superconductors  $\text{Nd}_{2-x}\text{Ce}_x\text{CuO}_4$  under the substitution of Nd by the rare-earth ions Pr, Sm, and Eu. The decrease of  $T_c$  with decreasing radii of RE ions and primitive cell volume may be related to the decrease of Cu-O bond length.

However, this simple picture contradicts the dependence of  $T_c$  on the external pressure. In the La-Sr compounds  $T_c$  increases with an increase of pressure, attaining a certain maximum, and even decreases at high pressures. Meanwhile, in Nd-Ce compounds  $T_c$  is pressure independent. This difference in the  $T_c/p$  dependence can be related to a special role of apex oxygen: In the tetragonal phase of La-Sr compounds there are two apex oxygens in the complete  $\text{CuO}_2$  octahedron, and in the tetragonal phase of Nd-Ce compounds the apex oxygen is absent.

An isovalence substitution of copper by the ions of  $3d$  metals has a much stronger effect on  $T_c$ . At a concentration of  $x = 5\%$  to  $7\%$  for Ni and Fe and  $x = 2\%$  to  $3\%$  for Zn ions, the superconductivity in LSCO disappears (45–47). While the disruption of superconductivity due to scattering of the Cooper pairs on magnetic impurities is well known, the suppression of  $T_c$  due to Zn impurities in  $\text{CuO}_2$  planes is characteristic of copper-oxide superconductors. Figure 16 shows the  $T_c(x)$  dependence for the Zn impurity in  $\text{La}_{1-x}\text{Sr}_{0.2}(\text{Cu}_{1-x}\text{Zn}_x)_2\text{O}_{4-y}$  (LSCO) and YBCO (47). It is possible that Zn ions, which have a filled  $3d$  shell ( $3d^{10}$ ), destroy a rather complicated correlation conduction band in  $\text{CuO}_2$  planes that is related to Cu-O charge transfer as shown in Fig. 17 (48). This leads to suppression of superconductivity. In this context, one can consider  $\text{Zn}^{2+}$  ions as effective magnetic scatterers, because they substitute  $\text{Cu}^{2+}$  ions, which have a local magnetic moment.



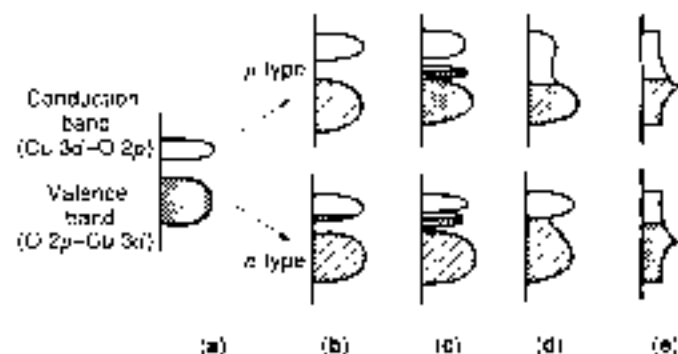


Figure 17. Schematic of electronic structure of cuprate superconductors accounting for (a) the correlation splitting of a  $pdx$  band, and some models for its change upon  $p$  or  $n$  doping; (b) rigid band; (c) impurity band; (d) filled gap; and (e) local density approximation. [After Pink et al. (48).]

### Role of Substituents in YBCO Compounds

A large number of studies have been carried out to investigate the role of constituents in YBCO compounds. Immediately after superconductivity with  $T_c = 90$  K in the YBCO system was observed, a large class of RBCO compounds with a similar  $T_c$  was prepared. This was done by substituting Y by lanthanides  $\text{Ln} = \text{La, Nd, Sm, Eu, Gd, Dy, Ho, Er, Tm, Yb, or Lu}$ . For these no separation of  $T_c$  in the compounds with rare-earth ions having a large magnetic moment has been observed. This indicates their weak coupling with the in-plane holes. The exceptions are Ce and Er ions, whose formal valence is +4, which violates the neutrality of the substitution of Y ions and can change the concentration of carriers in  $\text{CuO}_2$  planes. In order to investigate this concept, the system  $(\text{Y}_{1-x}\text{Ce}_x)_2\text{BaCu}_3\text{O}_{7-\delta}$  has been examined (49). The  $T_c(x)$  dependence for different  $x$  and  $T_c(x, y)$  at  $y = 0$  is shown in Figs. 18(a) and 18(b), respectively. The substitution of  $\text{Pr}^{3+}$  for Y decreases the number of in-plane holes, whereas substitution of  $\text{Ca}^{2+}$  has the opposite effect. This allows one to investigate separately the dependence of  $T_c$  on the concentration of holes  $n$  ( $y-x$ ) and on the magnetic scattering on the localized moments of Pr. Indeed, it follows from Fig. 18(a) that  $T_c(y)$  has a typical hole dependence with a maximum at an optimal concentration of holes  $n$  for various concentrations  $x$  of Pr ions. The location of these maxima shifts to larger values of  $y$  with an increase of Pr content, which proves directly that the concentration of carriers  $n$  ( $y-x$ ) decreases with an increase of  $x$ . At the same time, a decrease of the maximum value of  $T_c(x, y)$  for the optimal  $n$  is observed, which indicates a decrease of  $T_c$  due to magnetic scattering leading to the destruction of Cooper pairs. Figure 18(b) shows the effect of a change of the number of carriers on  $T_c(x, y = 0)$ . One can see that the maximum value of  $T_c$  in YBCO at  $T = 90$  K is attained under a small decrease of the number of in-plane holes due to the tetravalent impurities. In the course of investigations of the effect of impurity substitution on  $T_c$ , one must control the content of oxygen as it influences the number of carriers in  $\text{CuO}_2$  planes and determines the value of  $T_c$ .

An investigation of the effect of substitution of Ba in YBCO by the rare-earth ions has been carried out for the  $\text{La}_2\text{Ba}_{2-x}\text{Ln}_x\text{Cu}_3\text{O}_7$  compounds, where  $\text{Ln} = \text{La, Nd, Sm, Eu, Gd}$ . All the lanthanides produce an equal decrease in  $T_c$  with respect to the concentration of impurities  $x$ . This indicates a weak sensitivity of the superconducting transition to the appearance

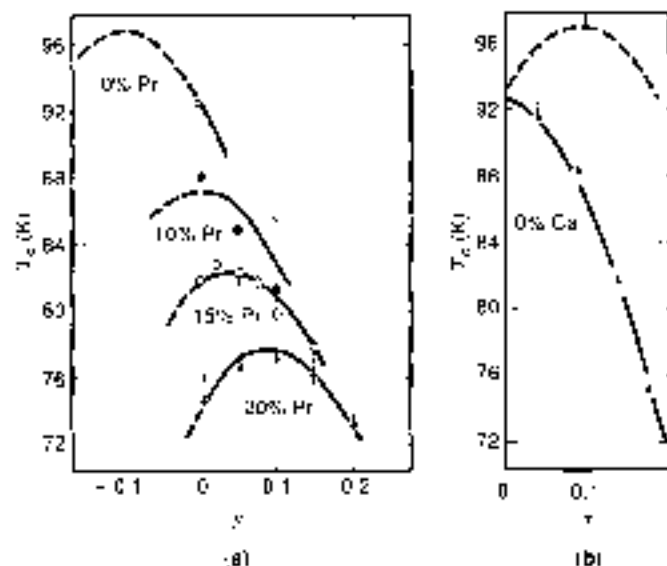


Figure 18. The dependence of  $T_c$  on  $\text{Y}_{1-x}\text{Ce}_x\text{Pr}_x\text{Ba}_2\text{Cu}_3\text{O}_{7-\delta}$  (a) on the concentration of Cu at a fixed content of Pr and (b) on the concentration of Pr for  $y = 0$ . The dotted line is the  $T_c$  dependence on the concentration of holes without magnetic scattering. [After Neumeier et al. (49).]

of magnetic moments on the Ln sites. A phase transition is also observed from orthorhombic to tetragonal (at  $x = 0.2$  to  $0.3$ ), which, however, did not have significant effect on the superconducting properties. In view of an often uncontrolled increase of oxygen content  $\delta = 4 - 7$  under the substitution of Ln for Ba and a complicated rearrangement of charge in the layers  $\text{Ba-O}(1)$ ,  $\text{Cu}(1-\text{O})$ ,  $\text{Cu}(2-\text{O})(2,3)$ , it is not possible to unambiguously conclude regarding the dependence of  $T_c$  on the concentration of holes in  $\text{CuO}_2$  planes.

In the YBCO compounds a more complicated effect of the substitution for copper ions has been observed. The complicated effect arises due to two reasons. First, in YBCO compounds there are two nonequivalent copper positions, whose substitution has a different effect on their electronic structure and superconductivity. Second, some impurities such as Fe and Co affect the oxygen content and the short-range order in  $\text{Cu}(1-\text{O})(1)$  layer, which may change the number of carriers in the  $\text{Cu}(2-\text{O})(2,3)$  planes.

It is interesting to study the effect of substituting Cu by ions of 3d elements  $\text{M}^{2+} = \text{Ti}(3d^2)$ ,  $\text{Cr}(3d^4)$ ,  $\text{Mn}(3d^5)$ ,  $\text{Fe}(3d^6)$ ,  $\text{Co}(3d^7)$ ,  $\text{Ni}(3d^8)$ , and  $\text{Zn}(3d^{10})$ . It has been observed that for  $x = 0.1$  in  $\text{YBa}_2\text{Cu}_{3-x}\text{M}_{0.1}\text{O}_7$ , the strongest suppression of  $T_c$  occurs for the ions of Fe and Co, having the maximum magnetic moment, and also for the Zn ions (50). However, some investigations have demonstrated that in case of Fe and Co an extra annealing in an oxygen atmosphere restores the value of  $T_c$  for  $x$  less than or equal to 0.35 almost completely, but does not relax the decrease of  $T_c$  by the impurities of Zn and Ni (51). Such different effects of annealing on  $T_c$  in impurity compounds of YBCO is explained by the fact that at low concentrations the ions of Co and Fe preferentially occupy the position of  $\text{Cu}(1)$ , and the ions of Zn and Ni the position of  $\text{Cu}(2)$ . Therefore, annealing in oxygen restores the oxygen content and its coordination in the  $\text{Cu}(1-\text{O})(1)$  layer for the Cu and Fe impurities, but has no effect on the plane  $\text{Cu}(2-\text{O})(2,3)$  for the Zn and Ni impurities.

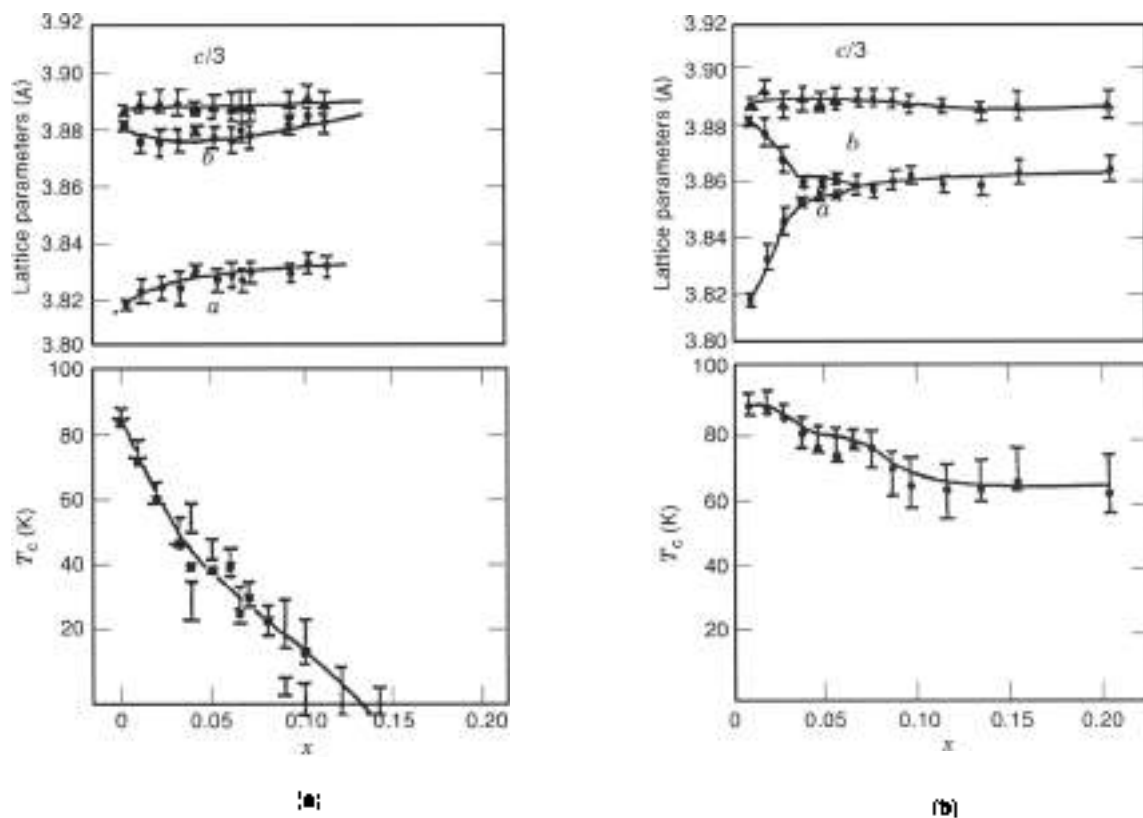


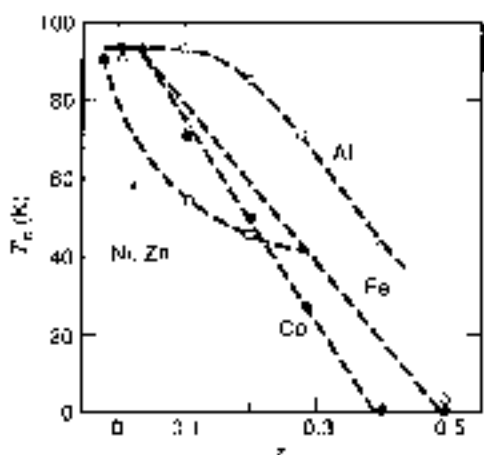
Figure 19. The dependence of lattice parameters and  $T_c$  on the concentration of impurities in  $\text{YBa}_2\text{Cu}_3\text{M}_2\text{O}_7$  for (a) Zn, and (b) Ga.

A more detailed analysis of the effect of diamagnetic impurities under a substitution for copper is also available in literature (52). In that study  $\text{Zn}^{2+}(\text{3d}^9)$  and  $\text{Ga}^{3+}(\text{3d}^0)$  impurities were used. Their ionic radii for a filled  $3d$  shell,  $R(\text{Zn}^{2+}) = 0.75 \text{ \AA}$ ,  $R(\text{Ga}^{3+}) = 0.62 \text{ \AA}$  are close to  $R(\text{Cu}^{2+}) = 0.73 \text{ \AA}$ , which allows one to obtain single-phase samples in a wide range of impurity concentrations. Neutron diffraction measurements have indicated that Zn primarily occupies the position of  $\text{Cu}(2)$ , and Ga the chain positions of  $\text{Cu}(1)$ . The dependence of  $T_c$  and lattice constants are shown in Fig. 19(a) for Zn and Fig. 19(b) for Ga. Zn impurities in the layers  $\text{Cu}(2)$  (0,2,3) have only a small effect on the parameters of the lattice, preserving the orthorhombic phase, but lead to a rapid suppression of  $T_c$ . The oxygen content remains close to optimal (6.8 and 7.0 for samples with Zn and Ga, respectively), and therefore the transition to the tetragonal phase in the case of Ga is not related to the deficiency of oxygen. The resistance  $\rho(x)$  in the normal phase increases much faster for Ga impurities than that for Zn, while they have the converse  $T_c(x)$  dependence. These experiments clearly demonstrate that the main role in the appearance of superconductivity in copper-oxide compounds is played by the  $\text{CuO}_2$  planes, where the specific properties of copper in  $3d^9$  states are necessary to attain high  $T_c$ .

According to studies of structural, magnetic and superconducting properties of  $\text{YBa}_2\text{Cu}_3\text{M}_2\text{O}_7$  compounds, paramagnetic and diamagnetic impurities affect these properties differently (49). Here  $\text{M} = \text{Ni, Fe, Co}$  represent paramagnetic impurities and  $\text{M} = \text{Zn, Al}$  represent diamagnetic impurities. It has been observed that the divalent impurities  $\text{Zn}^{2+}$  and  $\text{Ni}^{2+}$  do

not change the content in the sample and, when occupying positions in  $\text{CuO}_2$  planes, preserve the orthorhombic phase in the domain of single-phase states (for  $x < 0.8$  or the concentration  $x = z/8 < 0.1$ ).  $\text{Al}^{3+}$  impurities as well as Cu and Fe being in the trivalent state occupy the chain  $\text{Cu}(1)$  positions. They lead to an increase of oxygen content, filling the vacancies near the impurity site and leading to a transition to the tetragonal phase for  $x < 0.1$ . The dependence of  $T_c(x)$  for these impurity substitutions is shown in Fig. 20. It is clear from this figure that substituting in-plane  $\text{Cu}(2)$  by Zn impurities leads to a rapid decrease of  $T_c$ , while substituting the  $a$ -chain  $\text{Cu}(1)$  by trivalent impurities Al, Fe, Co has a much weaker effect on  $T_c$ . Apparently, both substitution of impurities for in-chain  $\text{Cu}(1)$  and changing the oxygen content have the same effect on the charge transfer from the chains to the planes, causing the transition from insulating to metallic phase which becomes superconducting. Note that in a series of experiments a more complicated dependence of  $T_c$  on the concentration of impurities that substitute for copper in YBCO compounds has also been observed (54, 55).

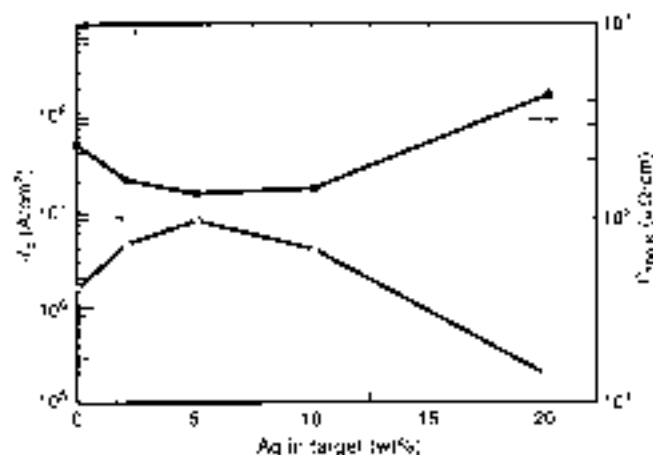
There are certain elements in the periodic table that have been shown to be unreactive with YBCO materials (56, 57). When such elements are added to YBCO compound, they do not substitute any of the elements in the lattice and tend to reside in intergranular regions (58). However, they have been found to affect significantly the superconducting properties of the YBCO material by modifying the grain-boundary properties. Silver is one such element that has most widely been used as a dopant to modify the grain-boundary properties of YBCO system (59). Shown in Fig. 21 is the variation of  $T_c$  and  $J_c$



**Figure 20.**  $T_c$  versus the concentration of impurities  $M = Ni^{2+}$  and  $Zn^{2+}$  and  $M = Al^{3+}$ ,  $Fe^{3+}$  and  $Co^{3+}$  in  $YBa_2Cu_3M_{1-x}O_{7-y}$ . [After Taraska et al. (55).]

of YBCO films doped with different amount of silver. All the films were deposited in situ on (100)  $LaAlO_3$  substrates using a pulsed laser deposition techniques. Different levels of doping in YBCO films were obtained by using targets containing different concentration of silver metal. Figure 21 shows that there is a one-to-one correlation between the resistivity and critical current densities. This correlation is due to weak-link coupling in which the critical current density of weak links should increase with lowered resistivity (60). Figure 21 also shows that films made using targets having 5% of silver gives highest  $J_c$  and the lowest room-temperature resistivity. This can be explained on the basis of the fact that a lower percentage of silver in the target may not be able to provide sufficient silver atoms required to enhance grain growth and liberate oxygen at the periphery of the grains of films, whereas a higher percentage of silver in the target would result in higher doping of YBCO films. Higher levels of silver in the targets and films may give rise to the formation of some impurity phases. According to Kao et al. (61), it is also possible that beyond the optimum Ag content, Ag atoms may enter the  $CuO$  basal planes in the orthorhombic structure and cause deterioration in  $T_c$  and  $J_c$  of the sample.

Since the grains in the superconducting films are coupled to each other through the materials in the grain boundaries,



**Figure 21.** Variation of room-temperature resistivity and critical current density ( $J_c$ ) at 77 K of films obtained from YBCO targets having different dopant levels of Ag. [After Kumar et al. (58).]

the presence of silver in the grain boundaries affects the properties of YBCO materials very significantly. The materials in the boundaries are the rejects of the grains, and they may be either metallic or insulating in nature. Depending upon the nature of the material in the grain boundary, the couplings are known as S-I-S (60) or S-N-S (62, 63) where S, I, and N stand for superconductor, insulator, and normal metal, respectively. The temperature dependence of the critical current density as a function of temperature gives an idea of the nature of the weak links present in superconducting samples with different amount of grain-boundary dopants (60). If the superconducting grains are coupled to each other by an insulating phase in the grain-boundary region, the current density of such junctions is determined by the expression suggested by Ambegaokar and Baratoff (60),

$$J_c(T) = n \Delta(T) / 2e R_j \tanh[\text{erfc}(\Delta(T) / 2k_B T)] \quad (1)$$

where  $R_j$  is junction resistance and  $\Delta(T)$  is energy gap. When  $T$  is close to  $T_c$ ,

$$J_c(T) \propto (1 - T/T_c) \quad (2)$$

If the suppression of the order parameter is taken into account, calculation through Ginzberg-Landau theory gives the following  $J_c(T)$  relation for an S-I-S network (64):

$$J_c(T) \propto (1 - T/T_c)^2 \quad (3)$$

However, if the superconducting grains are coupled to each other by normal-metal type material in the grain-boundary region, the junctions are known as S-N-S junctions and the current density of such junctions is given by the expression suggested by de Gennes (62) and Clarke (63),

$$J_c(T) = J_c(1 - T/T_c)^2 e^{-d/\xi_n} \quad (4)$$

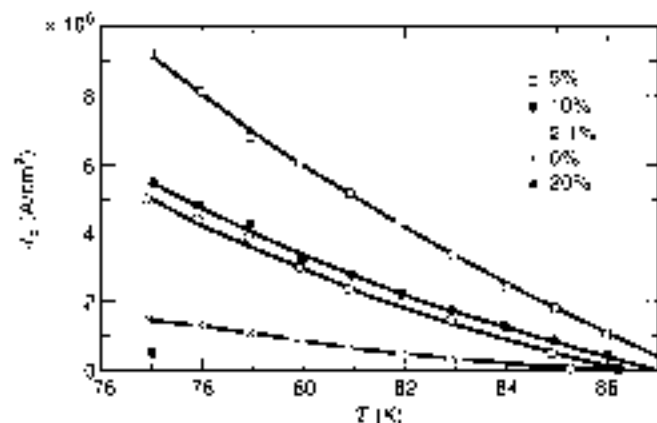
where  $d$  is the thickness of the grain boundary layer and  $\xi_n$  is the normal-metal coherence length. Ignoring the weak temperature dependence of  $J_c$  compared to the  $(1 - T/T_c)^2$  term, we obtain

$$J_c(T) \propto (1 - T/T_c)^2 e^{-d/\xi_n} \quad (5)$$

or

$$\sqrt{J_c(T)} \propto J_c(1 - T/T_c) e^{-d/2\xi_n} \quad (6)$$

The value of the slope of the  $\sqrt{J_c}$  versus  $T_c - T$  plot can be used as a figure of merit of the film, since the larger the slope of the plot the better the quality of the film. Having identified the nature of the weak link from the  $\sqrt{J_c}$  versus  $T_c - T$  plot, one can derive other useful information from this plot regarding the width of the grain boundary in different films. From Eq. (6), it is obvious that the slope of the  $\sqrt{J_c}$  versus  $T_c - T$  plot is proportional to the inverse exponent of the width of the grain boundary  $d$ . This concept was used by Kumar et al. (65) to understand the reason for the difference in critical current densities of undoped and Ag-doped films. They fabricated a series of YBCO films with different concentration of



**Figure 22.** Critical current density versus temperature plots for undoped and Ag-doped YBCO films on  $\text{LuAlO}_3$  substrates grown under identical conditions. [After Kumar et al. (1994).]

Ag. Plots of the temperature dependence of the critical current density of these films are shown in Fig. 22. Determining the values of  $\sqrt{J_c}$  and  $T_c - T$  from Fig. 22, the  $\sqrt{J_c}$  versus  $T_c - T$  plots of undoped and Ag-doped YBCO films were obtained and are shown in Fig. 23. It is clear from this figure that both undoped and Ag-doped YBCO films follow the S-N-S model, but their grain-boundary widths are different. The variation in grain-boundary widths in undoped and Ag-doped films is as follows:  $d_{\text{undoped}}/d_{\text{Ag-doped}} = 1.0/0.6/0.5/0.65/2.0$ . Owing to differences in grain-boundary width, the grain-boundary resistance and hence the strength of S-N-S coupling in these films are also different. In the present example, the coupling is maximum in the case of 5 wt % Ag-doped film and is minimum in the case of 20 wt % Ag-doped film, and therefore the former has the maximum current density and the latter has the least.

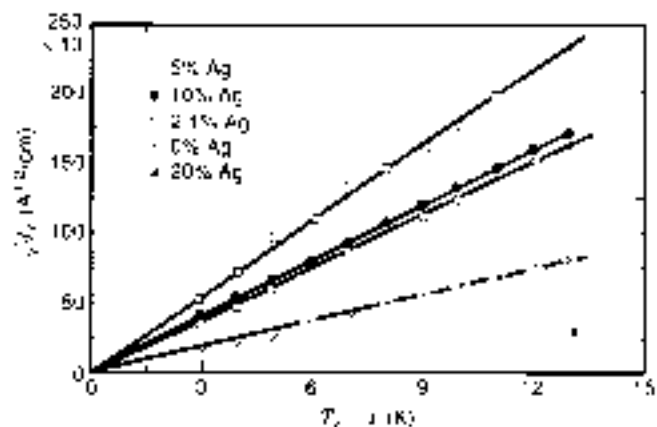
#### Role of Substituents in Bi-, Tl-, and Hg-Based Compounds

The copper superconductors containing layers of bismuth oxide, thallium oxide, or mercury oxide exhibit superconductivity at higher transition temperatures than YBCO. Because all three types of compounds have very similar structures, it is convenient to discuss the role of dopants in these compounds together. Among the bismuth-oxide-based superconductors, the

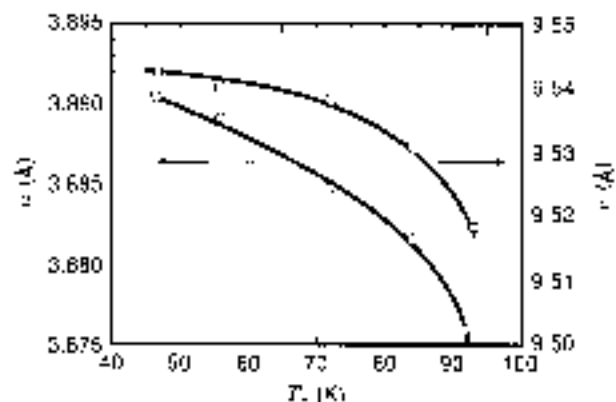
structures of greatest interest are  $\text{Bi}_2\text{Sr}_2\text{Ca}_1\text{Cu}_2\text{O}_{10}$  (BSCCO-1212) and  $\text{Bi}_2\text{Sr}_2\text{Ca}_2\text{Cu}_3\text{O}_{10}$  (BSCCO-2223). The most familiar substitution in bismuth-oxide-based superconductors is to replace bismuth with lead. This has relatively minor effects on the superconducting properties, but it strongly affects the kinetics and phase equilibrium of the mixture. The result is that  $(\text{Pb}, \text{Bi})_2\text{Sr}_2\text{Ca}_2\text{Cu}_3\text{O}_{10}$  is much easier to make than BSCCO-2223 without lead. In a similar analogy with the rare-earth substitutions in YBCO, the partial replacement of Bi with Pb makes little difference electronically, and thus changes the  $T_c$  only very modestly.

In both the thallium- and bismuth-based superconductors, the strategy of doping is driven by the conjectures of each investigator about what will produce desired properties. The number of substitutions possible on those unit cells, together with the complexity of their phase diagrams, requires verification of hypotheses by empirical means. The familiar series of thallium superconductors are made of Tl, Ba, Ca, Cu, and O. The transition temperatures of various phases are generally above 100 K. A less well known series is based on Tl-Sr-Ca-Cu-O (TSCCO) with Sr playing the equivalent role of Ba. TSCCO is nonsuperconducting (Sheng et al. 1986) have shown that the substitution of the trivalent Y for the divalent Ca brings an extra electron onto that layer in the unit cell, leading to the conversion of material to the superconducting phase with a  $T_c$  of 80 K. The goal of most substitutions is to change the number of carriers (holes, usually) by doping the various layers of the unit cell. As in YBCO compounds, the carrier concentration can be manipulated to optimize  $T_c$  by doping with elements of different valance. Using this concept a  $T_c$  of 110 K has been achieved in TSCCO by combining the lead-bismuth substitution for thallium with a substitution of yttrium for calcium (67, 68). The chemical formula is typically  $(\text{Tl}_{1-x}\text{M}_x)_2(\text{Sr}_2\text{Ca}_{1-y}\text{Y}_y)_2\text{Cu}_2\text{O}_{7+z}$ , with various choices of  $x$  and  $h$  between 0 and 1, and  $h$  around 0.1. The  $T_c$  of this series has been found to lie between 40 K and 100 K.

The first mercury-based HTSC is  $\text{HgBa}_2\text{CuO}_{4-x}$  (69). As in the case of  $\text{La}_2\text{CuO}_{4-x}$ , the primary doping mechanism is an interstitial oxygen. The maximum  $T_c = 95$  K when  $\delta$  is approximately equal to 0.06, but drops to 59 K when  $\delta = 0.01$ . The change in oxygen content changes the unit-cell  $a$  and  $c$  dimensions. The associated change in  $T_c$  correlates with these dimensions as shown in Fig. 24 (70). Experiments have shown that about 6% of copper occasionally substitutes on the mercury



**Figure 23.**  $\sqrt{J_c}$  versus  $T_c - T$  plots of undoped and Ag-doped YBCO films on  $\text{LuAlO}_3$  substrates grown under identical conditions. [After Kumar et al. 65.]



**Figure 24.** Lattice parameters  $a$  and  $c$  versus  $T_c$  for  $\text{HgBa}_2\text{CuO}_{4-x}$  samples with different oxygen contents. [After Wagner et al. (70).]

site, but a systematic pattern of changes in  $T_c$  is yet to be established. The new azate compounds with two or three  $\text{CuO}_x$  layers that have  $T_c \sim 133$  K also contain a slight excess of oxygen (6). It is expected that the effect of oxygen interstitial and vacancies may be similar to that found in previous HTSCs.

## CONCLUSION

This article deals with the structure and role of constituents in high-temperature superconductors. After introducing certain basic terminology, we have presented schematic drawings of the unit cells of several different high-temperature superconductors. The high-temperature superconductors have in common the presence of copper-oxide layers, with superconductivity taking place between these layers. The unit cells are not perfectly symmetric, which has important consequences for superconductivity. For the role of substituents, the range of possible substitution in high- $T_c$  superconductors is enormous. This article presented some examples with an effort to explain why researchers preferentially try particular kinds of doping. The motivation is to manipulate the interacting electrons so as to learn more about the metallic state of the ceramic oxides. The goal of doping research is not to raise  $T_c$  through empiricism but to establish an experimental foundation on which a theory can be built to explain the mechanism of high- $T_c$  superconductors.

## BIBLIOGRAPHY

- H. K. Onnes, Report on the research made in the Leiden Cryogenic Laboratory between the 2nd and 3rd Int. Congr. on Refrigeration, Suppl. No. 34b, 1919.
- J. G. Bednorz and K. A. Müller, *Z. Phys. B*, **64**: 183, 1986.
- X. D. Wu et al., *Appl. Phys. Lett.*, **51**: 861, 1987.
- H. Muedo et al., *Jpn. J. Appl. Phys.*, **27**: L2, 1988.
- Z. Z. Sheng and A. M. Hermann, *Nature*, **332**: 138, 1988.
- A. Schilling et al., *Nature*, **303**: 56, 1983.
- M. Nunez-Regueiro, *Science*, **262**: 97, 1993.
- C. W. Chu, *Nature*, **365**: 323, 1993.
- L. N. Cooper, *Phys. Rev.*, **104**: 1139, 1956.
- F. Landon and H. London, *Proc. R. Soc. London, Ser. A*, **140**: 7, 1935.
- J. C. Phillips, in *Physics of High T. Superconductors*, New York: Academic Press, 1988.
- A. Wattiaux et al., *C. R. Seances Acad. Sci.*, **310**: 1047, 1980.
- C. W. Chu et al., *Phys. Rev. Lett.*, **58**: 406, 1987.
- J. Fonteberta, L. Fábrega, and A. V. Narlikar, ed., *Studies in High Temperature Superconductors*, vol. 16, New York: Nova Science, 1986.
- T. Jerome, W. Kang, and S. S. P. Parkin, *J. Appl. Phys.*, **63**: 4005, 1988.
- J. M. Tranquada et al., *Phys. Rev. B*, **38**: 8691, 1988.
- D. Kumar, Electrical, structural and chemical properties of laser sputtered YBCO thin film. Ph.D. thesis, Indian Institute of Technology, Bombay, 1993.
- E. J. Rarhman et al., *Proc. DIMETA 66, Int. Conf. Diffus. Met. Alloys*, Hungary, 1969.
- R. J. Cava, *Nature*, **338**: 328, 1989.
- T. Siegrist et al., *Nature*, **384**: 231, 1989.
- C. C. Torardi et al., *Science*, **240**: 631, 1988.
- P. Chaudhari et al., *Phys. Rev. Lett.*, **60**: 1653, 1988.
- D. Dimos, P. Chaudhari, and J. Mannhart, *Phys. Rev. B*, **41**: 4038, 1990.
- J. Narayan et al., *Appl. Phys. Lett.*, **61**: 1845, 1992.
- H. Dijkamp et al., *Appl. Phys. Lett.*, **51**: 619, 1987.
- D. Dimos et al., *Phys. Rev. Lett.*, **61**: 219, 1988.
- K. Jagannadham and J. Narayan, *Philos. Mag.*, **59**: 917, 1989.
- W. T. Brad and W. Shirkley, *Phys. Rev.*, **78**: 275, 1950.
- K. Jagannadham, J. Narayan, and H. S. Kwok, ed., *Superconductivity and Applications*, New York: AIP, pp. 37-49.
- M. E. Chisholm and S. J. Pennycook, *Nature*, **351**: 47, 1991.
- V. L. Ginsberg and I. D. London, *Zh. Eksp. Teor. Fiz.*, **20**: 1064, 1950.
- G. Deutscher, *Physica C*, **153-155**: 15, 1988.
- K. Jagannadham and J. Narayan, *Mater. Sci. Eng. B*, **8**: 5, 1991.
- N. D. Browning et al., *Physica C*, **212**: 195, 1993.
- S. McKernan, M. G. Norton, and C. H. Carter, *J. Mater. Res.*, **7**: 1052, 1992.
- D. M. Krieger et al., *J. Appl. Phys.*, **64**: 351, 1988.
- S. Goryunov, et al., *Proc. Materials Meeting*, **4**, 1988.
- R. Kalyanasrman et al., *Mater. Res. Soc. Proc.*, **528**: in press.
- M. M. Gillbon, *Philos. Mag. A*, **73**: 625, 1996.
- S. E. Dabcock et al., *Nature*, **347**: 167, 1992.
- C. B. Eam et al., *Science*, **261**: 780, 1981.
- J. D. Jorgensen et al., *Phys. Today*, **44**: 34, 1991.
- J. D. Jorgensen et al., *Phys. Rev. B*, **36**: 3608, 1987.
- H. Takagi et al., *Phys. Rev. B*, **40**: 2254, 1989.
- I. H. Greene, B. G. Hagle, and T. M. Gruber, ed., *Physical Properties of High-Temperature Superconductors*, vol. 2, Singapore: World Scientific, 1990, p. 509.
- J. T. Marquart, Y. Dalichouh, M. B. Maple, and D. M. Ginsberg, ed., *Physical Properties of High-Temperature Superconductors*, vol. 1, Singapore: World Scientific, 1989, p. 509.
- A. V. Narlikar, C. V. N. Rao, S. K. Agarwal, and A. V. Narlikar, ed., *Studies of High Temperature Superconductors*, vol. 1, New York: Nova Science, 1989, p. 341.
- J. Pink et al., *Proc. Int. Semia High T<sub>c</sub> Supercond.*, Dubna, JINR, E17 90 472, 1990, p. 3.
- J. J. Neumeier, et al., *Phys. Rev. Lett.*, **63**: 2516, 1989.
- G. Xiao et al., *Phys. Rev. B*, **85**: 4742.
- Y. Oda et al., *J. Phys. Soc. Jpn.*, **67**: 4079, 1998.
- G. Xiao et al., *Phys. Rev. Lett.*, **60**: 1446, 1988.
- J. M. Taraska et al., *Phys. Rev. B*, **37**: 7453, 1988.
- A. M. Briagarov, J. Piechora, and A. Pajczkowska, *Solid State Commun.*, **78**: 407, 1991.
- S. Katsuyama, Y. Ueda, and K. Kosuge, *Mater. Res. Bull.*, **24**: 603, 1989.
- William Chaudhary, 1988.
- Kobaschewski Monck, 1979.
- D. Kumar et al., *Appl. Lett.*, **62**: 3522, 1993.
- J. Joo et al., *Appl. Supercond.*, **2**: 401, 1994.
- V. Ambegaokar and A. Baratoff, *Phys. Rev. Lett.*, **10**: 486, 1963.
- Y. H. Kao et al., *Appl. Phys. Lett.*, **67**: 363, 1995.
- P. G. de Gennes, *Rev. Mod. Phys.*, **36**: 225, 1964.
- J. Clarke, *Proc. R. Soc. London, Ser. A*, **308**: 447, 1969.
- M. Tinkham, in *Introduction to Superconductivity*, New York: McGraw-Hill, 1975.
- D. Kumar, P. H. Apte, and H. Patel, *J. Appl. Phys.*, **77**: 5802, 1995.
- J. P. Sheng et al., *Appl. Phys. Lett.*, **54**: 280, 1989.
- B. S. Liu et al., *Appl. Phys. Lett.*, **67**: 2432, 1995.

68. R. S. Liu et al., *Physica C*, **159**: 335, 1989.  
 69. S. N. Putilin et al., *Nature*, **382**: 226, 1990.  
 70. J. L. Wagner et al., *Physica C*, **210**: 147, 1993.

### Reading List

- IEEE Trans Electron Dev*, **34**, 1987, special issue.  
 Y. Jeon, G. Liang, J. Chen, M. Craft, M. W. Ruckman, D. DiMarzio, and M. S. Hegde, (1990), *Phys Rev B*, **41**, 4066.  
 W. Jin, M. H. Dagam, R. K. Kalia, and P. Vashita, *Phys Rev B*, **46**, 5535.  
 D. Kumar, (1994) in *Electrical, Structural and Chemical Properties of Laser Ablated YBCO Thin Films*, Ph. D. Thesis (Indian Institute of Technology, Bombay, India).  
 G. Xiao, Crepla's, D. Muzer, M. Z. A. Gavzin, E. H. Streitz, A. Bakhshai, C. L. Chien, J. J. Rhyne, and J. A. Gutkas, (1986), *Nature*, **322**, 236.

D. KUMAR  
 University of Dayton  
 RAJIV K. SINGH  
 University of Florida  
 R. KALANJAMAN  
 Northrop Grumman Science and  
 Technology Center  
 S. OKTYABRSKIY  
 K. JAGANNADHAM  
 J. NARAYAN  
 North Carolina State University

## Nb-Ti ALLOY SUPERCONDUCTORS

The most widely used superconducting materials are based on Nb-Ti alloys with Ti contents ranging from 46 wt. % to 50 wt. % Ti. These alloys of Nb and Ti have both high strength and ductility and can be processed to achieve high critical current densities that make them ideal candidates for magnets and applications. Nb-Ti-based superconductors are commercially produced in long uniform lengths and cost significantly less to produce than other superconductors. The main drawbacks of this material are a low critical temperature, typically requiring cooling by liquid helium, and a low upper critical field which limits the applied field at which they can be used to below 12 T.

Although several other ductile superconductors have been investigated (most importantly niobium-zirconium), the niobium-titanium alloy system has been the only ductile superconductor in use since the mid-1960s. The vast majority of all superconducting magnets have been constructed from a small Nb-Ti alloy range of 46 wt. % Ti to 50 wt. % Ti (62 at. % Ti to 66 at. % Ti), with most of the superconductor used being 47 wt. % Ti (63 at. % Ti). It is the strength, ductility, critical current density, and relatively low cost of this material that makes it stand out as a commercial superconductor. The critical temperature,  $T_c$ , and upper critical field,  $H_{c2}$ , of these Nb-Ti alloys are unexceptional ( $T_c \sim 9.3$  K,  $H_{c2}$  (4.2 K)  $\sim 11$  T,  $H_{c2}$  (12 K)  $\sim 14$  T), but their very high critical current densities ( $\sim 3000$  A/mm<sup>2</sup> at 5 T, 4.2 K) make them the ideal choice for applications in magnetic fields up to 9 T at 4.2 K (liquid helium coolant) and 12 T at 1.9 K (superfluid liquid helium coolant). The primary applications of Nb-Ti-based magnets are magnetic resonance imaging (MRI), nuclear magnetic resonance (NMR), particle accelerators, mag-

netic confinement for fusion reactors, arc separators, electrical power conditioners, and superconducting magnetic energy storage (SMES). Replacing some Nb with 15 wt. % Ta to 25 wt. % Ta further extends the upper critical field when cooled with superfluid liquid helium but has yet to achieve significantly higher critical current densities. This article will concentrate on Nb-47 wt. % Ti and similar alloys because of their complete dominance as ductile superconductors. Callings (1,2) has compiled extensive reviews of the properties and applications of ductile Ti-based superconductors and has covered the history of their development up to 1983 in an earlier volume (3).

Almost all Nb-Ti strands are manufactured as a composite of Nb-Ti filaments in a Cu (or occasionally Cu-Ni or Al) matrix. The number of filaments can vary from 60 to 6000, and their diameters can vary from 5  $\mu$ m to 50  $\mu$ m. The Cu matrix typically represents 50% to 70% of the strand volume and provides the composite with thermal and electrical stability. Subdividing the filaments provides adiabatic flux-jump stability and allows greater cold work to be applied to the strand, which is beneficial in achieving high critical currents. Commercial scale composites are manufactured on a scale of  $\sim 200$  kg billets, and the resulting wire lengths can exceed 10 km at final size without breakage.

In type II superconductors, such as Nb-Ti, high critical current densities in magnetic fields are only possible if fluxoid motion is inhibited. In Nb-Ti, strong fluxoid pinning is made possible by the creation of a finely dispersed nanometer scale non-superconducting phase that closely matches the fluxoid spacing. The method of creating that pinning microstructure distinguishes the two basic methods of producing Nb-Ti strand. Figure 1 shows such a microstructure achieved by heat-treating the strand during processing to produce Ti-rich  $\alpha$ -Ti which fold into a densely packed array of sheets when the Nb-Ti wire is drawn to final size. The  $\alpha$ -Ti sheets are typically 1 nm to 4 nm thick with a separation of 5 nm to 20 nm. The folding of the microstructure during wire drawing is a result of the limited deformation orientations available in the body-centered cubic (BCC) crystal structure of the Nb-Ti grains. Almost all commercial Nb-Ti strands are fabricated using precipitation heat treatment and are termed "conventionally processed" in order to distinguish them from the second and newer method of Nb-Ti production called Artificial Pinning Center (APC). In the APC process the pinning microstructure is engineered by



**Figure 1.** A transmission electron microscope image of the microstructure of a Nb-47 wt. % Ti superconductor in transverse cross section reveals a densely folded array of second phase pins which are 1 nm to 4 nm in thickness. For comparison a schematic illustration of the fluxoid diameter (18 nm) and spacing (22 nm) at 5 T and 4.2 K is superimposed in the top left-hand corner.

mechanically assembling rods or sheets of the component materials at a size large. Extrusion and wire drawing reduces the assembled array to the final nanometer dimensions. Like the conventionally processed strand, the deformation of the engineered microstructure is dominated by the folding of the Nb-Ti matrix grains, and the resulting APC microstructures are very similar in appearance to those produced by conventional processing. The APC approach allows a large degree of freedom in the selection of matrix and pinning materials as well as their ratios and physical distribution. Consequently, APC strands outperform conventionally processed strands in fields up to 5 T and promise greater performance at higher fields. The additional cost associated with assembling the pinning array and reducing it to nanometer scale has limited its commercial application.

## THE Nb-Ti ALLOY SYSTEM

### Stable Phases

The atomic volume difference between Ti and Nb is only about 2%, resulting in a  $\beta$ -isomorphous system where the  $\beta$  phase has a BCC structure with a lattice parameter of approximately 0.3285 nm. Figure 2 shows the variation in  $H_{c2}$  at 4.2 K (4),  $T_c$  (5) and resistivity,  $\rho$ , (6) with composition for single-phase  $\beta$ -Nb-Ti. The upper critical field peaks sharply in the range of 40 wt. % Ti to 50 wt. % Ti with a maximum value of 11.5 T at 44 wt. % Ti. The critical temperature drops continuously over this range with increasing Ti content. The only other stable phase in this system is the Ti-rich  $\alpha$  phase which has a hexagonal close-packed (HCP) structure and a composition of 1 at. % Nb to 2 at. % Nb. The low Nb content of the  $\alpha$ -Ti phase suggests that  $\alpha$ -Ti precipitates should have a low  $T_c$ , approaching the 0.89 K  $T_c$  of pure Ti, and should be nonsuperconducting under practical operating conditions. The beneficial role of precipitating the normal phase  $\alpha$ -Ti precipitates was first shown by Pfeiffer and Hillman in 1968 (7). The alpha phase is only stable below 562 C (at atmospheric pressure); and for the alloy composition range of interest,  $\alpha$ -Ti is only stable below 570 to 600 C. In Figure 3 the widely used high-temperature phase boundaries of Hansen et al. (8) are combined with the calculated low tem-

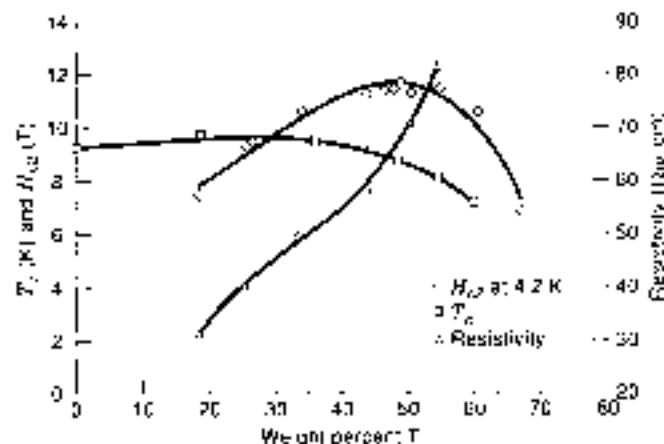


Figure 2. The variation in  $H_{c2}$  at 4.2 K (4),  $T_c$  (5) and resistivity (6) with composition for single-phase Nb-Ti.  $H_{c2}$  is defined as the linear extrapolation of the high-field pinning force ( $F_p$ ) to zero.

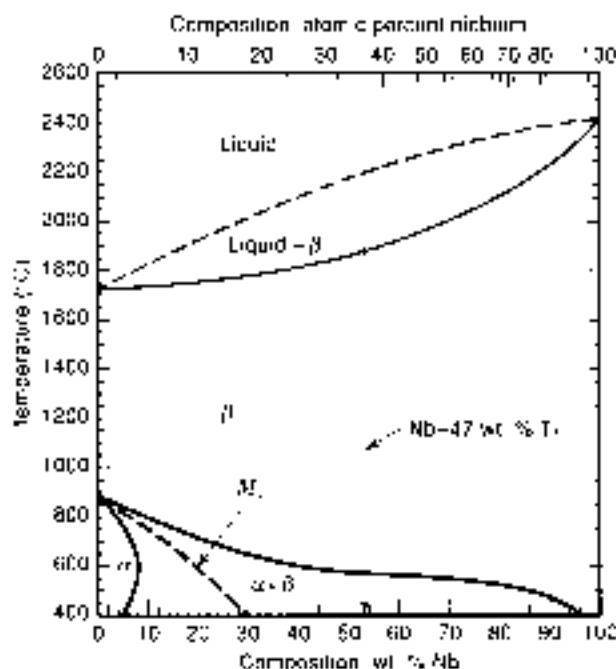


Figure 3. A hybrid equilibrium phase diagram for Nb-Ti combining the experimentally determined high-temperature phase boundaries of Hansen et al. (8), with the calculated low-temperature phase boundaries of Kaufman and Bernstein (9) modified by Moffat and Kattner (10). Also shown is the martensite transformation curve (M) of Moffat and Larbalestier (11).

perature boundaries of Kaufman and Bernstein (9) modified by Moffat and Kattner (10) to provide a composite equilibrium phase diagram that generally reflects production experience. The use of the calculated low-temperature phase boundaries is a result of the difficulty of achieving equilibrium at the low temperatures compared to the melting point at which the  $\beta$  to  $\alpha + \beta$  transformation occurs in Nb-Ti. The interdiffusion rate in  $\beta$ -Nb-Ti decreases exponentially with decreasing Ti content, and it was calculated in Ref. 10 that at 500 C it would take 10 years to reach the same condition in a 30 at. % Ti alloy that it would take in 3 s in an 80 at. % Ti alloy. The slow diffusion rates mean that in the composition range for superconducting application the single  $\beta$  phase can be retained even with relatively slow quenching from above  $\beta/\alpha + \beta$  boundary (600 to 650 C). The shape and position of the  $\beta/\alpha + \beta$  boundary is important because it determines the maximum volume of  $\alpha$ -Ti precipitate that can be formed for a given heat treatment temperature and alloy composition. The calculated boundary indicates that both increasing the Ti content and decreasing the heat treatment temperature should increase the maximum volume of precipitate that can be produced.

### Metastable Phases

There are three metastable phases of importance, two martensite ( $\alpha'$  and  $\alpha''$ ) and an  $\alpha$  phase. The  $\alpha'$  martensite is HCP with lattice parameters identical to  $\alpha$ -Ti, and it forms in alloys up to 7 at. % niobium. The orthorhombic  $\alpha''$  is transitional between the HCP  $\alpha'$  and the BCC  $\beta$  phase, and it forms at higher Nb concentrations. The martensite transformation boundary of Moffat and Larbalestier (11) is shown in Fig. 3 and shows that the

most commonly used Nb-Ti alloys are outside the range of the martensite transformation. The  $\alpha$  phase has a hexagonal crystal structure ( $c/a = 0.613$ ). It can be formed athermally in the alloy range 86 at. % Ti to 70 at. % Ti by quenching from the  $\beta$ -phase region, or it can be formed by aging in the temperature range of 100 to 500 C. The  $\omega$  phase is typically observed as small ellipsoids roughly 5 nm to 10 nm in their longest dimension. In cold-worked and heat-treated Nb-Ti strands, they can grow to 50 nm in diameter. All the metastable phases can be transformed to single-phase  $\beta$ -Nb-Ti or two-phase  $\alpha + \beta$  microstructures by heating long enough in the  $\beta$  or  $\alpha + \beta$  phase fields, respectively.

#### Cold-Worked Microstructures

In order to achieve high critical current densities a fine and homogeneous dispersion of flux pinning material must be introduced that is of sufficient volume for significant pinning but does not deleteriously affect the other  $H_{c2}$  or  $T_c$ . The process by which the first high critical current density microstructures were achieved was arrived at empirically before the resulting microstructures were characterized (12). The processing involved a high cold work strain followed by three or more heat treatments in the  $\alpha - \beta$  phase range, each separated by additional cold work with the final heat treatment being followed by another large cold-work strain. An understanding of the microstructural development was key, however, to the further optimization of Nb-Ti and the reproducible production of high critical current strand. Initial observation of the microstructure was hindered by the difficulty in preparing transverse cross sections of micron sized filaments suitable for examination by transmission electron microscopy (TEM). Once techniques had been developed to prepare the TEM specimens, it became clear that folded sheets of  $\alpha$ -Ti precipitates were the dominant microstructural features of the final strand (see Ref. 13). Systematic analysis of the production process (as in Ref. 14) revealed that the high prestrain heat treatments produced  $\alpha$ -Ti precipitates only at the intersections of grain boundaries. The location of precipitation at the grain boundary triple points meant that the precipitation was homogeneously distributed if alloy composition and grain size were uniform. The grain boundary triple-point  $\alpha$ -Ti was also sufficiently ductile that it could be drawn down to the nanometer scale with breaking up or causing the strand itself to become difficult to draw. This contrasted with the other commonly observed  $\alpha$ -Ti precipitate morphology, Widmanstätten  $\alpha$ -Ti, which formed in densely packed rafts in the interior of grains and resulted in a great increase in the filament hardness. The next section reviews each step of the process in more detail.

### THE CONVENTIONAL PROCESS

#### Alloy Fabrication

A high-purity fine-grained Nb-Ti alloy with chemical homogeneity over both a large and small microstructural scale is an essential starting point to the production of Nb-Ti strands. The large liquid-solid phase separation shown in the phase diagram (Fig. 3), along with the high melting point of the Nb, makes it particularly difficult and expensive to produce a high quality Nb-Ti alloy suitable for superconductor application. The main driving force for high homogeneity is the key role

that precipitate quantity and morphology play in determining critical current density, both of which are highly sensitive to composition. The development of a high-homogeneity Nb-Ti alloy was a crucial step in the advance toward high critical current Nb-Ti (see Ref. 15). The desired properties of the initial alloy billet are as follows.

1. The correct overall alloy composition to optimize  $H_{c2}$ ,  $T_c$ , and precipitation for pinning. The acceptable range is Nb-46 wt. % Ti to Nb-48 wt. % Ti.
2. Uniform composition over the entire billet to ensure optimum physical and mechanical properties over the entire filament.
3. Chemical homogeneity on a microstructural level in order to ensure uniform precipitation of the correct morphology (typically  $\pm 1.5$  wt. % Ti).
4. Low and controlled levels of impurity elements in order to ensure predictable superconducting and mechanical properties.
5. Elimination of hard particles (typically Nb-rich) because any particle that does not co-reduce with the alloy can result in filament drawing instability and ultimately strand breakage. The exterior of the final Nb-Ti rod must also be free of hard particles and must be smooth enough that it does not easily pick up particles during subsequent handling.
6. A fine (typically ASTM grain size 6 or smaller) and uniform grain size as it controls the distribution of precipitate nucleation sites. A fine grain size also improves diffusion barrier uniformity. Where high critical current is less important, a larger grain size has been used increase ductility.
7. Low hardness (typically a Vickers hardness number of 170 or less) to ease re-deformation with softer stabilizer material.

The Nb-Ti alloy is prepared from high-purity Nb and Ti by consumable electrode vacuum-arc melting (where the electrodes are composites of Nb and Ti) and by electron-beam or plasma-arc melting. It is usually necessary to remelt the ingot two or three times in order to achieve the necessary chemical homogeneity. Primarily produced for the aviation industry, the high-purity source Ti is reduced from  $TiCl_4$  by Mg (the Kroll process). High-purity Nb is refined from lower-purity Nb by two or three electron-beam remelts. The lower-purity Nb source itself is extracted from niobite-tantalite ( $Nb_2O_5$  and  $Ta_2O_5$ ) or pyrochlore (0.25% to 3%  $Nb_2O_5$ ) ores by way of an intermediate ferroniobium alloy which is used on a relatively large scale for steel production. Table 1 lists the typical allowable ranges for impurities, typified by specifications for the superconducting supercollider. The small level of allowable Ta has a historical origin, and it is unlikely that additions of less than 1.0 wt. % Ta will have a significant impact on superconducting or mechanical properties. Increasing the level of Fe from 200  $\mu\text{L/L}$  (the specification of the superconducting supercollider as found in Ref. 16) to 500  $\mu\text{L/L}$  ( $\mu\text{L/L}$ ) is equivalent to the more commonly used ppm; actually has a beneficial effect as shown in Ref. 17.

The chemical inhomogeneities that may be observed in the alloy at this stage in production can be divided into two types based on size: macroinhomogeneities (those visible to the eye)



**Table 1. Typical Nb-Ti Impurity Limits Based on Superconducting Supercollider Specifications**

| Impurity Element | Upper Limit<br>( $\mu\text{L/L}$ ) | Impurity Element | Upper Limit<br>( $\mu\text{L/L}$ ) |
|------------------|------------------------------------|------------------|------------------------------------|
| Ta               | 2500                               | Cu               | 100                                |
| O                | 1000                               | Ni               | 100                                |
| C                | 200                                | Si               | 100                                |
| Fe               | 200                                | Cr               | 60                                 |
| N                | 150                                | H                | 15                                 |
| Al               | 100                                |                  |                                    |

and microinhomogeneities (those requiring identification using microscopes). The most common macroinhomogeneities are Ti-rich "freckles" and hard Nb-rich particles. The Ti-rich freckles are so called from their appearance in ingot cross sections and are a result of Lorentz and buoyancy-driven flow of Ti-rich material between dendrites (see Ref. 18). Control of radial heat transfer and fluid flow in the melt pool eliminates the occurrence of freckles. Because of their relatively small size (typically 1 mm to 2 mm in diameter), compositional deviation (Ti-rich by 8 wt. % Ti to 10 wt. % Ti), and ductility, freckles are not in themselves particularly deleterious to strand production. The importance of the presence or rather the absence of Ti-rich freckles is as an indicator of good melt control. Ti-rich freckles are readily identified from flash radiographs of ingot cross sections. If flash radiography indicates that an ingot cross section is freckle free, it is likely that smaller-scale microinhomogeneities, which are more difficult and expensive to quantify, have been kept to a minimum. A more serious macroinhomogeneity is the presence of hard Nb-rich particles which result in strand breakage failures (see Ref. 19). Nb-rich regions are a result of the high freezing point of Nb and can be eliminated by good process control and remelting. The Nb-rich particles were the cause of many early strand failures but are rarely seen in modern production.

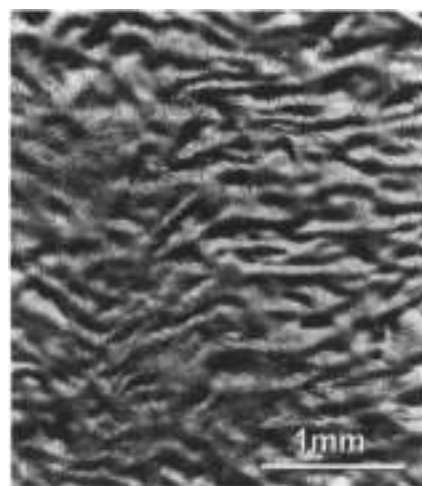
Microchemical inhomogeneity in the cast ingots of Nb-Ti is inevitable because of the coring produced by the large liquid-solid phase separation. The microinhomogeneity level can be quantitatively revealed by metallography using a composition-sensitive etch as shown in Fig. 4. In this example of a high-homogeneity-grade alloy, the microchemical variation is  $\pm 1$  wt. % Ti and has a wavelength of 100  $\mu\text{m}$  to 200  $\mu\text{m}$ . Commercial Nb-Ti alloys have microchemical variations of  $\pm 1$  wt. % Ti to  $\pm 4$  wt. % Ti, with higher-homogeneity alloys casting more. Where high critical current density is less important, reduced microinhomogeneity can be acceptable in order to reduce cost but not to an extent that will reduce strand yield by causing drawability problems during subsequent processing.

The diameter of the initial cast Nb-Ti ingot ranges from 200 mm to 600 mm, and this is typically reduced to 150 mm by hot forging before being fully annealed in the single-phase  $\beta$  region (approximately 2 h at 870 C). Extended anneals can be used to reduce microchemical inhomogeneity but will increase the grain size and consequently reduce the density of precipitate nucleation sites.

#### Composite Assembly

**Stabilizer.** All superconducting strands are fabricated as a combination of one or more continuous filaments of superconductor within a high electrical and thermal conductivity

matrix. The matrix material provides electrical and thermal stability and protection from burnout if the superconducting device reverts to the normal state ("quenches"). High-purity Cu is most commonly used because it has good electrical and thermal conductivity (a resistivity of 0.42 n $\Omega$  m and a thermal conductivity of 200 W/m-K at 6 T, 4.2 K), a high heat capacity, and good strength at both low temperature and during processing and it combines and processes well as a composite with Nb-Ti. A Cu-Ni stabilizer is used for alternating-current (ac) applications where a high-resistivity matrix is required to reduce eddy current loss in the matrix and coupling between the submicron filaments (e.g., Refs. 20 and 21). Where high transverse resistivity is not required, Mn additions to Cu can be used to suppress the proximity coupling of filaments (22). Combinations of high-purity Cu and Cu-Ni or Cu-Mn can be used within the same composite because the alloy stabilizer is only required between filaments. High-purity aluminum has a greater in-field thermal conductivity and electrical conductivity as well as a lower heat capacity, a lower density, and greater radiation resistance than Cu. Despite these considerable advantages, Al is rarely used as a stabilizer because it is extremely difficult to coprocess with Nb-Ti. Furthermore, conventional processing requires the use of precipitation heat treatments that further soften the Al with respect to the Nb-Ti. The use of APC processing avoids the need for precipitation heat treatment and has been used to fabricate Al-stabilized strand with 15 vol. % of Al (23). Additional stabilizer can be added after or late in

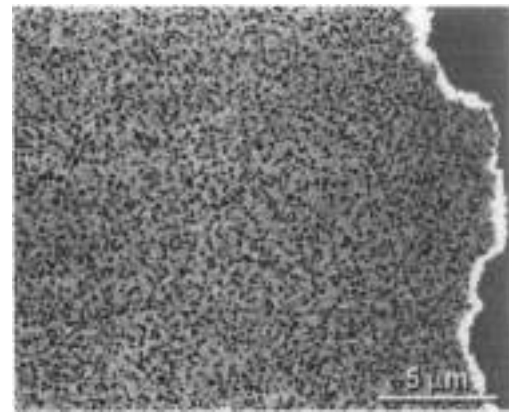


**Figure 4.** Microchemical inhomogeneity in an Nb-Ti alloy can be revealed using a composition-sensitive etch, as in this example of an high-homogeneity Fe-doped Nb-46 wt. % Ti alloy produced by Teledyne Wah Chang.

the fabrication process by external application, by soldering the strand into a channel machined in the external stabilizer, coextruding the strand inside the external stabilizer, or other cladding techniques (24). These methods of applying external stabilizer allow for a greater variety of conductor cross-section designs and stabilizer compositions.

**Diffusion Barriers.** Good fabrication practice results in an excellent metallurgical bond between the composite components by the stage of the process at which precipitation heat treatments are applied. At  $\alpha$ -Ti, precipitation temperatures Cu and Ti at an Nb-Ti/Cu interface will react to form hard intermetallic Ti-Cu (most commonly  $TiCu_3$ ) compounds (25,26). The hard intermetallics do not co-deform with the filaments and will agglomerate as final wire drawing proceeds. The agglomeration of hard particles results in filament nonuniform filament cross sections and ultimately can lead to strand failure. The practice of applying Nb diffusion barriers between the Nb-Ti and the stabilizer to improve filament uniformity and strand predates the acute intermetallic problems caused by aggressive multiple precipitation heat treatments (27). Nb diffusion barriers became commonplace after it was clear that they would be required to achieve the high critical current densities required by the superconducting supercollider project (28). Nb is effective at inhibiting the formation of the Ti-Cu intermetallics and is mechanically compatible with the Nb-Ti alloy rod.

The Nb barrier is normally applied as a sheet wrapped around the Nb-Ti rod before inserting into the stabilizer material. Because the Nb sheet does not contribute to the critical current density, the stability reduces the overall superconductor cross section and is kept to a minimum. The Nb sheet is not impervious to Cu or Ti diffusion and must be kept to a sufficient thickness so that Cu levels in the Nb-Ti do not reach high enough levels during heat treatment that  $TiCu_3$  is formed. As the composite is deformed to smaller and smaller cross section, so is the thickness of the diffusion barrier, thus the most critical processing stage for the Nb diffusion barrier is the final heat precipitation heat treatment. At this stage, not only is the barrier at its thinnest during heat treatment, but rather heat treatments have produced Ti-rich  $\alpha$ -Ti precipitates, some of which will be adjacent to the barrier. Although Cu and Nb have a very low mutual solubility, the high density of Nb and Nb-Ti grain boundaries in the cold-worked composites provides sufficient pathways for Cu diffusion. Faase et al. (29) have calculated that for an aggressive final heat treatment of 50 h at 420 °C a minimum barrier thickness of 0.6  $\mu\text{m}$  is required to prevent reaction layers in the filament (there is no significant reaction in the Cu for Cu-stabilized composites). For smaller filaments, this can represent a significant proportion of the non-stabilizer area. For instance, for a 6  $\mu\text{m}$  final filament diameter design, a 0.6  $\mu\text{m}$  barrier at final heat treatment size represents 4% of total superconductor area. For a 2.5  $\mu\text{m}$  final filament diameter the same thickness diffusion barrier represents a lost superconductor area of over 9%. The calculated predictions of required Nb thickness agree well the development of 6  $\mu\text{m}$  filament strands for the superconducting supercollider where it was shown that 4 area % Nb barriers were required as opposed to 2 area % barriers used in early prototypes (30). Faase et al. also observed that the  $\alpha$ -Ti precipitates adjacent to the Nb barrier were much smaller than elsewhere in the Nb-Ti. In addition to the formation of brittle intermetallics, Moreland et al. (31) found that Cu concentrations of > 3% in the Nb-Ti



**Figure 5.** High resolution back scattered electron scanning electron microscope image of the Nb-Ti adjacent to a Nb diffusion barrier (white) after final precipitation heat treatment. Nonuniformity of the barrier is shown along with a zone of reduced precipitate (black) next to the barrier.

significantly degraded the local energy gap. Reducing the temperature of the final precipitation heat treatment significantly reduces the required minimum barrier thickness but must be balanced against a reduced rate of precipitation. Figure 5 shows a transverse cross section of Nb-Ti filament adjacent to a Nb barrier after it has received the third of three precipitation heat treatments. The nonuniformity of the barrier thickness is clear in Fig. 5 as is a zone of reduced precipitate size adjacent to the barrier.

The grains of the BCC Nb barrier and the Nb-Ti superconductor deform under plane strain conditions such that their (100) directions align with the wire-drawing axis, and they thin in their (001) direction. Maintaining grain boundary continuity requires that both the Nb-Ti and Nb grains must intercurl around their neighbors during cold work (see Ref. 32). At the Nb-Ti/Nb interface the Nb grains intercurl with the Nb-Ti grains resulting in an uneven barrier thickness. As the grain size increases, the scale of Nb-Ti penetration into the barrier increases. Hesser et al. (33) showed that there was a linear relationship between increasing Nb-Ti grain size and increasing variation in Nb barrier thickness, providing an additional incentive to use fine-grained Nb-Ti alloy.

**Strand Geometry and Filament Spacing.** A wide variety of strand geometries are manufactured from Nb-Ti, ranging from large single filament conductors to multifilamentary strands with more than 40,000 filaments. The low-resistivity stabilizer tends to be much softer than the Nb-Ti superconductor, so attention has to be paid to producing a composite that is mechanically stable over the large extrusion and drawing reductions. Gregory et al. (34) established that the optimum mechanical stability for Nb-Ti-Cu multifilamentary composites was for a filament spacing-to-diameter ( $s/d$ ) ratio of 0.15 (ratios of 0.15 to 0.20 are now typical). This ratio fixes the local stabilizer-to-superconductor ratio, and if additional stabilizer is required, it can be designed outside the filament pack. Ghosh et al. (35) established that for direct-current (dc) magnet application a minimum Cu thickness of 0.4  $\mu\text{m}$  to 0.5  $\mu\text{m}$  was required to reduce the magnetization associated with proximity filament coupling. For  $s/d$  ratios of 0.15 to 0.20, the minimum Cu thickness requirement limits the minimum filament diameter to 2.7  $\mu\text{m}$

to 3.3  $\mu\text{m}$  when using a pure Cu matrix. If smaller filaments are required, Ni or Mn can be added to the Cu between the filaments as explained earlier.

**Composite Assembly.** In its simplest form, an Nb-Ti rod with or without a barrier wrap is coextruded inside a high-purity Cu fine-grained can. The warm extrusion produces a well-bonded monofilament. For a multifilamentary application the monofilament composite can be drawn to an intermediate size so that it can be restacked in another Cu can (typically 250 mm to 360 mm in diameter and shaped to fit the extrusion chamber) to form a second extrusion composite. The monofilaments for restacking are usually drawn through a final hexagonal die to facilitate uniform strand packing. In addition to the Nb-Ti/Cu subelements the stack will also contain shaped Cu spacers and Cu boxes to fill voids and provide additional stabilizer if necessary. Some composite designs utilize a central Cu core (usually assembled from hexagonal Cu rods) which is thought to reduce center burst during extrusion (superconducting supercollider specifications required a Cu core). The composite must be densely packed to avoid upset during extrusion. Even after packing with Cu shims, some void space is inevitable in any multifilamentary stack. Normally the amount of void space in a well-packed composite is not sufficient to cause problems during extrusion; but in some cases particularly, when there are large numbers of small filaments, the extrusion billet may have to be isostatically compacted prior to extrusion. Single stacks of more than 22,000 filaments have been successfully fabricated. Alternatively, composites with 200 or fewer filaments can be fabricated by inserting bare Nb-Ti rods into holes in a stabilizer billet prepared by gun-drilling (see Ref. 36 for more examples). A good metallurgical bond is required between the composite components if uniform filament diameters and high yield are to be achieved, and this requires that contamination of the surfaces be kept to a minimum. The mobility of oxygen in Cu at processing temperatures also means that surface oxides must be removed from the high-purity Cu in order to maintain low resistivity and good mechanical properties. When the monofilament is fabricated at large size (100 mm to 200 mm diameter), it benefits from a low surface-to-volume ratio for the components. The restacked monofilaments suffer from a relatively large surface-to-volume ratio that increases as the diameter of the stacked filaments decreases. The importance of reducing chemical and particulate contamination cannot be overemphasized. Long lengths of strand (typically 10 km or more) are desired for many applications, and small particulate pickup that significantly reduces yield by strand breakage can occur in even tightly controlled large-scale production (37).

After the billet is assembled, the filaments are sealed under vacuum by electron-beam (or sometimes tungsten inert gas) welding a Cu lid onto the billet.

### Extrusion

After preheating to 500 to 650°C the extrusion billets are conventionally extruded at ratios of 10:1 to 20:1 and then water-quenched. Lower temperatures are favored because they maintain a greater degree of cold work in the Nb-Ti but they require greater forces. Low extrusion speeds avoid excessive heating of the Nb-Ti. Hydrostatic extrusion may also be used when available and has the advantages of lower-temperature operation, higher yields available from higher length-to-diameter ratios,

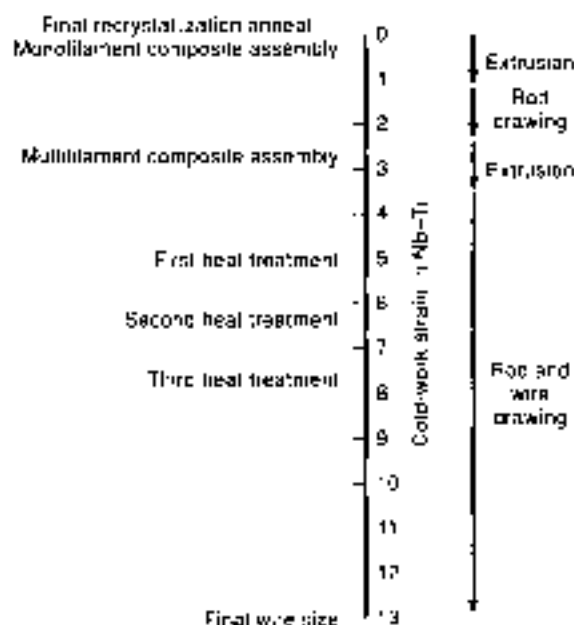


Figure 6. Schematic illustration of the conventional process for Nb-Ti superconductors in terms of cold-work strain.

and an ability to coextrude a wider variety of material combinations. After successful extrusion, the composite is a well-bonded composite of a suitable size (50 mm to 90 mm in diameter) for cold work to proceed.

### Cold Work

Superconducting Nb-Ti strands are fabricated in a cold-worked condition that is far greater than any other metallurgical process, with the possible exception of piano wire. The object of the cold work is to produce the final nanometer scale dispersion of pinning centers. Understanding the role and importance of this highly strained state in the production of Nb-Ti strands is the key to the conventional process. The amount of cold work in the Nb-Ti normally can be represented by the true strain,  $\epsilon_T$ , which is given by Eq. (1):

$$\epsilon_T = \ln A_0/A = 2 \ln d_0/d \quad (1)$$

where  $d_0$  and  $A_0$  are the starting diameter and cross-sectional area of the alloy rod at the final recrystallization anneal, respectively, and  $d$  and  $A$  are the diameter and area of the Nb-Ti after cold work. Figure 6 illustrates the conventional process in terms of cold-work strain for a composite fabricated from a cold-worked rod and given three heat treatments. The total available cold-work strain available for processing the Nb-Ti is termed the "strain space" and is limited by the initial annealed Nb-Ti rod size and the final filament diameter. Warm extrusions increase the cold work in the Nb-Ti at a reduced rate compared with similar reductions performed by rod or wire drawing and consequently reduce the effective strain space. Increasing the temperature of extrusion improves the bonding between the composite components but reduces the available cold-work strain.

The strain space can be further subdivided into three regions, each requiring a minimum cold-work strain to be effective.

1. Prestrain,  $\epsilon_p$ , the cold-work strain before the initial precipitation heat treatment.
2. The inter-heat treatment strain,  $\Delta\epsilon_{HT}$ , the strain required between precipitation heat treatments.
3. The final strain,  $\epsilon_f$ , the strain required to reduce the precipitate size to final optimum pinning size.

This cold work performs seven primary functions:

1. Encouraging the formation of the preferred precipitate phase and morphology.
2. Improving microchemical homogeneity by mechanical mixing both prior to heat treatment and after heat treatment when local Ti depletion has occurred.
3. Increasing the density of precipitate nucleation sites.
4. Increasing the grain boundary density, thereby increasing diffusion rates (grain boundary diffusion being considerably faster than bulk interdiffusion).
5. Reducing the average diffusion distance to the precipitate nucleation site.
6. Increasing the volume of precipitate by multiple strain/heat treatment cycles.
7. Reducing the precipitate dimensions from the precipitation scale of 100 nm to 300 nm diameter to the pinning scale of 1 nm to 5 nm.

The cold work is applied to the extruded composite by standard rod- and wire-drawing techniques. The rod or wire is pulled through a shaped die that uniformly reduces the cross-sectional area by 15% to 25%. As the strand becomes smaller in diameter, multiple die machines are used to apply more than one area reduction in a single pass. The amount of cold work in a single-phase Nb-Ti alloy can be monitored using hardness testing. As the strain in the Nb-Ti increases, so does the hardness [38].

Buckett and Larbalestier [39] established that a minimum cold-work strain of approximately 5 was required before precipitation heat treatment to ensure the optimum precipitation mode, that of  $\alpha$ -Ti precipitates located at grain boundary intersections. At lower strains,  $\alpha$  phase and/or intragranular Widmanstätten were formed, producing inhomogeneous microstructures that reduced workability. Lee et al. [40] systematically studied the relationship between prestrain, precipitate morphology, and alloy and showed a relationship between Nb-Ti composition and the required cold-work prestrain required to avoid strains  $\alpha$  phase and/or intragranular Widmanstätten  $\alpha$ -Ti [40]. Figure 7 illustrates the strong increase in required prestrain with increasing Ti content. This relationship is very important because it not only explains the importance of a homogeneous alloy composition but also explains the difficulties associated with processing high Ti alloys. An increase in  $\epsilon_p$  of roughly 0.77 is required for each wt. % Ti increase in composition. Figure 8 shows the large amount of  $\alpha$  phase and/or intragranular Widmanstätten  $\alpha$ -Ti that can be produced by heat treating a high-Ti alloy (in this case, Nb-54 wt. % Ti) at too low a prestrain. By a true strain of 5 to 7 a high-angle grain structure of regular, small diameter (50 nm to 100 nm) grains has been produced by the large cold-work strain. Increasing the prestrain further homogenizes and refines the microstructure.

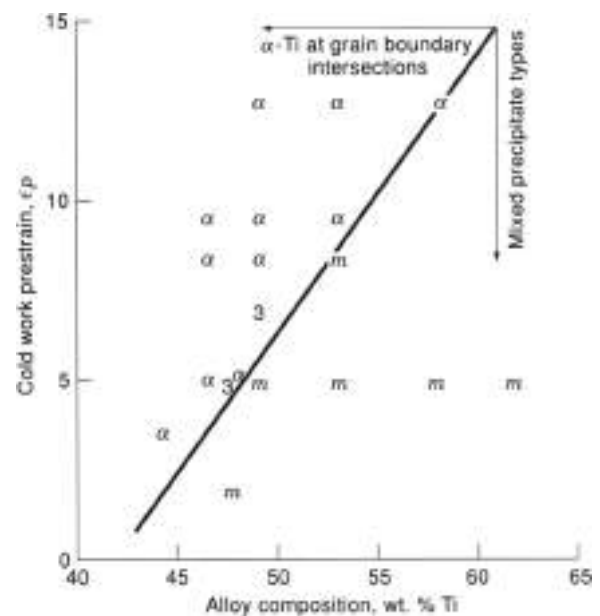


Figure 7. The morphology and location of precipitation in cold-worked Nb-Ti alloys are determined by the amount of cold work in the microstructure. As the Ti content is increased, the cold work required to produce optimum  $\alpha$ -Ti only at grain boundary intersections, increases. In this compilation of data [39,40,44], optimum precipitates is denoted by the symbol  $\alpha$ , and mixed precipitate morphology is denoted by the letter *m* for heat treatment at 420 C. The number 3 identifies where mixed-mode precipitation has also been observed for 375 C heat treatments.

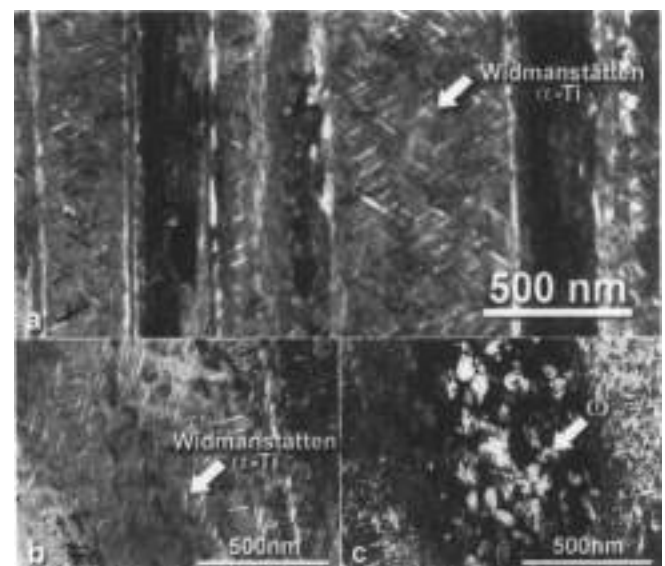


Figure 8. Transmission electron microscope image in bright field (a), (b), and dark field using an omega reflecting (c) of an Nb-54 wt. % Ti alloy after precipitation heat treatment. The heat treatment was applied at a cold-work prestrain of 5, which is insufficient to avoid the deleterious  $\alpha$  phase and/or intragranular Widmanstätten  $\alpha$ -Ti forms of precipitation. In (a) the microstructure is shown as longitudinal cross section with the drawing axis running down the page. Images (b) and (c) are of the same transverse cross-sectional area with the dark-field image bringing out contrast from the  $\alpha$  phase.

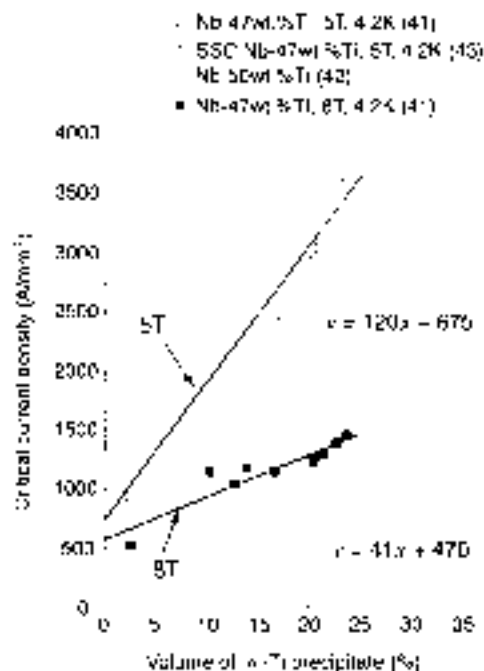


Figure 9. The peak critical current density increases linearly with volume percent of  $\alpha$ -Ti precipitates over the range of 3% to 25 vol. %  $\alpha$ -Ti for Nb-47 wt. % Ti (41,43). Similar results have been shown for 50 wt. % Ti (42). Least squares fit dependencies for the Nb-47 wt. % Ti data are shown for applied fields of 5 T and 8 T.

#### Precipitation Heat Treatment

Lee et al. (41) first established a linear relationship between the optimized critical current density and the volume of precipitate in a laboratory-scale monofilamentary composite fabricated from Nb-47 wt. % Ti alloy as shown in Fig. 9. The relationship extended from 0% of the strand volume being precipitated (non-heat-treated) to 25 vol. %, Chernyj et al. (42) extended this relationship to an Nb-50 wt. % Ti alloy and a maximum volume percent of  $\alpha$ -Ti of 25%. A wider study of strands produced by different manufacturers for the superconducting supercollider confirmed the linear relationship for industrial-scale strands (43). The importance of maximizing the amount of precipitate in the strand is unambiguous. Precipitate is produced in the Nb-Ti by heat treatments at 375–420°C for a duration of typically 40 h to 80 h. Increasing the temperature increases the precipitation rate but increases the precipitate diameter and the low-field  $J_c$  to high-field  $J_c$  ratio (44). The amount of precipitation is also dependent on the alloy composition; the quantity of  $\alpha$ -Ti produced by the first precipitation heat treatment increases strongly with Ti content (44) as shown in Fig. 10. This relationship shows how too low a Ti content in the Nb-Ti alloy can result in insufficient precipitation for high critical current density and how a large local variation in Ti content can lead to an inhomogeneous distribution of precipitates and, subsequently, flux-pinning sites.

After approximately 10 vol. % precipitate has been produced in the first heat treatment, it becomes very difficult to produce significantly more without excessively long heat treatment times. By applying additional cold-work strain to the microstructure, more precipitate is produced (as shown in Fig. 10 for the second heat treatment). An optimum balance between

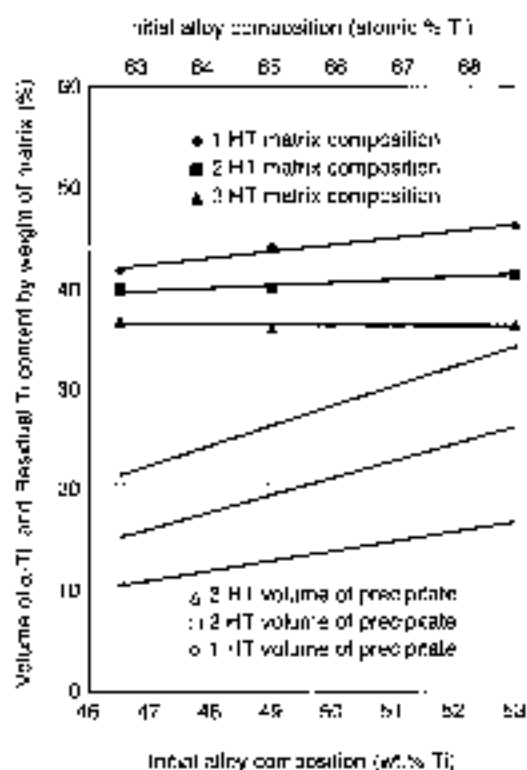


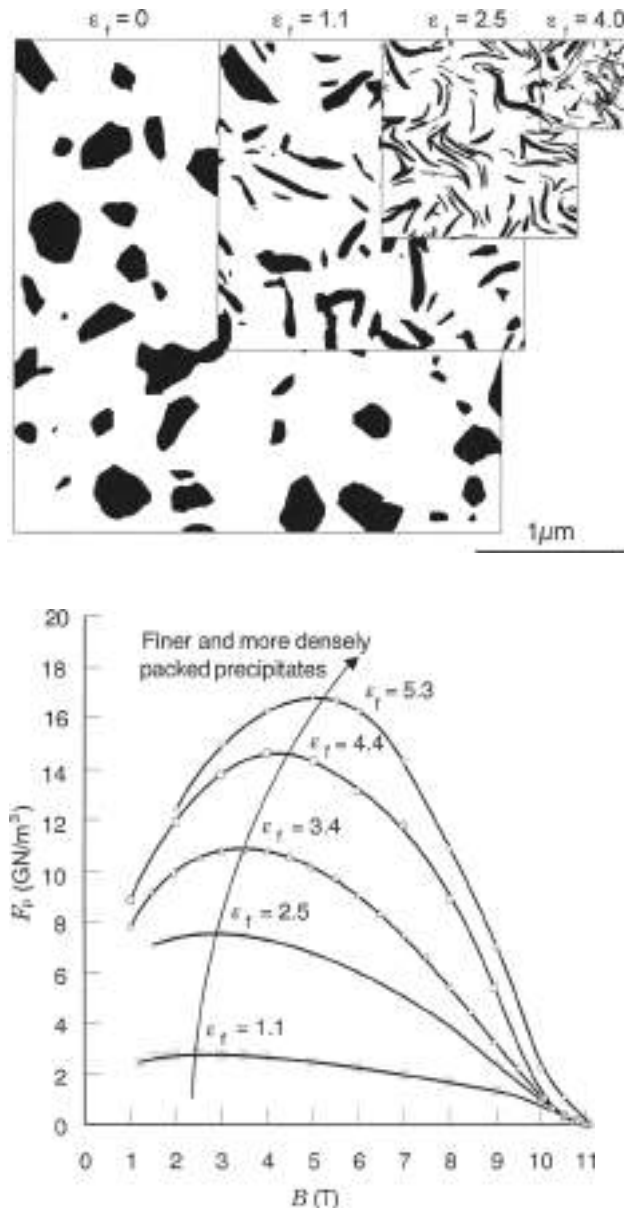
Figure 10. The precipitation rate in Nb-Ti increases strongly with Ti content in Nb-Ti alloys. Additional heat treatments further increase the amount of precipitate. As additional heat treatment and strain cycles are applied and more precipitate is produced, the residual Ti content of the  $\beta$ -Nb-Ti matrix drops and insufficient Ti is left to drive further precipitation. In this graph, data from one and two heat treatments (40) are combined with data from three heat treatments of 80 h at 420°C. The average residual matrix composition is calculated assuming an  $\alpha$ -Ti composition of Nb-47 wt. % Ti.

increased precipitate volume and minimum strain space is at a strain of approximately 1.2 (12). Three or more heat treatment and strain cycles are normally required to produce the 20 vol. % or more precipitate in the microstructure required for high critical current densities ( $J_c > 3000$  A/mm<sup>2</sup> at 5 T and 4.2 K). As the  $\alpha$ -Ti is precipitated, the composition of the  $\beta$ -Nb-Ti is depleted in Ti until it reaches between 36 wt. % Ti and 37 wt. % Ti, at which point there is insufficient Ti to drive further precipitation. More aggressive heat treatment is more likely to compromise the Nb diffusion barrier and coarsen the precipitate size.

After the final heat treatment the microstructure viewed transverse to the drawing axis consists of a uniform distribution of roughly equiaxed  $\alpha$ -Ti precipitates, 80 nm to 200 nm in diameter, in a matrix of equiaxed Nb-Ti grains of similar dimensions. Viewed in longitudinal cross section the  $\alpha$ -Ti and  $\beta$ -Nb-Ti grains are somewhat elongated along the drawing axis with an aspect ratio of 4 to 15 depending on the processing history. Further cold-work strain is required to reduce the dimensions of the precipitates so that they can pin flux efficiently. During the  $\alpha$ -Ti precipitation heat treatments, the  $\beta$ -Nb-Ti matrix has been depleted in Ti to a level of 37 wt. % Ti to 38 wt. % Ti, and the  $H_{c2}$  and  $T_c$  of the composite at this point in processing are the same as the values of single-phase material of these lower Ti levels (5).

### Final Wire Drawing

The plain strain imposed intercoiling of the Nb-Ti grains that is so deleterious to barrier uniformity also results in the distortion of the  $\alpha$ -Ti precipitates into densely folded sheets during final wire drawing. The folding process rapidly decreases the precipitate thickness and spacing and with a dependence of  $d^{-2}$  (where  $d$  is the strand diameter) and increases the precipitate length per area with a dependence of  $d^{-1/2}$  as measured by Meingast et al. (5). As the microstructure is refined toward optimum size the bulk pinning force increases and the peak in the bulk pinning force curves to higher field as shown in Fig. 11 (data from Ref. 45). The  $H_{c2}$  and the  $T_c$  gradually return to the



**Figure 11.** For conventionally processed Nb-Ti the bulk pinning force increases in magnitude with drawing strain after the last heat treatment. The increase occurs at all fields as the precipitate size and spacing are reduced to less than a coherence length in thickness (45). The refinement of the microstructure with increasing strain for the same strand is shown schematically in transverse cross-sections with the  $\alpha$ -Ti precipitates in black.

values of the original single phase starting alloy as the precipitates are refined toward and below the superconducting coherence length,  $\xi$  (5). The critical current density increases as the microstructure is refined until it reaches a peak, after which there is a steady decline. The peak in  $J_c$  for a monofilament or a multifilamentary strand with uniform filaments occurs at a final strain of approximately 5. If the filaments are nonuniform in cross section (sausaged), the peak occurs earlier and at a lower critical current density. A strand that has a premature (and lowered) peak in  $J_c$  during final drawing is described as extrinsically limited because it has not attained the intrinsic critical current of the microstructure. The most common source of extrinsic limitation is sausaging of the filaments due to intermetallic formation or lack of bonding between the components of the composite. The degree to which a composite has been extrinsically limited can be observed by examining the sharpness of the resistive transition when measuring the critical current. *J.* Volker (46) showed that the shape of the transition curve near its onset can be approximated by

$$V \propto I^n \quad (2)$$

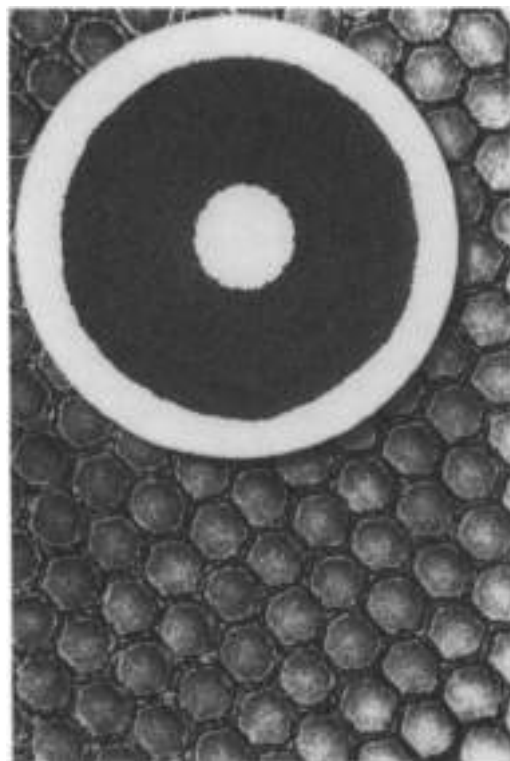
where  $V$  is the voltage across,  $I$  is the current in the superconductor, and  $n$  is the resistive transition index. For a nonextrinsically limited superconductor the value of  $n$  at 5 T, 4.2 K, can be 70 or higher. By quantifying the variation in filament cross-sectional area by image analysis, the amount of filament sausaging can be measured directly (47). A high critical current density superconductor with a high  $n$ -value is shown in Fig. 12. The strand is one of a number of high-performance wires developed for the superconducting supercollider (48). The filament sausaging in this strand has been reduced to a very low level: a coefficient of variation for the filament cross-sectional areas of approximately 2%. With tight quality control, uniform properties and piece lengths exceeding 10 km should be expected.

The specific pinning force for the  $\alpha$ -Ti precipitates falls from 360  $\text{N}/\text{m}^2$  for an average sheet thickness of 2.6 nm, to 200  $\text{N}/\text{m}^2$  for a 1 nm average sheet thickness (14,49) but this is more than compensated for by the increase in precipitate density caused by the continued folding of the  $\alpha$ -Ti sheets.

### Final Processing

**Twisting.** Just before a multifilamentary strand has reached final size, it is usually twisted about its drawing axis. The twisting is required to reduce flux-jump instability caused by varying external fields, reduce instabilities caused by self field, and reduce eddy-current losses. The tightness of the required twist increases with the expected rate of change of field. The required twist pitch for a superconducting supercollider strand, a relatively steady-state magnet, was approximately 50 rotations along the drawing axis per meter, while for an application with a similarly sized strand the number of twists per meter might be 300. The twisting occurs just before the strand is reduced to final size so that it can be locked in by a final drawing pass or by final shaping.

**Final Shaping.** The final shape of the strand cross section need not be round in cross section. It can also be shaped into square or rectangular cross section by the use of independently adjusted rollers operating along the strand surface.



**Figure 12.** A high-quality strand at final size manufactured by IGC-AS for the superconducting supercollider. The superconducting Supercollider Laboratory main ring dipole magnets would have required almost 1500 tons of 7500 to 8000 filament superconductor. An overview of the strand cross section is superimposed on a higher-magnification image showing the filament cross sections in more detail. Both the filament size (diameter of 6  $\mu\text{m}$ ) and spacing are very uniform. An Nb diffusion barrier can be seen surrounding each filament.

**Cabling.** Individual strands can be cabled or braided together to form a conductor with a higher current-carrying capacity. The most common design for Nb-Ti magnets is the Rutherford cable, which consists of two parallel flat layers of strands. Using this approach, high-aspect-ratio cables can be produced with as many as 48 strands (50). As was the case for the individual filaments, the strands are transposed around the cable, forming a densely packed square or rectangular cross-section spiral. The design considerations for Superconducting Supercollider Laboratory cable are discussed in Ref. 51. The compaction of the strand around the squared cable edges severely distorts the strand cross section, but the excellent mechanical properties of Nb-Ti/Cu composites combined with good strand design and advances in cabling technology have reduced cabling degradation to minimal levels (52).

### Nb-Ti-Ta

The addition of Ta to Nb-Ti alloys suppresses the paramagnetic limitation of  $H_{c2}$  by the large orbital moment of the alloys (53). Although Ta is only of benefit below 4.2 K (54), it has a relatively long history of study because it should extend the useful field range of ductile superconductors by 1 T or more (55). So far, however, improved  $H_{c2}$  has not translated effectively into improvements in  $J_c$ , except very near to  $H_{c2}$  (above 11 T). Lazarev et al. (56) were able to attain a critical current density (1000  $\text{A}/\text{mm}^2$  at a field of 11.5 T  $\cdot$  2.05 K) using an Nb-37 wt. % Ti-

22 wt. % Ta alloy. Ta has an even higher melting point than Nb, making the fabrication of chemically homogeneous ternary alloys particularly difficult. The behavior of Nb-Ta-Ta alloys under the conventional process is similar to that of binary alloys, but the precipitates do not appear to pin as efficiently (57).

### THE APC PROCESS

The quantity, composition, and distribution of pinning center as well as the composition of the matrix are limited, in the conventional process, by the thermodynamics of the Nb-Ti phase diagram. Additional precipitate can be produced by increasing Ti content of the alloy (as shown in Figure 10), but that is more than offset by the decrease in  $H_{c2}$  (Fig. 2). The result is a critical current limit in conventionally processed Nb-Ti superconductors of approximately 3500  $\text{A}/\text{mm}^2$ , at 4.2 K and 5 T. An alternative approach is to fabricate the microstructure by mechanically assembling the desired components of the microstructure at large size and reducing the microstructure to the appropriate size by extrusion and cold drawing (58,59). The engineered microstructural rods can be restacked into a composite just as for a conventional Nb-Ti superconductor, but no precipitation heat treatments are required. An intermediate approach developed by Supercon, Inc. (60) uses a low-temperature diffusion heat treatment to modify a densely packed microstructure fabricated from layer of pure Nb and Ti. The diffusion-modified APC has been successfully used in solenoid, model dipole (61), and MRI magnets (62). Round-wire APC superconductors and multilayers have developed zero-field  $J_c$  up to 10% of the theoretical upper limit provided by the depairing current density  $J_d$  ( $J_d \sim H_c/\lambda$ ) (e.g., Refs. 63 and 64), where  $\lambda$  is the penetration depth. APC superconductors fabricated with Nb pins perform particularly well at low fields (up to about 5 T to 7 T), and  $J_c$  values approaching 7500  $\text{A}/\text{mm}^2$  at 3 T (65,66) have been achieved (25% of Nb pinning center in an Nb-47 wt. % Ti matrix). Nb has been a preferred pinning material because of its mechanical compatibility with the Nb-Ti matrix. Even using Nb, however, poor workability and increased costs associated with assembly and yield have so far limited the commercial application of APC composites. The components of an engineered microstructure must initially be large enough to be stacked by hand (or possibly machined); consequently the engineered pins must undergo a far greater deformation to reach optimum size than for  $\alpha$ -Ti precipitates which start at 100 nm to 200 nm in diameter. The larger deformation and multiple extrusions and the restacks required by the APC process result in a microstructure that can be much less uniform than for the conventional process (67). For this reason, processes that can use smaller cross-sectional starting dimension, such as stacked or wrapped sheet, can result in superior properties such as the  $J_c$  of 4250  $\text{A}/\text{mm}^2$  at 5 T and 4.2 K, achieved by Matsumoto et al. (68) with stacked sheets of Nb-50 wt. % Ti and 28 vol. % of Nb shorts. Because of the large amount of cold work in the engineered microstructure, it is extremely sensitive to heating during extrusion; the highest round-wire  $J_c$  (5 T, 4.2 K) of 4600  $\text{A}/\text{mm}^2$  was achieved by Heuser et al. (69). For the Nb pins, similar volumes of pinning material are required as for conventionally processed materials; but by using ferromagnetic pins (Fe or Ni) the required pin volume to achieve high critical current density has been reduced to only 2 vol. % (70). Such developments suggest that there are still exciting advances

that can be made in the development of ductile Nb-Ti based superconductors.

## BIBLIOGRAPHY

1. E. W. Collings, *Applied Superconductivity, Metallurgy, and Physics of Titanium Alloys*, vol. 1: Fundamentals, New York: Plenum, 1986.
2. E. W. Collings, *Applied Superconductivity, Metallurgy, and Physics of Titanium Alloys*, vol. 2: Applications, New York: Plenum, 1986.
3. E. W. Collings, *A Sourcebook of Titanium Alloy Superconductivity*, New York: Plenum, 1983.
4. H. Müller, *The upper critical field of niobium-titanium*, PhD thesis, Univ. Wisconsin—Madison, 1980.
5. C. Meingast, P. J. Lee, and D. C. Larbalestier, Quantitative description of a high  $J_c$  Nb-Ti superconductor during its final optimization strain. I. Microstructure,  $T_c$ ,  $H_{c2}$  and resistivity, *J. Appl. Phys.*, **66**: 5962-5970, 1989.
6. D. L. Moffat, *Phase transformations in the titanium-niobium binary alloy system*, PhD thesis, Univ. Wisconsin—Madison, 1965.
7. I. Pfeiffer and H. Hillmann, Der Einfluss der Struktur auf die Supraleitungseigenschaften von NbTi<sub>50</sub> und NbTi<sub>65</sub>, *Acta Metall.* **16**: 1429-1439, 1968.
8. M. Hansen et al., Systems titanium-niobium and titanium-columbium, *J. Metals*, **3**: 881-883, 1951.
9. J. Kaufman and H. Bardeen, *Computer Calculations of Phase Diagrams*, New York: Academic Press, 1973.
10. D. L. Moffat and U. R. Kattner, The stable and metastable Ti-Nb phase diagram, *Metal. Trans.*, **19A**: 2389-2397, 1988.
11. D. L. Moffat and D. C. Larbalestier, The competition between martensite and omega in quenched Ti-Nb alloys, *Metal. Trans.* **19A**: 1677-1686, 1988.
12. L. Cheungren and D. C. Larbalestier, Development of high critical current densities in niobium 46.5 wt % titanium, *Cryogenics*, **27**: 171-177, 1987.
13. D. C. Larbalestier and A. W. West, New perspectives on flux pinning in niobium-titanium composite superconductors, *Acta Metall.*, **32**: 1871-1881, 1984.
14. P. J. Lee and D. C. Larbalestier, Development of nanometer scale structures in composites of Nb-Ti and their effect on the superconducting critical current density, *Acta Metall.*, **36**: 2526-2536, 1987.
15. D. C. Larbalestier et al., High critical current densities in industrial scale composites made from high homogeneity Nb46.5Ti, *IEEE Trans. Magn.*, **21**: 269-272, 1985.
16. *Specification for Niobium-Titanium Alloy for the Superconducting Super Collider*, MST Document Control, Dallas, Texas: Superconducting Super Collider Laboratory, 1992. SSC-Mag-M-4000A.
17. D. B. Smathers et al., Improved niobium 47 weight % titanium composition by iron addition, *Mater. Trans., Jpn. Inst. Metals*, **37** (3): 519-526, 1996.
18. R. Schaffer, Electrical and magnetic interactions in vacuum arc remelting and their effect on the metallurgical quality of specialty steels, *J. Vac. Sci. Technol.* **11**(6): 1047-1054, 1974.
19. H. Hillmann, Fabrication technology of superconducting material, in S. Foner and B. B. Schwartz (eds.), *Superconducting Materials Science*, New York: Plenum, 1981, p. 275.
20. J. Hlusnik et al., Properties of superconducting NbTi superfine filament composites with diameter  $\approx 0.1 \mu\text{m}$ , *Cryogenics*, **25**: 559-564, 1985.
21. J. R. Cave et al., Reduction of AC loss in ultra-fine multifilamentary NbTi wires, *IEEE Trans. Magn.*, **20**: 21: 1945-1948, 1983.
22. T. S. Kreilik et al., Reduction of coupling in line filamentary Cu/NbTi composites by the addition of manganese to the matrix, *Adv. Cryog. Eng.*, **34**: 895-900, 1988.
23. M. K. Rudzick, T. Wong, and J. Wong, Development of APC NbTi superconductors with internal high purity aluminum stabilizer, *IEEE Trans. Appl. Supercond.*, **7**: 1197-1209, 1997.
24. H. C. Keskkilä, D. Phillips, and B. A. Zettlin, Further development of aluminum clad superconductors, *IEEE Trans. Magn.*, **27**: 1800-1806, 1991.
25. D. C. Larbalestier et al., High critical current densities in fine filament NbTi superconductors, *IEEE Trans. Nucl. Sci.*, **32**: 3743-3745, 1985.
26. M. Garber et al., Effect of Cu<sub>2</sub>Ti compound formation on the characteristics of NbTi accelerator magnet wire, *IEEE Trans. Nucl. Sci.*, **32**: 3691-3693, 1985.
27. E. Gregory, in R. W. Meyerhuff (ed.), *Manufacture of Superconducting Materials*, Metals Park, OH: Amer. Soc. Metals, 1977, pp. 1-16.
28. R. M. Scanlan, J. Boyet, and R. Hannaford, Evaluation of various fabrication techniques for fabrication of fine filament NbTi superconductors, *IEEE Trans. Magn.*, **23**: 1719-1723, 1987.
29. K. J. Puase et al., Diffusion reaction rates through the Nb vapor in SSC and other advanced multi-filamentary Nb-46.5 wt % Ti composites, *Adv. Cryog. Eng.*, **38**: 723-730, 1992.
30. Y. High et al., Quantitative analysis of saussaging in Nb barrier clad filaments of Nb-46.5 wt % Ti as a function of filament diameter and heat treatment, *Adv. Cryog. Eng. (Mater.)*, **38**: 647-652, 1992.
31. J. Mousland, J. W. Ekin, and L. E. Goodrich, Electron tunneling into superconducting filaments: Depth profiling the energy gap of NbTi filaments from magnet wires, *Adv. Cryog. Eng. (Mater.)*, **32**: 1161-1168, 1986.
32. W. P. Bosford, Jr., Microstructure changes during deformation of (111) fiber textures metals, *Trans. Metall. Soc. AIME*, **230**: 32-35, 1964.
33. R. Heuser, P. Lee, and D. Larbalestier, Non-uniform deformation of niobium diffusion barriers in niobium-titanium wire, *IEEE Trans. Appl. Supercond.*, **3**: 757-760, 1993.
34. E. Gregory et al., Importance of spacing in the development of high current densities in multifilamentary superconductors, *Cryogenics*, **27**: 178, 1987.
35. A. K. Ghosh et al., Anomalous low field magnetization in fine filament NbTi conductors, *IEEE Trans. Magn.*, **23**: 1724-1729, 1987.
36. T. S. Kreilik, Niobium-titanium superconductors, in *Metals Handbook*, 10th ed., vol. 2: *Properties and Selection: Nonferrous Alloys and Special-Purpose Materials*, ASM Int., 1990, pp. 1043-1057.
37. P. Valaris et al., A statistical evaluation of recent SSC conductors produced at IGC/ASU, *IEEE Trans. Magn.*, **27**: 1752-1754, 1991.
38. J. Farrell, P. Lee, and D. Larbalestier, Cold work loss during heat treatment and extrusion of Nb-46.5 wt % Ti composites as measured by microhardness, *IEEE Trans. Appl. Supercond.*, **3**: 734-737, 1993.
39. M. J. Buckley and D. C. Larbalestier, Precipitation at low strains in Nb346.5 wt % Ti, *IEEE Trans. Magn.*, **23**: 1638-1641, 1987.
40. P. J. Lee, J. C. McKinell, and D. C. Larbalestier, Microstructure control in high Ti-NbTi alloys, *IEEE Trans. Magn.*, **MA6-25**: 1916-1924, 1989.
41. P. J. Lee, J. C. McKinell, and D. C. Larbalestier, Restricted novel heat treatments for obtaining high  $J_c$  in Nb-46.5 wt % Ti, *Adv. Cryog. Eng. (Mater.)*, **36**: 297-299, 1990.
42. O. V. Cherny et al., Nb-Ti superconductors of a high current-carrying capacity, *Supercond. Sci. Technol.*, **4**: 318-323, 1991.
43. P. J. Lee and D. C. Larbalestier, An examination of the properties of SSC Phase II R&D strands, *IEEE Trans. Appl. Supercond.*, **3**: 833-841, 1993.
44. P. J. Lee, J. C. McKinell, and D. C. Larbalestier, Progress in the un-



- derstanding and manipulation in high  $J_c$  Nb-Ti alloy composites, in Y. Murakami (ed.), *Proc. of New Developments in Appl. Supercond. Singapore* (World Scientific, 1989), pp. 257-262.
45. C. Meingast and D. C. Larbalestier, Quantitative description of a very high critical current density Nb-Ti superconductor during its final optimization stage: II. Flux pinning mechanisms, *J. Appl. Phys.* **68**: 5871-5883, 1990.
  46. F. Voelker, Resistance in small, twisted, multifilare superconducting wires, *Particle Accelerators*, **1**(3): 205-207, 1970.
  47. Y. High et al., Quantitative analysis of nanaging in Nb barrier clad filaments of Nb-46.5 wt% Ti as a function of filament diameter and heat treatment, *Adv. Cryog. Eng. Mater.*, **30**: 647-662, 1992.
  48. P. Valaris et al., SSC-1 inner cable development program at IGC Advanced Superconductors, Inc., *IEEE Trans. Appl. Supercond.*, **3**: 705-709, 1993.
  49. P. J. Lee and D. C. Larbalestier, Determination of the flux pinning force of a Ti ribbons in Nb-46.5 wt% Ti produced by heat treatments of varying temperature, duration and frequency, *J. Mater. Sci.*, **29**: 3971-3987, 1994.
  50. R. Scanlan et al., Design and fabrication of a high aspect ratio cable for a high gradient quadrupole magnet, *IEEE Trans. Appl. Supercond.*, **7**: 936-938, 1997.
  51. R. Christopherson et al., SSC 40 mm cable results and 60 mm design discussions, *IEEE Trans. Magn.*, **27**: 1881-1883, 1991.
  52. M. Garber et al., Superconducting wire and cable for RHIC, *IEEE Trans. Magn.*, **30**: 1732-1735, 1994.
  53. M. Suenaga and K. M. Ball, Some superconducting properties of Ti-Nb-Ta ternary alloys, *J. Appl. Phys.*, **40**: 4457-4463, 1971.
  54. D. G. Hawksworth and D. C. Larbalestier, Enhanced values of  $H_{c2}$  in Nb-Ti ternary and quaternary alloys, *Adv. Cryog. Eng.*, **26**: 429-446, 1990.
  55. H. H. Segal et al., NbTi liquid conductors for use in 12 Tesla toroidal field coils, *Proc. 8th Symp. Eng. Problems Fusion Res.* (New York: IEEE Press, 1980), pp. 255-259.
  56. R. G. Lazarev et al., The study of the microstructure and  $J_c$  in Nb-37Ti-22Ta superconductor produced with different duration of treatments, in H. W. Weber (ed.), *Proc. 7th Int. Workshop Critical Currents Supercond.* (Alpach, Austria: Singapore: World Scientific, 1994), pp. 601-604.
  57. P. J. Lee et al., Microstructure-property relationships in Nb-Ti-Ta, *IEEE Trans. Appl. Supercond.*, **3**: 1354-1357, 1993.
  58. G. L. Dudašev, E. Yu. Klumenko, and S. V. Enlov, Current carrying capacity of superconductors with surface pinning centers, in C. Mazurek and P. Weymuth (eds.), *Proc. 19th Int. Conf. Magnet Technol.* (Wolphen, Switzerland: Swiss Institute Nuclear Res., 1985), pp. 564-566.
  59. B. A. Zeitlin, M. S. Walker, and J. R. Matwadh, *Superconductors having controlled filament pinning centers, and method of manufacturing same*, US Patent 4,602,310, 1989.
  60. M. K. Hudziak et al., Development of APC Nb-Ti composite conductors at Supercon, Inc., *IEEE Trans. Appl. Supercond.*, **5**: 1539-1542, 1995.
  61. R. M. Scanlan et al., *IEEE Trans. Magn.*, **30**: 1627-1632, 1994.
  62. U. Reinold et al., First commercial application of NbTi superconductor employing artificial pinning centers, *IEEE Trans. Appl. Supercond.*, **5**: 1189-1192, 1995.
  63. G. Stojic et al., Effect of geometry on the critical currents of thin films, *Phys. Rev.*, **B40**: 1274-1286, 1984.
  64. F. Kadyrov, A. Gurevich, and D. C. Larbalestier, High critical current densities in Nb47%Ti multilayers with planar copper-Pb pinning nanostructure, *Appl. Phys. Lett.*, **68**: 1567-1569, 1996.
  65. L. R. Matwadh et al., Modification of NbTi with artificial pinning centers: The effect of alloy and porosity on the superconducting properties, *Appl. Phys. Lett.*, **61**(16): 991-993, 1992.
  66. H. C. Kanulki et al., Further developments in NbTi superconductors with Artificial Pinning Centers, *Adv. Cryog. Eng. Mater.*, **36**: 675-683, 1992.
  67. P. J. Lee, D. C. Larbalestier, and P. D. Jafarzadeh, Quantification of pinning center thickness in conventionally processed and powder processed artificial pinning center microstructures, *IEEE Trans. Appl. Supercond.*, **5**: 1701-1704, 1995.
  68. K. Matsumoto et al., Enhanced  $J_c$  properties in superconducting NbTi composites by introducing Nb artificial pins with a layered structure, *Appl. Phys. Lett.*, **64**(1): 115-117, 1994.
  69. R. W. Heuzarot et al., Increased critical current density in Nb-Ti wires having Nb artificial pinning centers, *Appl. Phys. Lett.*, **70**(7): 901-903, 1997.
  70. N. D. Kuzov et al., Ferromagnetic artificial pinning centers in superconducting Nb<sub>81</sub>Ti<sub>19</sub> wires, *Appl. Phys. Lett.*, **68**(15): 2265-2267, 1996.

PETER J. LEE  
University of Wisconsin - Madison

## STABILITY IN FORCED FLOW

Forced-flow-cooled conductors are the preferred choice for magnets that must operate in an electromagnetic and mechanically noisy environment, when pulsed operation requires minimization of ac losses, or whenever the operating conditions require a reliable and cost-effective design. In this article we review the guidelines that motivated the choice of forced-flow-cooled conductors to obtain an effective and stable superconductor design for large magnets, such as those for fusion, superconducting magnetic energy storage (SMES), particle detectors, or magneto-hydrodynamic (MHD) application. We will discuss the particular features of the stability margin in forced-flow-cooled conductors and the models commonly used to compute it.

## SUPERCONDUCTOR STABILITY

Superconductors exhibit zero resistance only within relatively narrow parameters of temperature, magnetic field, and transport current, below the so-called *critical surface*. When brought outside this region by a disturbance (e.g., by energy deposition stemming from a mechanical motion), superconductivity is lost and Joule heating is generated. If not prevented by other mechanisms, the superconductor cascades further from its nominal operating point into an irreversible process leading to the complete loss of superconductivity in the magnet. This process is commonly known as a *quench*. Even if the magnet is properly protected against damage, a magnet quench is an undesirable event in terms of availability and cost. A well-designed magnet will not quench under normal operating conditions. The study of stability pertains to the understanding of the processes and mechanisms whereby a superconductor will remain (or not) within its operating region, thus ensuring magnet operation without quench. This area of study has evolved through many years of experimentation and analysis.

### Stekly Criterion for Cryostability

The first superconducting magnets were cooled by immersion in a helium bath. As we will see later, classical stability theory as derived for these bath-cooled magnets does not directly extend

to forced flow conductors. It is nonetheless useful to review here the oldest and simplest stability criterion developed for bath-cooled conductors, the so-called Stekly criterion of cryostability (1). In their original development, Stekly and Zar (1) backed a superconducting material with a low-resistance copper shunt. The cross section  $A_{cu}$  of the shunt was such that in the case of a transition of the superconducting material to the normal state, the maximum Joule heating, obtained when the current  $I$  was completely displaced from the superconductor to the copper, was smaller than the heat removal capability at the conductor perimeter  $p_c$ , wetted by the helium. Under this condition the conductor always recovered from a perturbation, irrespective of the size of the disturbance that caused the quench. In brief, the conductor was unconditionally stable. Writing the simple power balance of heating and cooling, they came to the following criterion for cryostability, formulated using the so-called Stekly parameter  $\alpha$ :

$$\alpha = \frac{\rho_{cu} I^2}{h p_c A_{cu} (T_c - T_{op})} < 1 \quad (1)$$

where  $\rho_{cu}$  is the stabilizer resistivity,  $h$  the heat transfer coefficient between conductor and cooling bath,  $T_c$  is the critical temperature, and  $T_{op}$  the bath operating temperature.

#### Stability versus Perturbation Spectrum

Cryostable conductors have an exceptional tolerance to energy inputs. The drawback is that the resulting operating current density is low, and thus coil size and cost are large. The present approach is different, and consists in designing the conductor to be stable against the spectrum of energy disturbances expected in the magnet, instead of requiring the conductor to be stable against disturbances of arbitrary nature and intensity. This implies a comparison of the initial estimate of the energy release mechanisms and magnitude to the so-called energy margin  $\Delta E$  that we define as the maximum energy deposition that the conductor can tolerate while recovering from the superconducting state. Let us take as an example a conductor in which mechanical energy releases dominate. Following Kelin (2,3), we compare schematically in Fig. 1 the stability margin to the disturbances that can potentially drive the conductor normal. As shown there, the stability margin  $\Delta E$  decreases with increasing current  $I$ , while mechanical disturbances (the curves labeled "D") increase with the current. In this case, the conductor is no longer unconditionally stable but it has an upper

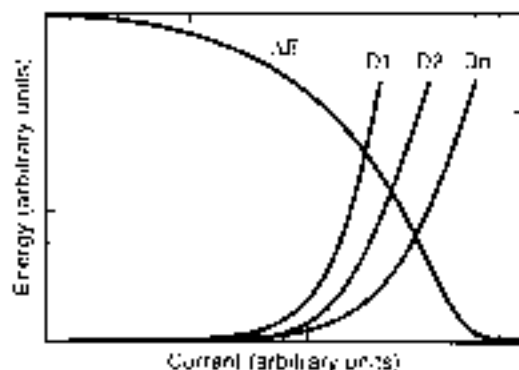


Figure 1. Typical curves for the conductor stability margin ( $\Delta E$ ) and mechanical disturbance spectrum (D).

stability margin, traditionally expressed in  $\text{mJ}/\text{cm}^3$  of metal in the cable. The disturbance spectrum must be interpreted as the energy released in each event. At increasing current a single event releases an increasing energy because of the proportionality to either the Lorentz forces ( $I^2$ ) or to the strain energy in the cable ( $I^4$ ). As the magnet is charged, the two curves approach, until eventually the spectrum of mechanical disturbances (D) equals and surpasses the stability margin ( $\Delta E$ ). At this point the magnet will quench as soon as a perturbation event will take place. The most likely event during magnet charge-up will be the one associated with the largest energy release. Most mechanical energy inputs are associated with irreversible processes such as stick-and-slip motions and cable compaction. Therefore, once this event has taken place, and the associated energy has been released, the energy perturbation spectrum at the following charge-up will be diminished. We illustrate this situation in Fig. 1 by the set of curves D1, D2, D3 that represent the perturbation spectrum at successive charge-ups 1, 2, ...  $n$ . The intersection of these curves with the energy margin curves towards higher currents, and we see from this elementary example a simple explanation of the phenomenon of *training* that disappointed early builders of superconducting magnets.

We see from this simple example that we have two possibilities to guarantee the stable performance of the conductor. The first is to decrease the energy perturbations (movements, cracks, ac losses) as much as possible so that the highest possible operating current can be achieved. This solution can be adopted for small- to medium-scale magnets operating in a quiet environment where, for instance, the perturbation energy input can be limited by properly fixing the cable in the winding pack. On the other hand, large size magnets, as typical of SMES systems, thermonuclear fusion experiments or MHD applications, operate in a mechanic and electromagnetic noisy environment (e.g., rapidly changing magnetic fields or large stress cycles) that per force results in a minimum value of the perturbation spectrum. In this case, the designer must increase the stability margin to tolerate the existing perturbation spectrum. This can be achieved by increasing the heat sink associated with the cable.

In the temperature range typical of the operation of a superconducting cable, generally from 2 to 4 K, all solid materials are known to have a very small heat capacity. In the same temperature range, helium uniquely possesses a volumetric heat capacity two to three orders of magnitude larger than solids. Naturally, cable designers tend to take advantage of this feature, trying to make an effective use of the heat sink provided by adding a limited amount of helium to the cable. To achieve this, it is necessary to increase the heat transfer coefficient at the wetted surface of the conductor, at the conductor surface to volume ratio, or both. Forced-flow-cooled conductors are designed along this line to make the most effective use of the helium heat sink.

Stability in forced-flow-cooled conductors is different from classical stability theory in adiabatic and bath-cooled wires, cables, or built-up manifolds mainly for three reasons:

- The largest heat sink providing the energy margin is the helium, and not the enthalpy of the strands themselves or conduction at the end of the heated length.
- This heat sink is limited in amount.

- The helium behaves as a compressible fluid under energy inputs from the strands, implying additional feedback on the heat transfer coefficient through heating induced flow.

The main issue is, therefore, the heat transfer from the strand surface to the helium flow and the thermodynamic process in the limited helium inventory.

## FORCED-FLOW-COOLED CONDUCTORS

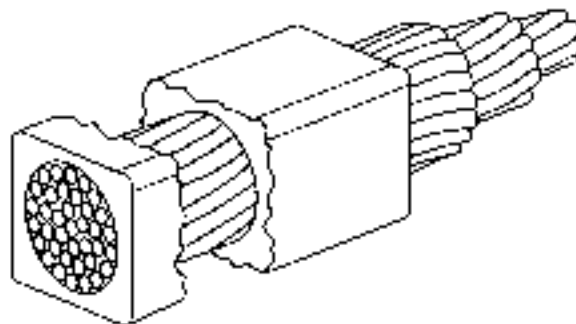
In a forced-flow conductor, the helium and the superconductor form a single unit, with the coolant flowing inside a pipe also housing the superconductor, or with parallel cooling and electrical paths in close thermal contact. The most common design for this class of conductors is at present the cable-in-conduit conductor (CICC), in which a superconducting cable is placed inside a conduit that also serves as helium containment. We will use this particular type of conductor as a prototype for the general discussion on stability in forced-flow-cooled conductors.

### History of Cable-in-Conduit Conductors

The CICC concept evolved from the internally cooled superconductors (ICS), which had found application in magnets of considerable size between the late 60s and early 70s (see in particular the work of Merzjurgs (4)). In ICS, the helium was all contained in the cooling pipe, very much like standard water-cooled copper conductors. The conductor could be wound and insulated using standard technology and the magnet would be stiff both mechanically and electrically, a considerable advantage for medium and large systems requiring, with increasing stored energy, high discharge voltages. Control of the heat transfer and cooling conditions was achieved using supercritical helium, thus avoiding the uncertainties related to a flowing two-phase fluid.

A major drawback of this concept was that according to heat transfer predictions, a large helium massflow would have been required in order to achieve good stability and thus high operating current density. This would require large pumping work and eventually impair the efficiency of the cryogenic system. Chester (5) readily recognized the advantage of the increase in the wetted perimeter obtained by subdivision of the strands. Subdivision dramatically increases the surface-to-volume ratio, thus improving heat transfer for a given cable cross section. Hoenig et al. (6–8) and Dresner (9–11) developed models for the local recovery of ICSs after a sudden perturbation, where they found that for a given stability margin the mass flow required would be proportional to the 1.5th power of the hydraulic diameter as the fixed superconductor inventory is divided in finer and finer strands. This consideration finally brought Hoenig, Iwasa, and Montgomery (6,7) to present the first CICC prototype idea, shown in Fig. 2.

Although many variants have been considered, the basic CICC geometry has changed little since. A bundle conductor is obtained, cabling superconducting strands, with a typical diameter in the millimeter range, in several stages. The bundle is then *jacketed*, that is, inserted into a helium-tight conduit, which provides structural support. Helium occupies the interstitial spaces of the cable. With the cable void fractions of about 30 to 40% commonly achieved, the channels have an effective hydraulic diameter of the order of the strand diameter, while the wetted surface is proportional to the product of the strand



Round bundle with 37 strands enclosed in rectangular conduit. Showing transposition of strands.

**Figure 2.** The original concept of CICC, as presented by Hoenig, Iwasa, and Montgomery (7). Reproduced from Ref. 7 by permission of Servizio Documentazione CRE ENEA Frascati. Copyright 1978 CRE ENEA Frascati.

diameter and the number of strands. The small hydraulic diameter ensures a high turbulence, while the large wetted surface achieves high heat transfer, so that their combination gives excellent heat transfer properties.

Because of the limited helium inventory, a sufficiently large energy input will always cause a quench in a CICC. This behavior has been defined by Dresner (11,12) as *metastable*. The question is the magnitude of the minimum energy input producing a quench in a particular operating condition. This parameter, the *stability margin*  $\Delta E$ , was identified by Hoenig (7) as fundamental to the design of a CICC. It is usually measured as an energy per unit strand volume (traditionally expressed in  $\text{mJ}/\text{cm}^3$ ). In its original definition, the energy input was thought to happen suddenly, and initial experiments and theory concentrated on this assumption. Throughout this chapter we will use the same definition of the stability margin, extending it to an arbitrary energy deposition time scale.

The heat transfer mechanisms determining stability in supercritical He-I and superfluid He-II are different enough to warrant a separate treatment. The phenomenology of each case and the experimental data supporting the stability calculations are presented in the next sections.

## STABILITY MARGIN OF CABLE-IN-CONDUIT CONDUCTORS IN SUPERCRITICAL HELIUM

### Dependence on the Mass Flow

Measurements of the stability margin of CICC's in supercritical helium started early in their history (18–19). One of the aims was to study the dependence of stability on the coolant flow, to determine the influence of the turbulent heat transfer coefficient and the thermal coupling of strands and helium. The first surprise came with the observation by Hoenig (3) that the stability margin was largely independent of the operating mass flow (see the curves reported in Fig. 3), a result soon duplicated by Lue and Miller (17). These results showed that the heat transfer at the wetted surface of the strands during a temperature excursion was only weakly correlated to the steady state mass flow and the associated boundary layer.

As discussed by Dresner (20) and Hoenig (16), during a strong thermal transient the heat transfer coefficient  $h$  at the

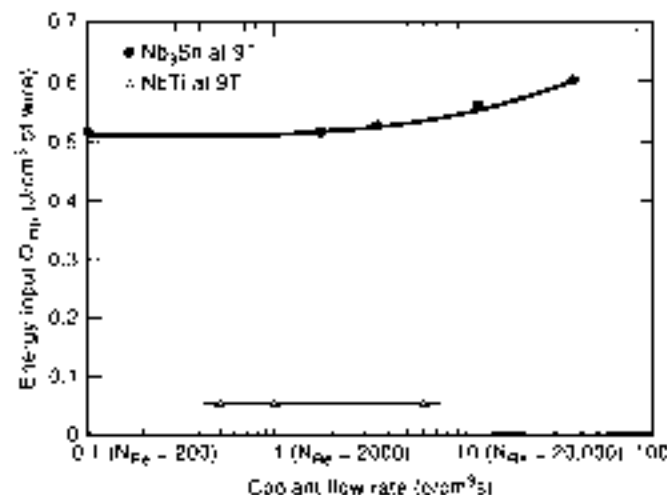


Figure 3. Stability margin of a NbTi and a Nb3Sn CICC as a function of the steady state helium flow measured by Hoemig, Montgomery, and Waldman (14). Reproduced from Ref. 14 by permission of IEEE. Copyright 1999 IEEE.

strand surface changes mainly because of two reasons (see also appendix, Transient Heat Transfer, below)

1. thermal diffusion in the boundary layer (a new thermal boundary layer is developed and thus  $h$  increases compared to the steady state value), and
2. induced flow (31) in the heated compressible helium (associated with increased turbulence and thus again with an increase in  $h$ ).

The concurrence of these two effects was advocated to explain the weak dependence of  $\Delta E$  on the steady mass flow rate.

**Dependence on the Operating Current.** A second parameter of major interest in the experiments on stability was the operating current of the cable. Several experiments (see the vast amount of data presented in Refs. 22 through 27) have revealed a characteristic behavior of the stability margin as a function of operating current. As we show schematically in Fig. 4, at low operating current a region with high stability margin is observed. We name this region, following Schultz and Minervini (28), the *well-cooled regime*. In this regime, the stability margin is comparable to the total heat capacity available in the cross section of the CICC, including both strands and helium, between operating temperature  $T_{op}$  and current-sharing temperature  $T_{cs}$ . At increasing current, a fall in the stability margin to low values, the *ill-cooled regime*, is found. In this regime, the stability margin is lower than in the well-cooled regime by typically one to two orders of magnitude, and depends on the type and duration of the energy perturbation.

The transition between the two regimes was identified by Dresner (20) to be at a *limiting operating current*  $I_{lim}$ :

$$I_{lim} = \sqrt{\frac{h \rho_w A_{cs} (T_{cs} - T_{op})}{\beta \rho_w}} \quad (2)$$

The above definition of the limiting current  $I_{lim}$  is obtained equating the Joule heat generation to the removal at the strand

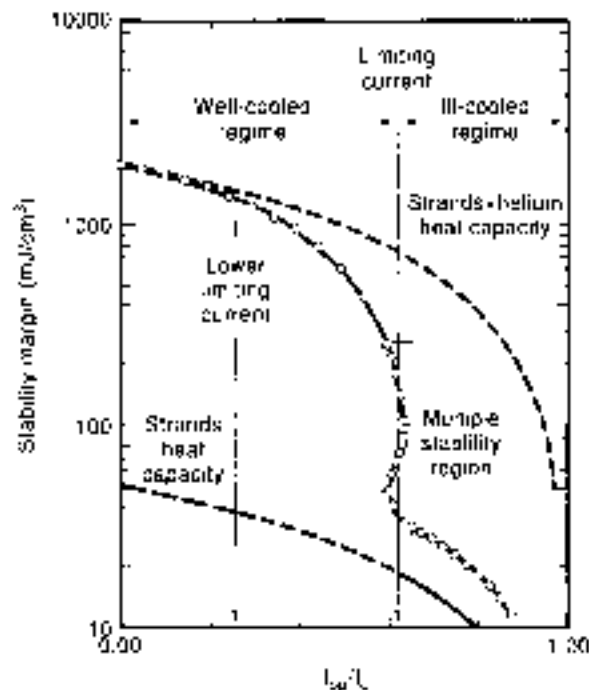


Figure 4. Schematic behavior of the stability margin as a function of the cable operating current.

surface, assuming that the helium temperature is constant, and is therefore equivalent to the Stekly criterion of Eq. (1). As discussed later, the heat transfer coefficient  $h$  is not constant, but it is a complex function of time and several other parameters such as heating pulse waveform and strength, heating induced flow, and details of the cable design. Let us assume for the moment that the heat transfer coefficient is constant in time and equal to an *effective value*. As shown by Luo (29), it is possible to estimate this effective value of  $h$ , deducing it from the location of the limiting current in several experiments. For operating currents below  $I_{lim}$  (i.e., in the well-cooled regime), the heat generation is smaller than the heat removal to the helium. A normal zone recovers, provided that the helium is a sufficiently large heat sink capable of absorbing the heat pulse and the subsequent Joule heating. On the other hand, above  $I_{lim}$ , in the ill-cooled regime, a normal zone generates more heat than it can exchange to the helium, and therefore recovery is not possible.

This observation indeed explains the behavior of the energy margin below and above  $I_{lim}$ . In the well-cooled regime, recovery is unconditional: the cable can transfer a large heat pulse to the helium and still recover at the end of the pulse, provided that the helium temperature has not increased above  $T_{cs}$ . Therefore, the energy margin is of the order of the total heat sink in the cable cross section between the operating temperature  $T_{op}$  and  $T_{cs}$ , including, obviously, the helium. In the ill-cooled regime an unstable situation is reached as soon as the strands are current sharing, and therefore the energy margin is of the order of the heat capacity of the strands between  $T_{op}$  and  $T_{cs}$ , plus the energy that can be transferred to the helium during the pulse. As mentioned earlier, in practical cases, the heat capacity of the helium in the cross section of a CICC is the dominant heat sink by two orders of magnitude and more, and this explains the fall in the stability margin above  $I_{lim}$ .

The transition between the well-cooled and ill-cooled regimes happens in reality as a gradual fall from the maximum

heat sink values to the lower limit [Miller, 1951]. An intuitive explanation of this fall can be given using again the power balance at the strand surface. For the derivation of Eq. (2) it was assumed that the helium has a constant temperature  $T_{he}$ . In reality, during the transient, the helium temperature must increase as energy is absorbed and power is transferred under a reduced temperature difference between strand and helium. Two limiting cases can be defined. The first is the ideal condition of helium at constant temperature, giving the limiting current of Eq. (2), for which, however, the energy absorption in the helium is negligible. Operation exactly at  $I_{lim}$  results thus in a stability margin at the lower limit—the ill-cooled value. The second limiting case is found when the Joule heat production can be removed even when the helium temperature has increased up to  $T_{co}$ . This second case is obtained for a current of (and below)

$$I_{lim}^{co} = \sqrt{\frac{h \rho_w A_c (T_{co} - T_c)}{\rho_c}} \quad (3)$$

that we call *lower limiting current* for analogy to Eq. (2) and because  $I_{lim}^{co}$  is always smaller than  $I_{lim}$ . For operation at (and below)  $I_{lim}^{co}$ , the full heat sink can be used for stabilization and the stability margin is at the upper limit—the well-cooled value. Between the two values  $I_{lim}$  and  $I_{lim}^{co}$ , the stability margin falls gradually.

#### Multiple Stability

Near the limiting current the balance between heat production and removal becomes critical. Indeed, in some cases, a multi-valued region can be found in the vicinity of  $I_{lim}$ , as schematically shown in Fig. 4. As mentioned earlier, supercritical helium behaves as a compressible fluid in the typical range of operation of a magnet. Therefore, any heat pulse causes a heating-induced flow driven by the fluid expansion and proportional to the pulse power. The flow in turn modifies the heat transfer at the wetted surface of the conductor, enhancing the heat transfer coefficient. Let us concentrate on the close vicinity of the limiting current, just above  $I_{lim}$  on the ill-cooled side. A large heating power, above the ill-cooled stability margin, can result in a significant heating-induced flow and thus a large enhancement of the heat transfer coefficient. Hence the power balance can be tipped in the direction favorable to recovery, and a second stable region appears. This is what has been observed by Lee, Miller and Dresner [18,19] in experiments on single triplex NbTi cables. Figure 5 reports one such multiple stability curve, as measured as a function of the operating current. This situation is evidently not agreeable for reliable operation and should be avoided in a sound design by remaining safely below the limiting current.

#### Dependence on the Operating Field

The stability margin as a function of the background field  $B$  mainly through the dependence on critical and current-sharing temperatures. A higher  $B$  causes a drop both in the limiting current (through a decrease of  $T_c$  and increase of  $\rho_{cs}$ ) and in the energy margin (through a decrease in  $T_{cs}$ ). Therefore,  $\Delta E$  drops as the field increases. An interesting feature is that the limiting current only decreases with  $(T_c - T_{cs})^{1/2}$ , that is, with a dependence on  $B$  weaker than that of the critical current. At

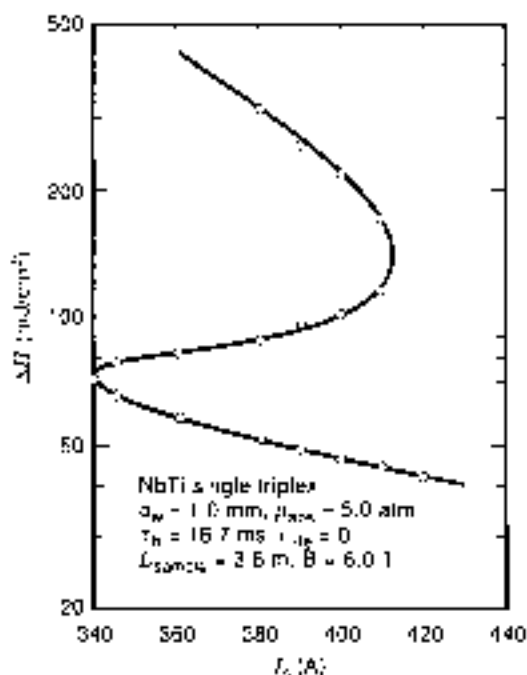


Figure 5. Stability margin of a NbTi CICC as a function of the operating current, measured by Lee and Miller [19]. The experiment was performed on a single triplex CICC of 3.6 m length ( $L_{cable}$ ), with strand diameter of 1 mm ( $d_w$ ), under zero imposed flow ( $v_0 = 0$ ), at a helium pressure of 5 bar ( $p_{helium}$ ). The background field was 6 T ( $B$ ), and resistive heating took place in 16.7 ms ( $\tau_h$ ). Reproduced from Ref. 19 by permission of IERX. Copyright 1981 IERX.

large enough  $B$  we will always have that  $I_{lim}$  is larger than  $I_c$  and the cable will reach the critical current while still under well-cooled conditions.

**Dependence on the Heating Time Scale.** The stability margin depends on the duration of the heating pulse, as shown experimentally by Miller et al. [17], and reported here in Fig. 6. A change in the heating duration for a given energy input corresponds to a change in the pulse power. In the well-cooled regime, that is, for low operating currents in Fig. 6, the heat balance at the end of the pulse is in any case favorable to recovery, and therefore the energy margin does not show any significant dependence on the pulse duration. On the other hand, when the conductor is in the ill-cooled regime, its temperature can increase to or slightly beyond  $T_{cs}$  without causing a quench. This limits the heat flux per unit length at the wetted surface to roughly  $h \rho_w (T_{cs} - T_{he})$ . The consequence is that energy transferred to the helium, and thus the energy margin, will grow at increasing pulse duration, until it becomes comparable to the total heat capacity available (as in the well-cooled regime). This effect is partially balanced for very fast pulses, because the heat transfer coefficient can exhibit very high values at early times (see the appendix, Transient Heat Transfer, below), which would shift the well-cooled/ill-cooled transition at higher transport currents. In principle, higher energy margins should be expected in this range. However, the high input powers in this duration range tend to heat the conductor above 20 K, in a temperature range where the stabilizer resistivity grows quickly, and the power balance is thus strongly influenced. This effect causes the saturation of the energy margin for extremely fast pulses (well below 1 ms duration).

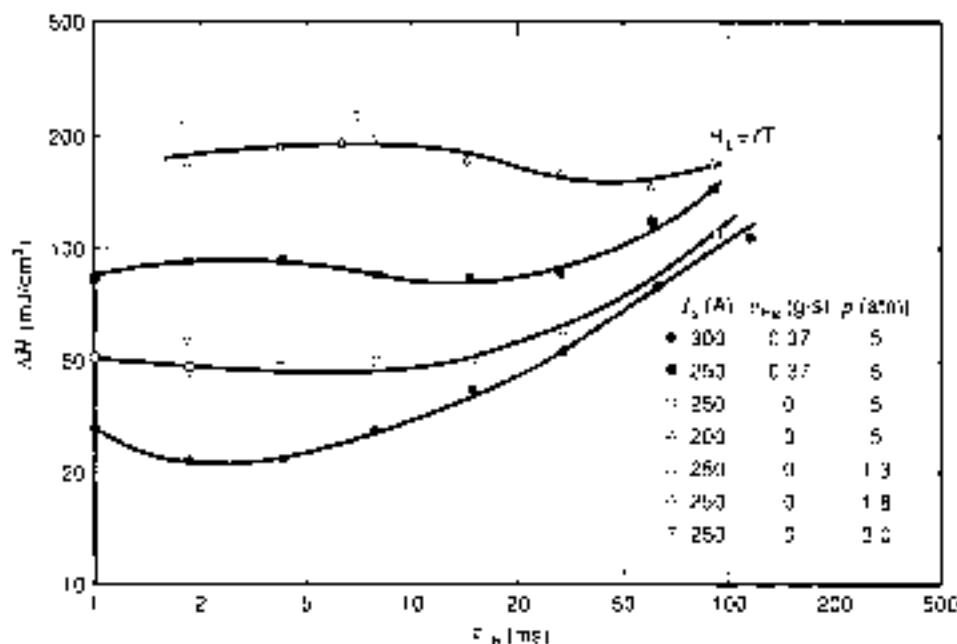


Figure 8. Dependence of the stability margin for a CICC (indicated on this plot as  $\Delta H$ ) on the heating time scale  $\tau$ , as measured by Miller et al. (17). The parameters varied in the experiment, indicated in the inset, are the transport current in the sample  $I_s$ , the helium flow velocity  $v_{He}$ , and the pressure  $p$ . Reproduced from Ref. 17 by permission of IEEE. Copyright 1979 IEEE.

#### Dependence on Operating Temperature and Pressure

The dependence on the operating temperature and pressure in supercritical conditions is not easily quantified. The reason is that the helium heat capacity in the vicinity of the usual regimes of operation (operating pressure  $p_{op}$  of the order of 3 to 10 bar and operating temperature  $T_{op}$  around 4 to 6 K) varies strongly with both  $p_{op}$  and  $T_{op}$ . This affects both the heat sink and the heat transfer coefficient (through its transient components). An increasing temperature margin under constant operating pressure gives a higher  $\Delta E$ . But a simultaneous variation of  $p_{op}$  and  $T_{op}$  under a constant temperature margin can produce variations of  $\Delta E$  as large as a factor two in the range given above [see Miller (25) and Chaniotakis (30)].

#### STABILITY MARGIN OF CABLE-IN-CONDUIT CONDUCTORS IN SUPERFLUID HELIUM

If the operating temperature is lowered below the so-called *lambda* value  $T_\lambda$  (e.g., 2.17 K at 1 atm), helium undergoes a state change and becomes a quantum fluid: *superfluid* helium, or He-II. He-II has unique properties and its physical behavior is very different from that of *normal* helium (or He-I). For our purposes, the most remarkable fact is that He-II does not obey the traditional Fourier law of conduction (proportionality between heat flux and temperature gradient), but rather follows a nonlinear law of the form

$$q = -K(\nabla T)^2 \quad (4)$$

where  $q$  is the heat flux,  $K$  is a parameter that depends on the thermodynamic state of the He-II, and  $\nabla T$  is the temperature gradient. For heat fluxes of practical interest, the He-II properties are such that the "equivalent thermal conductivity"

is extremely high, orders of magnitude higher than in He-I. "Superfluid heat conduction" refers to the ability of stagnant He-II to provide significant heat removal over long lengths of narrow channels without an appreciable temperature gradient, making it an attractive alternative for superconducting magnet cooling (31,32).

Similarly to the behavior in He-I, the stability margin of a CICC operating in He-II is determined by the balance between Joule heating and the ability of the helium to provide enough enthalpy margin, given that metal-to-helium heat transfer is sufficient. Lotan and Miller (27) have measured stability margin of a CICC at different operating temperatures, both in supercritical and superfluid helium. We show typical results of this experiment in Fig. 7. The stability margin behaves at low current in a way similar to what would be expected in the case of He-I operation. However, at the ill-cooled transition—the first drop in the stability curve taking place at similar currents both in He-I and He-II—the stability margin shows a peculiar behavior (owing to the large heat transfer capability in He-II, the power balance at the strand surface remains favorable for recovery as long as the bulk helium is in the He-II phase). Therefore, in a first approximation, the full heat sink between the initial operating point and the transition temperature  $T_L$  is still available at levels of the operating current at which the conductor would have turned to be ill-cooled for operation in He-I. In other words, the conductor can still be considered as *well-cooled* for temperature excursions up to  $T_L$ . We can call this region of operation the *superfluid* stability regime. As the helium undergoes a phase transition at the temperature  $T_L$ , the available heat sink is significant, of the order of 200 mJ/cm<sup>3</sup> of helium volume. At increasing current, finally, the power balance can eventually become unfavorable, as soon as the heat removal capability of He-II reaches its upper limit. There the final transition to the ill-cooled regime of operation takes place. Unlike

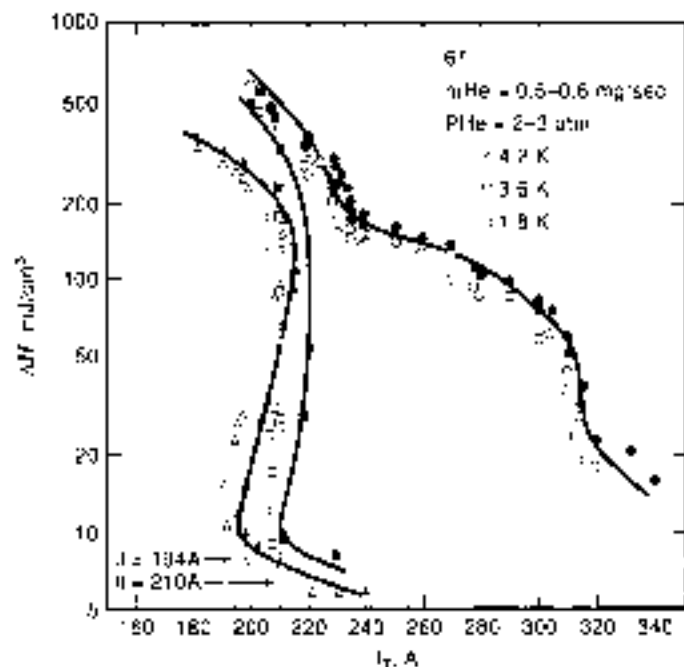


Figure 7. Stability margin of a NbTi CICC as a function of the operating current, measured by Laton and Miller (27), at different temperatures in supercritical and superfluid helium (filled in symbols are quenches, open symbols are recoveries). Reproduced from Ref. 27 by permission of IEEE. Copyright 1989 IEEE.

the He-I case, the transition from high to low stability-margin regimes occurs gradually without a region of multivalued stability owing, on the one hand, to a small coefficient of thermal contraction (compared to the large thermal expansion in He-I), and, on the other hand, to a very high effective thermal conductivity of He-II, both preventing the onset of large heating induced flows.

In the superfluid well-cooled stability regime, where the conductor takes advantage of the superfluid helium properties only, the stability margin has a distinct behavior (33-37): that, as indicated by Seyfert, Lafframandier, and Claudet (33), is mainly determined by the near-field effects of surface heat transfer to He-II. Following a disturbance, the metal temperature exceeds  $T_c$  almost immediately, and the helium in contact with the strand undergoes the transition to He-I. However, at some distance from the strand surface, there is an He-I/He-II transition where the temperature is  $T_c$ . It can be shown that, given the heat transfer properties of He-II, the layer of He-I adjacent to the strand is negligibly thin (33). The consequence is that, for all practical purposes, heat transfer in the superfluid bath takes place "as if" the strand surface had its temperature clamped at  $T_c$ . This is indeed the basis for the simplified model used to calculate the heat flux limits in He-II and the behavior of the stability margin in the superfluid region.

#### CALCULATION OF THE STABILITY MARGIN IN HE-I

The calculation of the stability margin in a CICC is a difficult task, involving accurate computation of compressible helium flow and heat diffusion in a complex geometry. For practical purposes, several simplified models have been developed. These models make extensive use of the experimental evidence discussed in the previous sections as a basis for introducing

and justifying several simplifications. For the purpose of introducing the reader to the concepts involved in the calculation of the energy margin, this presentation will start with a very simplified model (an integrated energy balance), and then proceed to introduce refinements to the model (a zero-dimensional (0-D) energy balance model, and a one dimensional (1-D) flow model).

#### Energy Balance

The simplest stability model is that considering the energy balance for the combined helium/strand system integrated over the duration of the disturbance. This method gives a rough estimate of the stability margin in the well- and ill-cooled regimes (called here  $\Delta E_w$  and  $\Delta E_i$ ) based on the available heat capacities and the location of the well-cooled/ill-cooled boundary (neglecting the dual-stability region), and has the advantage of producing easily applicable design criteria for the selection of the cable layout. We introduce the maximum heat sink in the cable cross section (referred to the unit strand volume):  $\Delta E_{max}$ :

$$\Delta E_{max} = \int_{T_c}^{T_0} \frac{A_{He}}{A_{cs}} C_{He} dT + \int_{T_c}^{T_0} C_s dT \quad (5)$$

where  $C_{He}$  and  $C_s$  are respectively the volumetric heat capacity of helium and of the strands, while  $A_{He}$  and  $A_{cs}$  are their cross sections. The estimate of  $C_{He}$  can pose some questions. As known from thermodynamics, the volumetric heat capacity in a compressible fluid depends on the process assumed. Two limiting cases can be identified: a process at constant volume, where we have that  $C_{He} = \rho c_v$ ; or the case of constant pressure, where we have that  $C_{He} = \rho c_p / \rho_v$  and  $c_p$  represent the helium constant-volume and constant-pressure-specific heat. The proper selection depends on the comparison of the characteristic times involved. In a transient where the flow characteristic times are much longer than the heating and recovery time (i.e., for long-heated zone or fast-heating pulse), the process will be at constant volume. Approximate constant pressure conditions will be found when the flow characteristic times are much shorter than the heating time (short heated zone or long pulse). The real process will be between these two extremes, but generally a conservative estimate is obtained choosing the constant volume process.

In the well-cooled regime we will have

$$\Delta E_w \leq \Delta E_{max}$$

that is, the energy margin is at most equal to the available heat sink up to  $T_c$ , and in general smaller than  $\Delta E_{max}$ . A first reason is that during the heat pulse  $t_p$  and the recovery time  $\tau_r$ , the Joule heat generated by the current sharing strand consumes the available heat capacity. An approximation of the Joule heat contribution normalized to the strand volume is given by

$$Q_{Joule} = \int_0^{t_p + \tau_r} \frac{I_{cs}^2 P}{A_{cs} A_{cs}} dt \quad (6)$$

This contribution increases at increasing operating current and increasing energy deposition time, although the above approximation tends to give only an upper limit and overestimates the real contribution (the strands are assumed fully normal for the whole transient). Still, for fast and for most common heating

pulses (typically in the 1 to 10 ns range) the term above is small. It is then justifiable to neglect the gradual fall of  $\Delta E_s$  and to take

$$\Delta E_s \approx \Delta E_{max} \quad (7)$$

For operation in the ill-cooled regime at and above  $I_{lim}$ , the energy margin can be approximated as the sum of the strand heat capacity up to  $T_{sc}$  and the energy transferred from the strand to the helium during the heat pulse (36), again expressed per unit of strand volume:

$$\Delta E_{tr} = \int_{T_{sc}}^{T_{st}} C_{p1} dT + \frac{P_{sc}(I_{sc} - T_{sc})}{A_{sc}} \int_0^{\tau} h dt \quad (8)$$

where the second term on the right-hand side is an approximation of the energy transferred to the helium under the assumption that the strands rise instantaneously to  $T_{sc}$  and the helium temperature  $T_{sc}$  does not change significantly. For short energy pulses, the use of Eq. (8) shows that generally  $\Delta E_s \ll \Delta E_{max}$ . The energy margin given by Eq. (8) tends to increase when the energy deposition time  $\tau$  increases, which is consistent with the experimental results quoted earlier. For very long pulses, the power input to the strand can be transferred to the helium without a significant temperature difference. At the limit of long pulse times, the whole heat capacity is used again and we have that  $\Delta E_s \approx \Delta E_{max}$ . In any case, the value of the maximum heat sink  $\Delta E_{max}$  of Eq. (5) remains the absolute upper limit of the stability margin. In summary, Eqs. (7) and (8) give the estimated energy margin respectively below and above the limiting current of Eq. (2).

#### Zero-Dimensional Model

The next level of complexity and accuracy in the calculation of the stability margin consists of introducing time as a variable to capture the distinction between short and long duration pulses while neglecting heated-zone length effects. Maintaining the fundamental distinction between strand and helium temperature, it is possible to write this 0-D balance as follows:

$$A_{sc} C_{p1} \frac{\partial T_{s1}}{\partial t} = \dot{q}_{ext} + \dot{q}_{Joule} - \rho_w h (T_{sc} - T_{h1}) \quad (9a)$$

$$A_H C_{pH} \frac{\partial T_{h1}}{\partial t} = \rho_w h (T_{s1} - T_{h1}) \quad (9b)$$

The rightmost terms in Eqs. (9a) and (9b) represent the thermal coupling of strands (at temperature  $T_{s1}$ ) and helium (at temperature  $T_{h1}$ ) at the wetted perimeter  $\mu_w$  with a heat transfer  $h$ . In Eq. (9a) we have in addition the external and Joule heat sources (per unit conductor length)  $\dot{q}_{ext}$  and  $\dot{q}_{Joule}$ , respectively. The Joule heating can be computed once the critical current dependence on the temperature  $I_c(T)$  is known. Note that an accurate calculation of  $\dot{q}_{Joule}$  is necessary to describe the recovery phase properly. This model is attractive because of its simplicity; it can be solved efficiently and used routinely. It is accurate in describing the local energy balance on the time scale of recovery, but some care must be taken in the selection of the parameters in order to capture flow-related physical features that only a 1-D model can include.

The first parameter to be chosen properly is the volumetric helium heat capacity, as we discussed earlier. The second

parameter that requires care is the heat transfer coefficient, changing in time during the transient. While the boundary layer formation and the associated diffusive component of the heat transfer coefficient can be approximated in a local treatment as a variable thermal resistance between strands and helium, the heating-induced flow and its effect on stability are not amenable to local treatment. An average value for this component is a reasonable choice, but the actual modeling is to a large extent left to empiricism [see Luc (29)]. This is, in fact, one of the research areas on stability margin in CICC's.

The search of the stability margin with the 0-D model is the *virtual* analogue of the experimental technique. A trial-and-error search is done on the energy input, increasing or decreasing it as a function of the quench or recovery result at the end of the transient.

#### One-Dimensional Model

With a typical hydraulic diameter in the millimeter range, the overall helium flow in a CICC can be expected to be one-dimensional, with a good approximation already over flow lengths of the order of 1 m. As the helium flows generally in turbulent regime, the helium temperature is nearly uniform in the cross section of the CICC. Therefore, the temperature gradients in the cable cross section reduce to those across the strand, and are negligibly small. We assume that the current distribution is uniform in the strands. In well-designed CICC's, the current can redistribute over typical lengths of the order of some centimeters in times of the order and below 1 ms. In this case, the heat generation in the CICC cross section during current sharing is also uniform. This is not the case for CICC's with insulated strands or high transverse resistance, where the current redistribution can take several seconds over lengths of several meters. In this case, an homogenized treatment is not appropriate and the stability margin is actually strongly degraded. We will therefore drop this case in the following treatment. As the stability transients are fast compared to the thermal diffusivity of the conduit materials (e.g., steel), the conduit contribution to the energy balance is neglected also.

These assumptions lead to a much simplified 1-D model of the CICC, where two constituents are identified, the helium and the strands. Both are at uniform, but distinct, temperature. The compressible flow equations in the helium (mass, momentum, and energy balances) are written to include wall friction, modeled using a turbulent friction factor. Strand and helium exchange heat at the wetted surface, and the thermal coupling is usually modeled using the correlation for the heat transfer coefficient  $h$  discussed in the Appendix. The system is then described by the equations:

$$\frac{\partial \rho}{\partial t} + \frac{\partial \rho v}{\partial x} = 0 \quad (10a)$$

$$\frac{\partial \rho v}{\partial t} + \frac{\partial \rho v^2}{\partial x} + \frac{\partial p}{\partial x} = 2 \rho f \frac{v^2}{D_h} \quad (10b)$$

$$\frac{\partial \rho e}{\partial t} + \frac{\partial \rho v e}{\partial x} + \frac{\partial p v}{\partial x} = \frac{\mu_w h}{A_{sc}} (T_{s1} - T_{h1}) \quad (10c)$$

$$A_{sc} C_{p1} \frac{\partial T_{s1}}{\partial t} = A_{s1} \frac{\partial}{\partial x} \left( K_{sc} \frac{\partial T_{s1}}{\partial x} \right) \\ = \dot{q}_{ext} + \dot{q}_{Joule} - \rho_w h (T_{s1} - T_{h1}) \quad (10d)$$



where  $\rho$  is the helium density,  $p$  its pressure and  $u$  is the flow velocity,  $f$  the friction factor and  $D_h$  is the hydraulic diameter of the conductor. The total specific energy  $e$  is defined as the sum of the internal specific energy  $i$  and the kinetic specific energy, that is,

$$e = i + \frac{u^2}{2}$$

Finally, the strand heat balance of Eq. (10d) takes into proper account the contribution of the heat conductivity  $K_{eff}$  along the cable length.

The 1-D model introduced above is widely used for detailed calculations of stability margin. When the numerical solution technique to account for the different time scales involved is properly selected, the model can predict the heating-induced flows responsible for multivalued stability, and can be adapted directly to follow the evolution of the normal zone when the energy input is large enough and the coil quenches. The only significant modification in this case is the need to take into account the additional heat capacity of the conduit material. This modification is straightforward and consists of adding a temperature diffusion equation to the system. Because of the level of fine details, even within the simplification of the 1-D assumption, this model gives the possibility of wide parametric analysis. Its main drawback is that, dealing with largely different time scales, it is slow and not easy to handle.

#### CALCULATION OF STABILITY MARGIN IN HE-II

A CICC operating in superfluid helium is most efficient if it is designed to take advantage of the large heat removal capability at the strand wetted surface. As we discussed in the description of the general features of the stability margin in He-II, it is possible to operate the cable at a current density significantly higher than in the case of supercritical helium. However, we recall that in the *superfluid* well-cooled regime, the upper stability margin is determined by the helium enthalpy available between operating temperature and the lambda point, of the order of 200 mJ/cm<sup>3</sup> of helium volume. Therefore, in the design for operation in He-II, we assume implicitly that if the Joule heating is enough to drive the bulk temperature above  $T_\lambda$ , cooling in He-I is so reduced in relation to superfluid cooling that the conductor will not recover.

As for operation in supercritical He-I, the stability margin of a CICC in He-II can be computed at different levels of approximation and complexity. We need in this case to modify the heat transfer coefficient (see the Appendix) and the helium energy balance to take into account the equivalent thermal conductivity given by Eq. (4). However, the model used in practice to design CICCs for stable operation in He-II is different from that discussed in the previous section. In this case, we concentrate on the heat removal capability, with the aim of maintaining it sufficiently high so that the full helium heat sink up to the lambda point is available for stabilization.

The essence of the simplified model, due to Dresner (36,37), consists in solving the 1-D heat transport equation in stagnant He-II in a channel in order to obtain the effective cooling capacity. The model implies that even if the helium goes through the lambda transition at the conductor surface, the bulk of the coolant remains in the superfluid state and cooling is deter-

mined by "conduction" along the channel (in the case of a CICC, the "channel" is the imaginary annulus of helium surrounding each strand). This is consistent with Seyler's observation that the layer of He-I around each strand is negligibly thin (33). Dresner's model is based on the analytical solution of the nonlinear heat "conduction" in the annulus of He-II around the strands, and the ability of the helium to absorb the heat flux stemming from the Joule heating (for details of the derivation see Refs. 36 and 37). In that context, it is useful to define the following two quantities:

$$E = \frac{\Delta c_p}{\rho_w} \Delta \mathcal{E} \quad (11a)$$

$$\mathcal{E} = [h_{10} T_o - h_{10} T_\lambda] L \quad (11b)$$

where the quantity  $E$  is just a different scaling of the stability margin (the quantity of interest in this calculation),  $\Delta c_p$  is the helium enthalpy per unit volume, and  $L$  is the effective "channel" length, so that  $\mathcal{E}$  represents the maximum enthalpy available between the operating temperature and the lambda temperature in the annulus of helium around each strand (the total heat sink). The equivalent channel length  $L$  depends on the helium cross section as follows,

$$\rho_w L = A_{He} = \frac{f_{10}}{1 - f_{10}} A_{cs} \quad (11c)$$

where  $f_{10}$  is the CICC void fraction. The analytical solution of the Dresner model leads to a relationship between the stability margin and the design current density that can be expressed using the nondimensional groups  $\mathcal{E}/K_c$  and  $q_c/q$ , with the latter terms defined as follows

$$q_c = \frac{\Delta c_p I^2}{\rho_w \Delta c_p} \quad (11d)$$

$$q = \frac{K C_{10}^{-1} T_o - T_w}{14 K_c T_o} T_w^2 \quad (11e)$$

in which  $q_c$  represents the Joule heating in the cable per unit of cooling surface. The quantity  $q$  is a fiducial heat flux that depends on the effective thermal conductivity of superfluid helium ( $K$ ), the volumetric specific heat of the helium ( $C_{10}$ ), the temperature difference between the operating point and the lambda transition, and  $K_c$  (defined above). This quantity stems from the solution of the nonlinear heat conduction equation for He-II. The ratio  $q_c/q$  determines the severity of the Joule heating pulse that needs to be transported and absorbed by the superfluid helium.

The model predicts an expression relating  $\mathcal{E}/K_c$  as a function of  $q_c/q$ , which is shown in Fig. 8. The results of this stability model indicate that, as in He-I operation, there are two distinct regimes, at low currents the conductor is *well-cooled* and essentially the entire helium enthalpy margin (up to the transition to He-I) is available to stabilize the cable; at higher currents the cable becomes *ill-cooled* and the stability margin is significantly lower because of the inability of the He-II to conduct the Joule heat flux over the length of the equivalent "channel."

The criterion used in practice then stems directly from the results of this model. To ensure the highest possible stability margin (the entire helium enthalpy from the operating point to the lambda point), the conductor needs to be designed so that  $q_c$

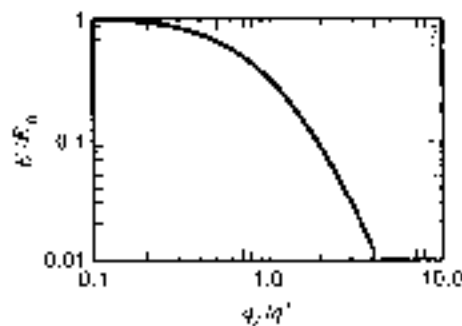


Figure 6. Semiempirical representation of the stability margin for a CICC operating in He-II.

$= 0.2q$ , for at higher fractions of  $q$  the margin decreases rapidly. Note that as a result, the conductor design (and the stability margin) depends only on the physical properties of the He-II and on the current density in the stabilizer. The current sharing temperature does not enter into the determination of stability as a result of assuming that the conductor cannot recover once the lambda transition is crossed during a disturbance. The present model has been successfully used in the design of a large (200 kA) CICC for application in SMES (19,40).

## CONCLUSIONS AND RESEARCH DIRECTIONS

This article has presented the basic considerations and models that go into the design of CICC's for stable operation. Enough is known of the mechanisms determining stability so that, in conjunction with other constraints, CICC's can be designed successfully and even optimized (41). However, this does not mean the field is not open to new areas of research. As new magnet designs are proposed, and as more stringent requirements are imposed on the designer, areas of further study continue to open—in particular, work toward improving our understanding of stability under transient operating conditions (e.g., ac losses and stability) or for more complex cable geometries (e.g., CICC with central channels).

As may be clear after review of the literature, stability depends in a synergistic manner on the dc and ac operating conditions of the cable in the coil (42–44). This is the main direction of the actual research in the field of CICC stability. In particular, in view of the applications to pulsed magnets, the interaction of stability, current distribution, and ac losses in the cable is one of the main topics. The so-called *compromise* limit of operation for pulsed magnets (43–45) (a decrease in the maximum achievable current at increasing field change rate) is an outstanding example of this synergistic interaction. The appearance of such a phenomenon, explained so far in terms of nonuniform current distribution and a degradation of the stability margin of the cable, has alerted us to the difference between dc stability, with constant operating current and background field, and ac stability of the cable.

However, while dc operating conditions are easier to produce and simulate, ac stability is difficult to measure and poses some basic problems in the interpretation of the data. The simulation and prediction of ac stability are therefore object of an intense activity in the field of transient electromagnetics in superconductors and thermohydraulics.

## APPENDIX: TRANSIENT HEAT TRANSFER

The main issue in heat transfer is fast transients, for their relevance to stability. A strong variation of the transient heat transfer to a flow of supercritical helium was demonstrated by Giarratano (46) and Bloem (47) in dedicated measurements on short test sections. The experiments showed an initial peak in the heat transfer coefficient at early times, below 1 ms. At later times, in the range of some ms to about 100 ms, the initial peak decreased approximately with the inverse of the square root of time. The behavior could be explained in terms of the diffusion of heat in the thermal boundary layer. Using the analytical solution of diffusion in a semi-infinite body (the helium due to a heat flux step at the surface), the effective heat transfer coefficient can be computed as [Bloem (47)]:

$$\bar{h}_t = \frac{1}{2\lambda} \frac{\overline{K_{H,0}} \overline{c c_p}}{\sqrt{t}} \quad (12)$$

where  $K_{H,0}$  is the heat conductivity of helium. The expression above is shown to fit properly the experimental data for times longer than 1 ms and until the thermal boundary layer is fully developed. At earlier times, Eq. (12) would tend to predict an exceedingly high heat transfer coefficient, consistent with the assumptions of the analytical calculation. In reality, the early values of  $\bar{h}$  are found to be limited by the Kapitza resistance (48) at the contact surface of strand and helium, which gives a significant contribution only when the transient heat transfer coefficient is in the order of or larger than  $10^4 \text{ W m}^{-2} \text{ K}^{-1}$  (in the case that the wetting helium is in the superfluid state). A suitable expression for the heat transfer coefficient in the Kapitza resistance can be obtained using

$$h_K = 200 T_{H,0}^2 + T_{H,0}^2 (T_{H,0}^2 + T_{H,r}^2) \quad (13)$$

which, in fact, approximates a radiation-like phenomenon at the conductor surface with an equivalent heat transfer coefficient.  $T_{H,0}$  and  $T_{H,r}$  are the strand and the helium temperature, respectively.

At later times, usually around 10 to 100 ms, the thermal boundary layer is fully developed and the steady state value of  $\bar{h}$  is approached. Its value appears to be well approximated by a correlation of the Dittus-Boelter form, as shown by Yaskin (49) and Giarratano (50). A best fit of the available data is obtained with the following expression (neglecting corrections due to large temperature gradients at the wetted surface):

$$h_s = 0.0259 \frac{K_{H,0}}{D} \text{Re}^{0.75} \text{Pr}^{-1} \quad (14)$$

An empirical expression for the heat transfer to supercritical helium during a transient finally can be obtained modeling the Kapitza resistance and the helium boundary layer as series thermal resistances, and taking

$$\bar{h} = \max \left\{ \frac{h_s \bar{h}_t}{h_s + \bar{h}_t}, \frac{\bar{h}_t h_K}{\bar{h}_t + h_K} \right\} \quad (15a)$$

This expression is in good agreement with the experimental results quoted above, and shows that, for short pulses, the heat transfer coefficient only depends on the helium state and not

on the flow conditions. At temperatures below the lambda point (superfluid helium), the Kapitza resistance is the only limit to the heat transfer at the strand-wetted surface. In this case, we approximate the heat transfer coefficient simply as

$$h = h_K \quad (15b)$$

Equations (15a) and (15b) can then be used to approximate the heat transfer coefficient at the strand-helium interface for the entire transient in the supercritical regime, or to cover the entire operating temperature range, from below the lambda point to 4.2 K and above. It should be noted, however, that the expressions above are only a convenient tool to implement in stability models covering most regimes of interest, and do not capture all the complexities encountered, for instance, when operating in atmospheric superfluid, when phase changes may be taking place at the interface during a transient.

## BIBLIOGRAPHY

1. Z. J. J. Stokly and J. L. Zar, Stable superconducting coils, *IEEE Trans. Nuc. Sci.*, **14**, 361, 1965.
2. V. E. Keilin, Experimental features of superconducting windings training by balance of acting and permissible disturbances, *IEEE Trans. Appl. Supercond.*, **3**: 297-300, 1993.
3. V. E. Keilin, An approach to deduce quench fields of superconducting magnets from model coils testings, *IEEE Trans. Magn.*, **30**: 1962-1965, 1994.
4. M. Mungai, The design of the superconducting magnet for the "Omega" project, particle accelerator, vol. 1, Glasgow, UK: Gordon and Breach Publishers (printed by Bell and Bain Ltd.), 1970.
5. P. F. Chester, Superconducting magnets, *Reports on Progress in Physics*, **XXXII**, 561-614, 1967.
6. M. O. Hoernig and D. B. Montgomery, Dense supercritical helium cooled superconductors for large high field stabilized magnets, *IEEE Trans. Magn.*, **11**: 569, 1975.
7. M. O. Hoernig, Y. Iwasa, and D. B. Montgomery, Supercritical-helium cooled "bundle conductors" and their application to large superconducting magnets, *Proc. 6th Int. Conf. Magn. Tech.*, Rome, Italy, 21-25 Apr. 1975, in N. Sacchetti, M. Spadani, and S. Stipcich (eds.), Frascati Lab. Naz. del CERN, 519, 1975.
8. M. O. Hoernig et al., Supercritical helium cooled cabled, superconducting hollow conductors for large high field magnets, *Proc. 6th Int. Cryo. Eng. Conf. Grenoble, France*, 11-14 May 1976, in K. Mundelsohn (ed.), Guildford, Surrey UK: Science and Technology Press, 310, 1976.
9. L. Dresner, Stability-optimized, force-cooled, multifilamentary superconductors, *IEEE Trans. Magn.*, **13**: 670, 1977.
10. L. Dresner and J. W. Lee, Design of forced-cooled conductors for large fusion magnets, *Proc. 7th Symp. on Eng. Probl. of Fus. Res.*, **1**: 263, 1977.
11. L. Dresner, Stability of internally-cooled superconductors: A review, *Cryogenics*, **20**: 555, 1980.
12. L. Dresner, Superconductor stability 1983: A review, *Cryogenics*, **24**: 243, 1984.
13. M. O. Hoernig and D. B. Montgomery, Cryostability experiments of force-cooled superconductors, *Proc. 7th Symp. on Eng. Probl. of Fus. Res.*, **1**: 789, 1977.
14. M. O. Hoernig, D. B. Montgomery, and S. J. Waldman, Cryostability in force-cooled superconducting cables, *IEEE Trans. Magn.*, **16**: 792, 1979.
15. M. O. Hoernig, Internally cooled cabled superconductors—Part I, *Cryogenics*, **20**: 373-399, 1980.
16. M. O. Hoernig, Internally cooled cabled superconductors—Part II, *Cryogenics*, **20**: 427-431, 1980.
17. J. R. Miller et al., Measurements of stability of cabled superconductors cooled by flowing supercritical helium, *IEEE Trans. Magn.*, **16**: 451, 1979.
18. J. W. Lee, J. R. Miller, and L. Dresner, Stability of cable-in-conduit superconductors, *J. Appl. Phys.*, **51**: 772, 1980.
19. J. W. Lee and J. R. Miller, Parametric study of the stability margin of cable-in-conduit superconductors: Experiment, *IEEE Trans. Magn.*, **17**: 1991.
20. L. Dresner, Parametric study of the stability margin of cable-in-conduit superconductors: Theory, *IEEE Trans. Magn.*, **17**: 753, 1981.
21. L. Dresner, Heating induced flow in cable-in-conduit conductors, *Cryogenics*, **19**: 653, 1979.
22. J. R. Miller et al., Stability measurements of a large Nb<sub>3</sub>Sn force-cooled conductor, *Adv. Cryo. Eng.*, **26**: 634, 1990.
23. J. W. Lee and J. R. Miller, Performance of an internally-cooled superconducting solenoid, *Adv. Cryo. Eng.*, **27**: 227, 1982.
24. J. V. Minervini, M. M. Scèveas, and M. O. Hoernig, Experimental determination of stability margin in a 27 strand bronze matrix, Nb<sub>3</sub>Sn cable-in-conduit conductor, *IEEE Trans. Magn.*, **21**: 439, 1985.
25. J. R. Miller, Empirical investigation of factors affecting the stability of cable-in-conduit superconductors, *Cryogenics*, **25**: 552, 1985.
26. T. Ando et al., Investigation of stability in cable-in-conduit conductors with heat pulse duration of 0.1 to 1 ms, *Proc. 10th Int. Cryo. Eng. Conf.*, 756, 1986.
27. J. C. Luttin and J. R. Miller, Stability of internally-cooled superconductors in the temperature range 1.4 to 4.2 K, *IEEE Trans. Magn.*, **19**: 430, 1983.
28. J. H. Schultz and J. V. Minervini, Sensitivity of energy margin and cost figures of internally-cooled cabled superconductors (ICCS) to parametric variations in conductor design, *Proc. 9th Magn. Tech. Conf.*, 643-646, 1986.
29. J. W. Lee, Review of stability experiments on cable-in-conduit conductors, *Cryogenics*, **34**: 779-786, 1994.
30. E. A. Charinakis, Energy margin of cable-in-conduit conductor as a function of operating pressure and initial heated zone, *IEEE Trans. Magn.*, **32**(4): 2960-2969, 1996, also E. A. Charinakis, Effect of operating pressure and heated length on the stability of ICCCs, *J. Fusion Energy*, **14**(1): 61-73, 1995.
31. G. Claudet et al., Superfluid helium for stabilizing superconductors against local disturbances, *IEEE Trans. Magn.*, **15**: 340, 1979.
32. S. W. Van Sever, Stability of superconductors cooled internally by He II heat transfer, *Cryogenics*, **31**: 516, 1991.
33. P. Seyfert, J. Lallierrandere, and G. Claudet, Time-dependent heat transport in subcooled superfluid helium, *Cryogenics*, **22**: 601, 1982.
34. C. Meuris et al., Transient stability of superconductors cooled by superfluid helium at atmospheric pressure, *Proc. Int. Inst. Refrigeration*, 217-223, 1981.
35. C. Meuris, Experimental study of the stability of a superconductor cooled by a limited volume of superfluid helium, *IEEE Trans. Magn.*, **19**: 272, 1983.
36. L. Dresner, A rapid, semiempirical method of calculating the stability margin of superconductors cooled with subcooled He II, *IEEE Trans. Magn.*, **23**: 318-321, 1987.
37. L. Dresner, *Stability of Superconductors*, New York: Plenum Press, 1976.
38. U. Schmidt, Stability of superconductors in rapidly changing magnetic fields, *Cryogenics*, **30**: 501, 1990.

- 39 D. L. Walker et al., SMES conductor design, *IEEE Trans. Magn.*, **26**: 1596, 1989.
- 40 S. D. Peck, J. C. Ziegler, and C. A. Luongo, Tests on a 200 kA Cable-in-Conduit Conductor for SMES Application, *IEEE Trans. Appl. Supercond.*, **4**: 199, 1994.
- 41 L. Bottura, Stability protection and ac loss of cable-in-conduit conductors—A designer's approach, *Fus. Eng. Des.*, **20**: 351–362, 1995.
- 42 E. Tada et al., Downstream effect on stability in cable-in-conduit superconductor, *Cryogenics*, **29**: 830, 1999.
- 43 N. Kozumi et al., Experimental results on instability caused by non-uniform current distribution in the 30 kA NbTi Domo Poloidal Coil (DPC) conductor, *Cryogenics*, **34**: 156, 1994.
- 44 N. Kozumi et al., Current imbalance due to induced circulation current in a large cable-in-conduit superconductor, *Cryogenics*, **36**: 409–413, 1996.
- 45 S. Jennig et al., Ramp rate limitation experiments using a hybrid superconducting cable, *Cryogenics*, **36**: 628–633, 1996.
- 46 P. J. Garrattano and W. G. Steward, Transient forced convection heat transfer to helium during a step in heat flux, *Trans. ASME*, **106**: 350–357, 1983.
- 47 W. B. Dinsen, Transient heat transfer to a forced flow of supercritical helium at 4.2 K, *Cryogenics*, **26**: 300–306, 1986.
- 48 S. Van Siver, *Helium Cryogenics*, New York: Plenum Press, 1979.
- 49 L. A. Yaskin et al., A correlation for heat transfer to superphysical liquid helium in turbulent flow in small channels, *Cryogenics*, **17**: 549–562, 1977.
- 50 P. J. Garrattano, V. D. Arp, and R. V. Smith, Forced convection heat transfer to supercritical helium, *Cryogenics*, **11**: 396–399, 1971.

LUCA BOTTURA  
CEPN  
CESAR LUONGO  
Bechtel

## STABILIZATION AGAINST FLUX JUMPS

### INTRODUCTION

Superconducting conductors used for ac and dc magnets are usually composite conductors consisting of many "intrinsically" stable fine superconducting filaments in a normal metal matrix.

It is well known that a simple twisting of the wire as a whole cannot ensure an equal current distribution between each filament or each layer of filaments. It appears that during energization the current fills first the outer layers of filaments up to the critical current density, whereas hardly any current flows in the inner layers.

It has already been shown that this particular current distribution with the field totally shielded from the interior of the wire could result in instabilities and premature quenching of large multifilamentary wires as it does in a large monofilament. It has already been observed and explained how a highly conductive matrix can slow the process and eventually totally damp the flux jump. Several laboratories have also observed how highly resistive matrix wires could hardly be stable. This was not very surprising and could be explained in agreement with the theories of stability of wires made with many filaments in homogeneous matrix (copper or manganickel).

This work was mainly devoted to superconducting wires made of filaments embedded in a resistive matrix (such as CuNi) to reduce coupling current losses with an outer shell of low resistivity (such as Cu) to provide enough stability. In order to help our understanding, this study was carried out in line with previous work on stability (1–6). It concerns the so-called "intrinsic stability." We have shown how most of the classical criteria (usually unrelated) can be integrated into one unique and general expression. In particular, in the case of superconducting composites with a highly resistive matrix, we investigated how the different physical parameters can be optimized for achieving stabilization.

*Intrinsic stability* of a superconducting composite refers, in this approach, to hampering the development of flux jumps, the origins of which lie in the nonuniform current distribution inside the wire. This means that a wire is intrinsically stable if a "small" perturbation does not lead the conditions to enter in a diverging spiral,  $\Delta T \rightarrow -\Delta J_c \rightarrow \Delta B \rightarrow E \rightarrow Q \rightarrow \Delta T$ , caused by the decrease of critical current  $J_c$  with respect to temperature  $T$ .

The sources of perturbations are mainly of internal origin (change in current, mechanical hysteresis) or, by extension, any "small" change in temperature due to an external source (field change, for instance).

In the last section of this article, a more complete discussion is given about what can be considered to be a "small" perturbation with respect to the actual relationship between electric field and current density. Thus, intrinsic instability is directly ascribed to the nonuniform current distribution inside the wire (a direct consequence of superconductivity), which stores a source of magnetic energy, and the decrease in critical current density  $J_c$  with respect to temperature  $T$ .

Two ways of preventing this phenomenon are known: by absorbing heat generation itself (adiabatic stabilization) and by removing it by using the enthalpy of external materials or coolants (dynamic stabilization). In fact, the second process is different from the first only in that the flux jump has slowed enough to provide time for heat to be transferred to the coolant. This results only in obtaining more "available" enthalpy than the enthalpy of the wire itself, to absorb the magnetic energy stored in the system.

### FUNDAMENTAL EQUATIONS AND CALCULATIONS

For a given transport current  $I_t$  less than the critical current  $I_c$ , the current density and field profile are shown in Fig. 1. (The transport current  $I_t$  and the background field  $B_0$  are kept constant.) The critical current density is a unique function of temperature. For simplicity, the dependence on temperature is taken to be linear:

$$J_c = J_c \frac{T_c - T}{T_c - T_0} \quad \text{for } T \leq T_c$$

$$J_c = 0 \quad \text{for } T > T_c$$

In the frame of a simplified critical model,  $J_c$  is a step function versus electric field:

$$J_c(E = 0) = 0$$

$$J_c(E \neq 0) = J_c$$

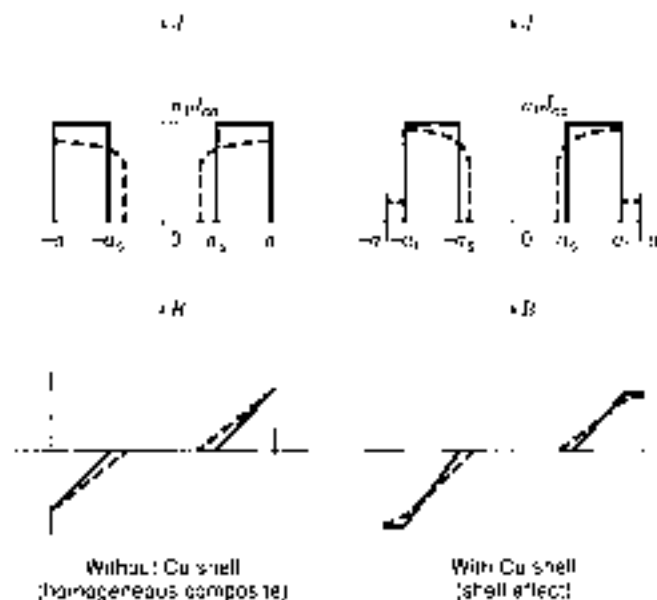


Figure 1. Schematic representation of current distribution and field profiles in a composite before and after a temperature perturbation.

The effects of more realistic expressions of current density, such as  $j_{0i} = j_{0i} E_i / E_i^{1/2}$ , are discussed in the final section.

In previous work (devoted to copper matrix composites), the resistivity of the normal metal matrix was taken into account by the expression (5):

$$E = \rho_e (j - \eta_s j_s)$$

where  $\rho_e$  is the effective or average resistivity of the wire in the critical state,  $j$  is the average current density in the wire, and  $\eta_s$  is the superconductor space factor.

In contrast to the previous work, we now consider the case of a multifilament wire made of two different regions: a central region consisting of the superconducting filaments in a highly resistive matrix (CuNi), for which a justified assumption is made in a later step in the calculations that the resistivity of this central zone is infinite, and an outer normal region with good electrical conductivity (Cu).

Calculations are derived for a slab model. Owing to the self-field effect, for a given transport current  $I$ , the field penetrates to radius  $a$ , as shown in Fig. 1, with

$$j = \frac{I}{l} = \begin{cases} \frac{a}{cr} & r < a \\ 0 & r > a \end{cases}$$

where  $a$  is the limit of the elementary region and  $a$  is the outer "dimension" of the wire ( $a = R\sqrt{2}$  for equivalence with a round wire).

According to Fig. 1, there are three zones of study:

- 0  $\leq x \leq a_1$ ,
- $a_1 \leq x \leq a$
- $a \leq x \leq a_2$

A small temperature perturbation  $dT$  in the composite results in a magnetic flux motion together with an induced electric field. The deviations from the steady state are related by two

electromagnetic equations and one thermodynamic equation in each region.

In the central region (filamentary zone with no current)

$$\begin{aligned} \frac{\partial H}{\partial z} &= 0 \\ j &= 0 \\ C_1 \frac{\partial T}{\partial t} &= \text{div}(K \text{ grad } T) \end{aligned}$$

In the central region carrying the current,

$$\begin{aligned} \text{rot } E &= -\mu_0 \frac{\partial H}{\partial t} \\ \text{rot } \frac{\partial H}{\partial z} &= \eta \frac{\partial j}{\partial t} \\ C_2 \frac{\partial T}{\partial t} &= \eta_1 j_s E - \text{div}(K \text{ grad } T) \end{aligned}$$

In the outer copper shell,

$$\begin{aligned} \text{rot } E &= -\mu_0 \frac{\partial H}{\partial t} \\ \text{rot } H &= j_0 \\ C_3 \frac{\partial T}{\partial t} &= \eta_0 j_0^2 - \text{div}(K \text{ grad } T) \end{aligned}$$

For simplicity, the physical properties of materials are assumed to be independent of temperature. Equivalent and average characteristics are derived for each zone after a homogenization that takes into account the local structure of the composite strand.

In each zone, according to the differential equations, the solution for the temperature can be developed in the form of a sum of terms with time and space, for separate variables:

$$T(x, t) = \sum X_i(x) \exp(\lambda_i t) \quad (1)$$

where

$$\tau = \frac{C_i a_i^2}{K_i} \quad (2)$$

A dimensionless differential equation with respect to space can be derived in the filamentary region:

$$X'' + (1 - \epsilon)X' + \epsilon(\lambda - \mu)X = 0 \quad (3)$$

with the fundamental parameters

$$\begin{aligned} \epsilon &= \frac{\eta_0 \eta_s^2 j_0^2 a^2}{C_1 (T_c - T_0)} \\ \mu &= \frac{K_1 \eta_0}{C_2 \rho_f} \end{aligned} \quad (4, 5)$$

Note that  $\eta_0 a = \eta_s a$  and that

$$\eta = \frac{I_0 c}{D_{sc}}$$

is the ratio of the thermal diffusivity over the magnetic diffusivity.

In composites with a highly resistive matrix,  $\epsilon$  is much less than 1, which allows a further simplification of the differential equation

$$X^{(4)} - \lambda X^{(2)} - \gamma \beta X = 0 \quad (6)$$

In the outer copper shell, if the Joule heating is neglected, the heat equation becomes

$$X'' - \frac{\lambda}{r} X = 0 \quad (7)$$

with

$$\gamma = \frac{K_c C_T}{C_n K_1} \quad (8)$$

For a given constant transport current, there are eight boundary conditions among the three zones concerning the electric field, magnetic field, temperature, flux, and heat transfer to the outer coolant. The general solutions of the differential equations are combinations of eight hyperbolic functions, the coefficients of which can be derived from the set of eight linear equations.

In order to ensure self-consistency and to obtain a nonunique trivial solution for  $\lambda$ , the determinant of the system must be equal to zero. From Eq. (1), it can be inferred that any positive  $\lambda$  value leads to an irrevocable increase of temperature with time. The first positive value for  $\lambda$  determines the limit between stability and instability (see also Refs. 2 and 3).

Our calculations were carried out in the same way as in previous papers (2-7). One of the goals of this work was to find an *analytical* expression for the stability conditions that included the parameter  $\beta$  and other physical characteristics.

As stability is violated for the very first positive  $\lambda$  value, a power-series expansion of the hyperbolic functions can be made. Let us set the dimensionless parameters

$$\gamma = \frac{K_c C_T}{K_1 C_n}, \quad \alpha = \frac{K_c}{K_1}, \quad \delta = \frac{a_1 C_1}{a_2 K_1}, \quad \epsilon = \frac{a_n}{a_r}, \quad h = \frac{h_0 \pi}{K_T}$$

When  $\lambda$  tends to zero, it can be written in the form

$$\lambda = 4R/a_2 \left/ \left[ 4\beta^2 (1 - \delta \epsilon) \left( \frac{R_0^2}{3} - 1 \right) - \alpha \left[ \frac{4\beta \epsilon}{r} \left( 1 - \frac{h_0}{2} \right) + h \left( 1 + \frac{2\beta \epsilon^2}{\delta} + 2\delta + \frac{4\beta \epsilon}{\delta} + \frac{\beta^2}{2} \epsilon^2 \right) \right] \right] \right. \quad (9)$$

## GENERAL STABILITY CRITERION

From our final expression Eq. (9), we can find a condition for  $\beta$  when an instability occurs ( $\lambda \geq 0$ ). Conversely, a general sta-

bility criterion can be written in the form

$$\beta < \frac{A + (A^2 - 4h\beta)^{1/2}}{2B} \quad (10)$$

with

$$A = 4 \left( 1 + \frac{h_0}{\alpha} \right) + 2 \frac{h_0}{\delta} (2\delta + \epsilon) + 2\delta + 4\alpha \left( 1 + \frac{h_0}{2\alpha} \right) \quad (11)$$

$$B = \frac{4}{3} \left( 1 + \frac{h_0}{\alpha} \right) \epsilon^2 + \frac{h}{2} \epsilon^2 \quad (12)$$

For greater clarity, it is convenient to present criterion Eq. (10) in terms of an energy balance. The magnetic energy released by the flux jump should be less than the available enthalpy in the system (composite and exchange to the surrounding helium layer):

$$\frac{\alpha \mu_0 I_0^2 \Omega^2}{3} \epsilon^2 < C_T (T_c - T_0) (1 - f) h_0 K_c C_T \rho \quad (13)$$

where  $f$  is a function containing all the extra terms issuing from Eq. (10). To confirm the general character of this criterion for composites with a highly resistive matrix, we can show how some usual criteria can be found, at the cost of a few approximations.

We must bear in mind in any case that the first assumption that has been made in our calculations is  $\epsilon \rightarrow 0$  in the filamentary zone ( $r_0 \rightarrow \infty$ ).

Let us recall a few typical orders of magnitude for physical properties (Table 1). It can be seen that thermal and electrical diffusivities are exactly permutable for the inner region and the outer shell.

It takes 0.5  $\mu$ s to diffuse heat over 0.5 mm in the copper or magnetic flux in the CuNi matrix, whereas it takes a much longer time (0.3 ms) to diffuse magnetic flux in the copper or heat in the CuNi matrix. This means that the process in the filamentary region is almost locally adiabatic. There is hardly any current generated in the inner core due to the fact that at the initiation of the flux jump the electric field in the inner core is zero when it is at a maximum at the interface between the filamentary region and the outer copper shell. The self-field effect tends to expel the excess current to the periphery.

## APPLICATIONS TO A FEW SIMPLIFIED OR USUAL CRITERIA

### Adiabatic Criterion

If one assumes  $h_0 \rightarrow 0$  (no heat exchange with the helium bath), Eq. (10) becomes

$$\beta < \frac{1}{\beta^2} \left( 1 - \frac{\epsilon_n C_n}{a_1 C_r} \right) \quad (13')$$

and using Eq. (11) we obtain

$$\frac{\alpha \mu_0 I_0^2 \Omega^2}{3} \epsilon^2 < C_T (T_c - T_0) \left( 1 + \frac{\epsilon_n C_n}{\alpha C_r} \right) \quad (14)$$

**Table 1. Physical Properties of the Materials of Superconducting Composites**

| Zone       | $K$ (W · mK <sup>-1</sup> ) | $C$ (J · m <sup>-1</sup> · K <sup>-1</sup> ) | $\rho$ (Ω · m)     | $D_r$ (m <sup>2</sup> · s <sup>-1</sup> ) | $D_n$ (m <sup>2</sup> · s <sup>-1</sup> ) |
|------------|-----------------------------|--|--------------------|---|---|
| NbTi, CuNi | 1.0                         | 1500   | $3 \times 10^{-4}$ | $8 \times 10^{-4}$                        | 0.55                                      |
| Cu         | 360                         | 1000   | $3 \times 10^{-8}$ | 0.3                                       | $2.5 \times 10^{-4}$                      |

This criterion indicates the role of the enthalpy of the outer normal metal shell. With  $\epsilon = 1$  and  $\epsilon_n = 0$  (no shell), we obtain the usual so-called adiabatic criterion (1-4). The "stable" parameter  $\beta$  varies with  $\epsilon$  as  $1/\epsilon^2$ . A direct application of this expression is the evaluation of the maximum average current density  $J$  that can be carried in a wire of half-dimension  $a$ . Letting

$$J = \eta_c J_c \quad \text{and} \quad C = C_0 \left( 1 + \frac{\epsilon_n C_n}{\alpha_r C_r} \right)$$

yields

$$J < \left( \frac{C_0 T_c - T_h}{\mu_0} \right)^{1/2} (\eta_c J_c)^{1/2} a^{-1/2} \quad (15)$$

Whereas using a wire with the highest possible critical current density seems important, the gain in stable density is not so significant, however. On the other hand, the stability is not as dependent on the size  $a$  as is often considered. Although it is true that the stability parameter  $\beta$  varies as  $a^2$ , the maximum stable average current density varies as  $a^{-1/2}$ . For instance, doubling the thickness  $a$  results in a stable average current density multiplied by 0.63.

**Dynamic Criterion**

Several workers have proposed analytical expressions for stability criteria of homogeneous composites (6-8). To our knowledge, they have always been presented for  $\epsilon = 1$  (critical current) and for nonzero values for  $\nu$  in the case of exact solutions obtained along a similar approach.

Let us consider a typical composite wire as a guide to justify some approximations, with  $\alpha_r = 0.7 \times 10^3$  m,  $\nu_r = 0.3 \times 10^{-3}$  m,  $\nu_n = 0.2$ ,  $\eta_c \nu_n = 0.5$ ,  $\nu_r \nu_n = 0.8$ , and  $h_1 = 10^8$  W · m<sup>-1</sup> · K<sup>-1</sup>. The basic assumption ( $\nu = 0$ ) is satisfied by  $\rho_0^2 \nu_n \gg K_r/C_r$ . To derive a dynamic criterion, it is necessary to assume that the heat conductivity in the inner region is not negligible ( $\nu_r$  is kept infinite). The two terms  $h_1 \alpha_r^2 K_r$  and  $h_2 \nu_n^2 K_n$  are considered to be much less than 1.

The general criterion can then be written as follows:

$$\frac{\mu_0 \eta_c J_c a^2 \omega^2}{3} < C_0 (T_c - T_h) \left( 1 + A_1 \frac{h_1 \alpha_r}{K_r} + A_2 \frac{h_2 \nu_n}{K_n} \right) \quad (16)$$

with

$$A_1 = \frac{2 + 3 \left( \frac{\epsilon_r C_r}{\alpha_r C_0} \right) \cdot \frac{\epsilon'}{3}}{4 \left( 1 + \frac{\epsilon_n C_n}{\alpha_r C_r} \right)^2} - \frac{1}{3} \quad (17)$$

$$A_2 = \left[ \left( 1 + \frac{\epsilon_n C_n}{\alpha_r C_r} \right) \left[ 2 + \frac{\mu_0 K_n}{\nu_r C_r} \left( \frac{\epsilon_n}{\alpha_r} + 2\epsilon \right) + \frac{\nu_n C_n}{\alpha_r C_r} \right] + 2 - 3 \left( \frac{\epsilon_n C_n}{\alpha_r C_r} + \left( \frac{\epsilon_n C_n}{\alpha_r C_r} \right)^2 \right) / 4 \left( 1 + \frac{\epsilon_n C_n}{\alpha_r C_r} \right)^2 + \frac{\mu_0 K_n}{\nu_r C_r} \left( \frac{\nu_n}{\alpha_r} + 2\epsilon \right) \left( 2 + \frac{\epsilon_r C_r}{\alpha_r C_r} \right) \right] / 4 \left( 1 + \frac{\epsilon_n C_n}{\alpha_r C_r} \right)^2 - 1 \quad (18)$$

Equation (16) emphasizes the respective influences of the inner region and of the outer shell. From Eq. (16), two particular cases can be derived as follows.

- 1. Letting  $\epsilon_n = 0$  (no outer normal shell) and  $\epsilon = 1$ , we find

$$\frac{\mu_0 \eta_c J_c a^2 \omega^2}{3} < C_0 (T_c - T_h) \left( 1 + \frac{5}{24} \frac{h_1 \alpha_r}{K_r} \right) \quad (19)$$

This expression is similar to the expression obtained previously (6) in the case of a homogeneous multifilamentary composite for low values of the ratio  $\nu$ :

$$\frac{\mu_0 \eta_c J_c a^2 \omega^2}{3} < C_0 (T_c - T_h) \left( 1 + \frac{7}{20} \frac{h_1 \alpha_r}{K_n} \right) \quad (20)$$

The small discrepancy between the two expressions can be mainly ascribed to the fact that the calculations performed previously (6) were only correct for  $\nu \neq 0$ , whereas the present calculations were carried out with  $\nu = 0$ .

- 2. If we keep  $\epsilon_n \neq 0$  and assume a very low resistivity for the outer normal metal shell (made of copper, for instance), Eq. (16) becomes

$$\frac{\mu_0 \eta_c J_c a^2 \omega^2}{3} < C_0 (T_c - T_h) \left( 1 + \frac{\epsilon_n C_n}{\alpha_r C_r} \right) \left( 1 - A_2 \frac{h_2 \nu_n}{K_n} \right) \quad (21)$$

with

$$A_2 = \frac{\mu_0 K_n}{\alpha_r C_r} \left( \frac{\epsilon_n}{\alpha_r} + 2\epsilon \right) \frac{3 + 2 \left( \frac{\epsilon_r C_r}{\alpha_r C_r} \right)}{4 \left( 1 + \frac{\epsilon_n C_n}{\alpha_r C_r} \right)^2} \quad (22)$$

Again we find a simplified criterion that emphasizes the beneficial influence of a low resistivity  $\nu_r$  as in the case of homogeneous composites. This is already visible in Eqs. (21) and (22). It can be pointed out more clearly for  $\epsilon_n/\alpha_r \ll 1$ , which permits another simplification step. It follows

that

$$\frac{\rho_0 l^2 \eta_0 \dot{I}_0 a^2}{3} \beta < C_1 (T_c - T_b) \left[ 1 + \frac{\eta_0 C_1}{\alpha_r C_r} + \frac{3 \mu_0 h_0 \alpha_r}{4 \rho_r C_r} \left( \frac{\alpha_r}{\alpha_c} - 2 \right) \right] \quad (23)$$

We can see that the simplified Eq. (23) of our general criterion Eq. (10) can also be put in a form similar to more conventional criteria derived for homogeneous composites. For instance, for  $\epsilon = 1$  (and in the frame of the particular assumptions), we find stability for

$$\frac{\rho_0 l^2 \eta_0 \dot{I}_0 a^2}{3} < (T_c - T_r) \left( C_1 + C_2 \frac{\alpha_r}{\alpha_c} - \frac{3}{2} \frac{\mu_0 h_0 \alpha_r}{\rho_r} \right) \quad (24)$$

when it was determined for homogeneous composites in previous work.

$$\frac{\rho_0 l^2 \eta_0 \dot{I}_0 a^2}{3} < (T_c - T_b) \left( C_m + \frac{3}{10} \frac{\mu_0 h_0 \alpha_r}{\rho_r} \right)$$

(exact solution in Ref. 6) and

$$\frac{\rho_0 l^2 \eta_0 \dot{I}_0 a^2}{3} < (T_c - T_b) \left( C_m - \frac{4}{\pi^2} \frac{\mu_0 h_0 \alpha_r}{\rho_r} \right)$$

(approximate solution in Ref. 8). Both expressions were given with no outer copper shell.

We can see that the criterion given in Eqs. (23) and (24) indicates the enhancement of stability by the heat removal to the coolant due to the "shell effect." In addition to the enthalpy of the composite, the third term in Eq. (24) represents the enthalpy transferred to the helium during the diffusion time of the current in the copper shell.

The enhancement of stability with increase in the Cu shell thickness is clear. Evidently a good heat transfer is of an help when no copper can damp or slow the development of the flux jump. In contrast, with a sufficient copper thickness, time is provided to take advantage of the heat transfer to the coolant and therefore for accounting for the enthalpy absorbed by helium in the heat energy balance.

#### Similarity with the Cryogenic Criterion

Equation (23) contains two terms: the enthalpy of the composite and the enthalpy absorbed by the helium. If the resistivity of the shell is small enough or the heat transfer large enough, the enthalpy of the composite becomes negligible, and stability is completely ensured by the transfer to helium.

The criterion becomes:

$$\frac{\rho_0 l^2 \eta_0 \dot{I}_0 a^2}{3} \beta < \frac{3 \mu_0 h_0 \alpha_r}{4 \rho_r} \left( \frac{\alpha_r}{\alpha_c} - 2 \right) (T_c - T_b) \quad (25)$$

Under these conditions, the criterion can be written, to a first approximation, in the form of generated and exchanged heat

fluxes. For a thin outer copper shell

$$\frac{4 \rho_0 l^2 \eta_0 \dot{I}_0 a^2}{9 \rho_r} < (T_c - T_b) \quad (26)$$

#### Effect of Thermal Conductivity

In the general expression Eq. (10), assuming now a very high heat exchange to the helium ( $h_0 \rightarrow \infty$ ) and good thermal and electrical properties of the steel (high  $K_1$  and low  $\rho_1$ ), we obtain

$$A \approx \frac{4 \mu_0 h_0 l^2 \rho_0}{\rho_r C_1}$$

$$B \approx \frac{h_0 \alpha_r}{2 K_1} \beta^2$$

hence

$$A^2 \gg 4 \frac{h_0 \alpha_r}{K_1} B$$

In particular, when  $\epsilon = 1$ , it can be found that the condition for stability becomes

$$\frac{\rho_0 l^2 \eta_0 \dot{I}_0 a^2}{6 K_1 (T_c - T_b)} \frac{\rho_r}{\alpha_r} < 1 \quad (27)$$

This expression can be compared with the expression that was regarded as being the dynamic stability criterion.

$$\frac{\rho_0 l^2 \eta_0 \dot{I}_0 a^2}{3 K_m (T_c - T_b)} < 1$$

This criterion, developed by several groups, quoted in Ref. 8, assumes an infinite heat transfer coefficient to the coolant. As a result, it is too optimistic in most practical cases. However, it is correct in that it determines the limit for the existence of stationary solutions for temperature-dependent critical current profiles inside the filamentary region. It points out clearly, in this case, the effect of poor thermal conductivity.

#### Simplified Analytical Expression for Homogeneous Composites

In the general criterion Eq. (10) written in a dimensionless form, stability is ensured when  $\beta < Q(\beta)$ , where

$$Q(\beta) = \frac{A - (A^2 + 4 B \beta)^{1/2}}{2 B}$$

The coefficients  $A$  and  $B$  include all the physical parameters, the heat transfer, and the reduced transport current. For a given set of physical parameters, it is possible to plot the expression  $Q(\beta)$  as a function of the current  $i$  in the form of a boundary line between stability and instability for any given value of the pair  $(\beta, i)$ .

In order to compare with theoretical and numerical calculations performed previously (6), we give here an analytical expression for the particular case of a highly resistive matrix superconducting composite without a stabilizing shell ( $\rho_m = 0$ ).



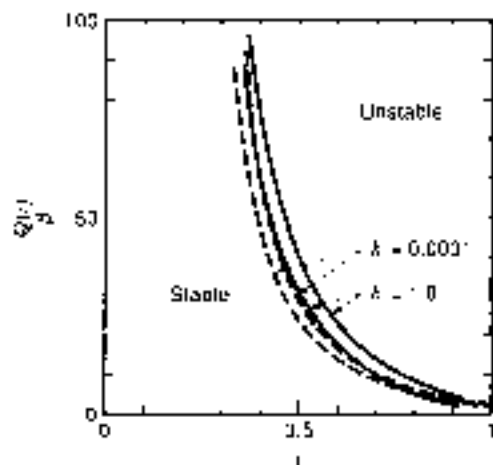


Figure 2. Comparison of stability curves for composites with an outer shell with axial curves given for homogeneous composites. Dashed lines, this work ( $\epsilon = 0$ ), solid lines, Ref. 5 ( $\epsilon = 0.1$ )

In the general criterion Eq. (10), let us set  $\epsilon_1 = 0$  and define  $h = h_1 a_1 / K_1$ . We find

$$\beta = \frac{3(2h - 2) - 12h^2(2 + z^2) + 16h^3 + z^2(3z + 16)^2}{6 + 36h} \quad (28)$$

This expression can be compared with the expression obtained for the homogeneous composite studied previously (6) in the case of a low ratio of thermal diffusivity to magnetic diffusivity ( $\epsilon$ ):

$$\beta = \frac{D_1}{D_r}$$

Figure 2 shows the stability function  $Q(z)$  for two values of reduced heat transfer  $h$ . In the present approach  $\epsilon = 0$  because

$\mu_1$  is infinite. In the previous approach (6), the limiting case  $\epsilon = 0.1$  was considered. It can be seen that the results are similar if it is remembered that the theory of the homogeneous composite was developed for a round wire (a correction coefficient of  $\pi/4$  was applied to the curves plotted previously (6) to obtain the curves plotted in Fig. 2).

Developing Eq. (28) for  $z = 1$  and for small values of the reduced heat transfer  $h$  leads to

$$\beta_1 = 3 \left( 1 + \frac{5h}{24} \right)$$

(another form of Eq. (19)).

Our expression especially developed for the case of an outer good conducting shell is in perfect agreement with previous models devoted to homogeneous composites.

SUMMARY

At the cost of some particular assumptions, it has been shown that the general criterion Eq. (10) encompasses the usual criteria that have been already developed in various particular cases. This can be represented schematically as in Fig. 3.

INFLUENCE OF THE OUTER SHELL THICKNESS AND OF THE HEAT TRANSFER COEFFICIENTS

A set of curves have been plotted in Fig. 4 using the more general criterion Eq. (10) directly. The particular case under investigation corresponds to typical orders of magnitude of  $\gamma = 200$ ,  $\alpha = 200$ , and  $\lambda = 1$ .

The thickness of the shell and the heat transfer coefficients are given in the form of the dimensionless parameters  $\epsilon = \epsilon_1 / a_1$  and  $h = h_1 a_1 / K_1$ .

It is clearly seen in Fig. 4(a), for  $\epsilon = 0$ , that because of the low thermal conductivity of the filamentary region, little

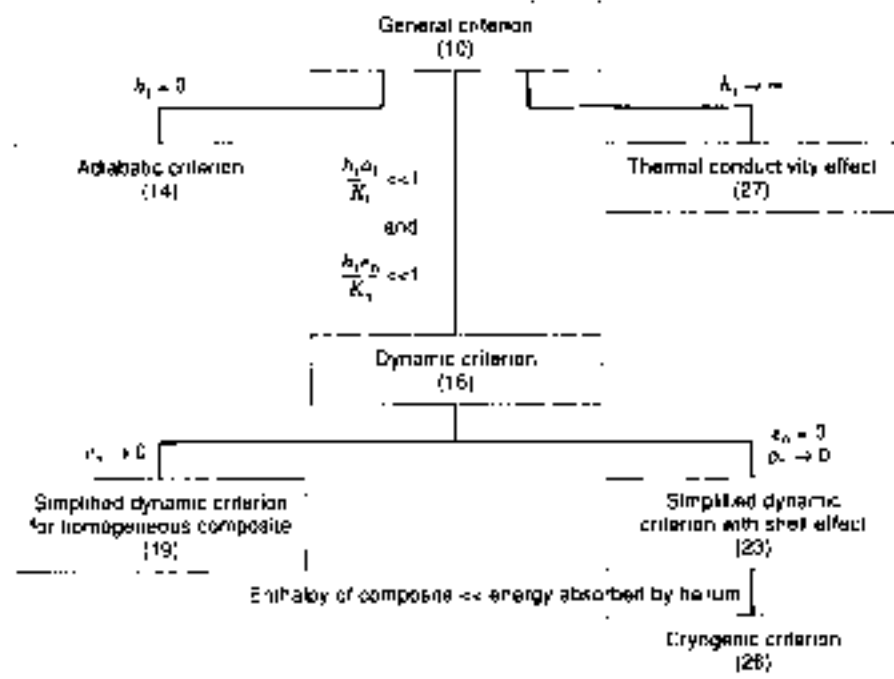


Figure 3. Summary of conditions to deduce usual criteria from the general stability criterion

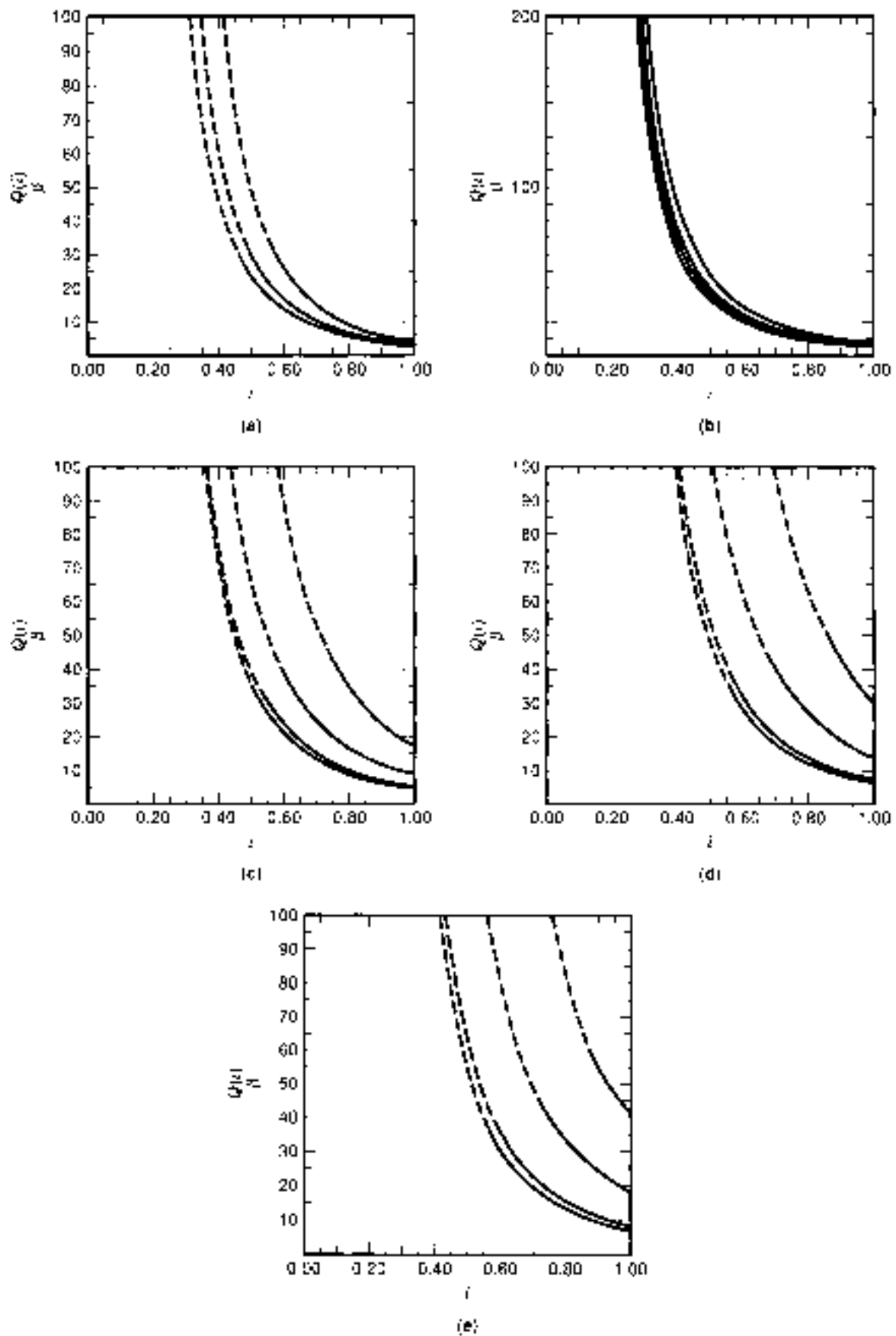


Figure 4. Stability is improved when both copper shell thickness and the heat transfer coefficient; at least, significant values for copper shell thickness  $\delta = \nu_{Cu} \alpha_T^{-1}$  (a) 0.1; (b) 0.3; (c) 0.6; (d) 1.0; (e) 3. Lines from top to bottom:  $h = 10, 1, 0.1, 0.01$ .

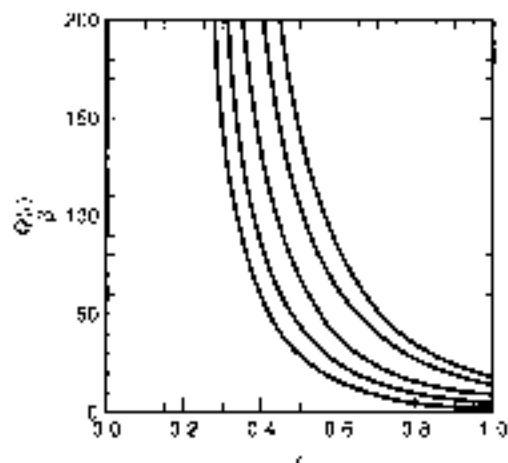


Figure 5. Influence of the copper shell thickness  $\epsilon = \epsilon_0/\mu_0$  on stability curves for  $h = 1$ . Lines from left to right:  $\epsilon = 0, 0.3, 0.6, 1.0, 1.3$ .

as expected in improving the stability from a good heat transfer to the coolant. Figures 4(a)–4(c) show the significant role of the heat transfer coefficient when  $\epsilon$  is not negligible (for low values of  $h$ , improvement is only provided by the increase in the equivalent heat capacity of the composite). A noticeable improvement in stability is provided when both  $\epsilon$  and  $h$  simultaneously attain significant values.

The copper fraction in the conductor is given by  $\eta_{Cu} = \epsilon/(1 + \epsilon)$ . This expression means that the amount of copper is totally concentrated in the outer shell. It is, for instance, 23% for  $\epsilon = 0.3$  and 56% for  $\epsilon = 1.3$ . The reduced heat transfer is about  $h = 1$  for  $k_1 = 10^5 \text{ W} \cdot \text{m}^{-2} \cdot \text{K}^{-1}$  and  $\mu_0 = 0.6 \times 10^{-6} \text{ m}$ .

The influence of the shell thickness is summarized in one set of curves for  $h = 1$  in Fig. 6.

#### DETERMINATION OF THE STABLE CURRENT $i_c$ IN A PARTICULAR EXAMPLE

For any given composite there is no special difficulty in plotting the general curve  $Q(i_c)$  of criterion Eq. (10). Then, the evaluation of the parameter

$$\beta = \frac{\mu_0 i_c \epsilon_0 k_1 a^2}{C_p (T_c - T_b)}$$

leads immediately to the determination of the maximum stable current  $i_c$ .

As a practical example, let us consider a multifilamentary composite with a cupronickel matrix surrounded by a copper shell with the two fixed parameters  $2R = 1.35 \text{ mm}$  (diameter) and  $\eta_c = 0.20$  (superconductor volume fraction).

In order to point out again the influence of the outer copper, let us present the curves in terms of copper volume fraction (the CuNi is entirely located in the filamentary region):

$$\beta^2 = (1 - \eta_c) R^2 \quad \text{and} \quad \eta_{Cu} = 1 - \eta_c / \epsilon_0$$

The equivalent dimensions used in the slab model are given by

$$a = \frac{\sqrt{\pi}}{2} R \quad \text{and} \quad a_0 = \frac{\sqrt{\pi}}{2} R_0$$

respectively. All the parameters intervening in the expressions for  $Q(i_c)$  can be derived as equivalent physical properties.

The expressions  $Q(i_c)$  are plotted in Figs. 6(a)–6(e) with the copper fraction as a parameter for given values of the actual heat transfer coefficient  $h_1$  (up to a heat transfer of  $10^5 \text{ W} \cdot \text{m}^{-2} \cdot \text{K}^{-1}$ ), the composite can be considered to behave as in adiabatic conditions.

In order to determine the stable operating current, we can proceed as follows. Calculate the parameter

$$\beta = \frac{\mu_0 i_c \epsilon_0 k_1 a^2}{C_p (T_c - T_b)}$$

and then, for a given copper fraction and a given heat transfer coefficient, determine the maximum operating current. For instance, Fig. 6(d) shows the evaluation of the stable current for two amounts of copper for a particular case. In a field of 11 T and if one assumes the strand to be cooled in superfluid helium to 1.8 K ( $h_1 = 5 \times 10^5 \text{ W} \cdot \text{m}^{-2}$ ),  $\mu_0 = 90$ , which means that the transport current can be increased safely up to 60% of the critical current for 60% of copper, whereas it is 77% of the critical current for 70% of copper.

#### SMALL PERTURBATION AND THE CRITICAL STATE MODEL

The theoretical calculations that have been discussed in the previous sections assume an ideal critical state model, that is, the electric field and the current density in the superconductor are related simply by  $E \neq 0 \rightarrow J = J_c$ . Actually, in composites made up with a large number of fine filaments, a continuous dependence of  $E$  on  $J$ , caused mainly by inhomogeneities, can be observed.

One of the expressions that is most often accepted is

$$E = E_0 \left( \frac{J}{J_c} \right)^n$$

where  $n$  is of the order of 10–100 depending on the field level and on the quality of manufacture. Under these conditions, the "critical current" is defined for a given electric field. For instance, the conditions  $E_c = 10^{-5} \text{ V} \cdot \text{m}^{-1}$  and  $J_{c0} = 2 \times 10^{10} \text{ A} \cdot \text{m}^{-2}$  result in another definition of critical current for  $E = 10^{-11} \text{ V} \cdot \text{m}^{-1}$ :  $J_c = 2.16 \times 10^6 \text{ A} \cdot \text{m}^{-2}$  for  $n = 30$  and  $J_c = 2.046 \times 10^6 \text{ A} \cdot \text{m}^{-2}$  for  $n = 100$ .

One of the consequences is that the concept of a small perturbation is more difficult to define, since there is no step function in the  $E(J)$  relationship. Another way of pointing out this effect is to compare the apparent resistivity in this resistive "critical state" with the resistivity  $\rho_{Cu}$  of the stabilizing copper.

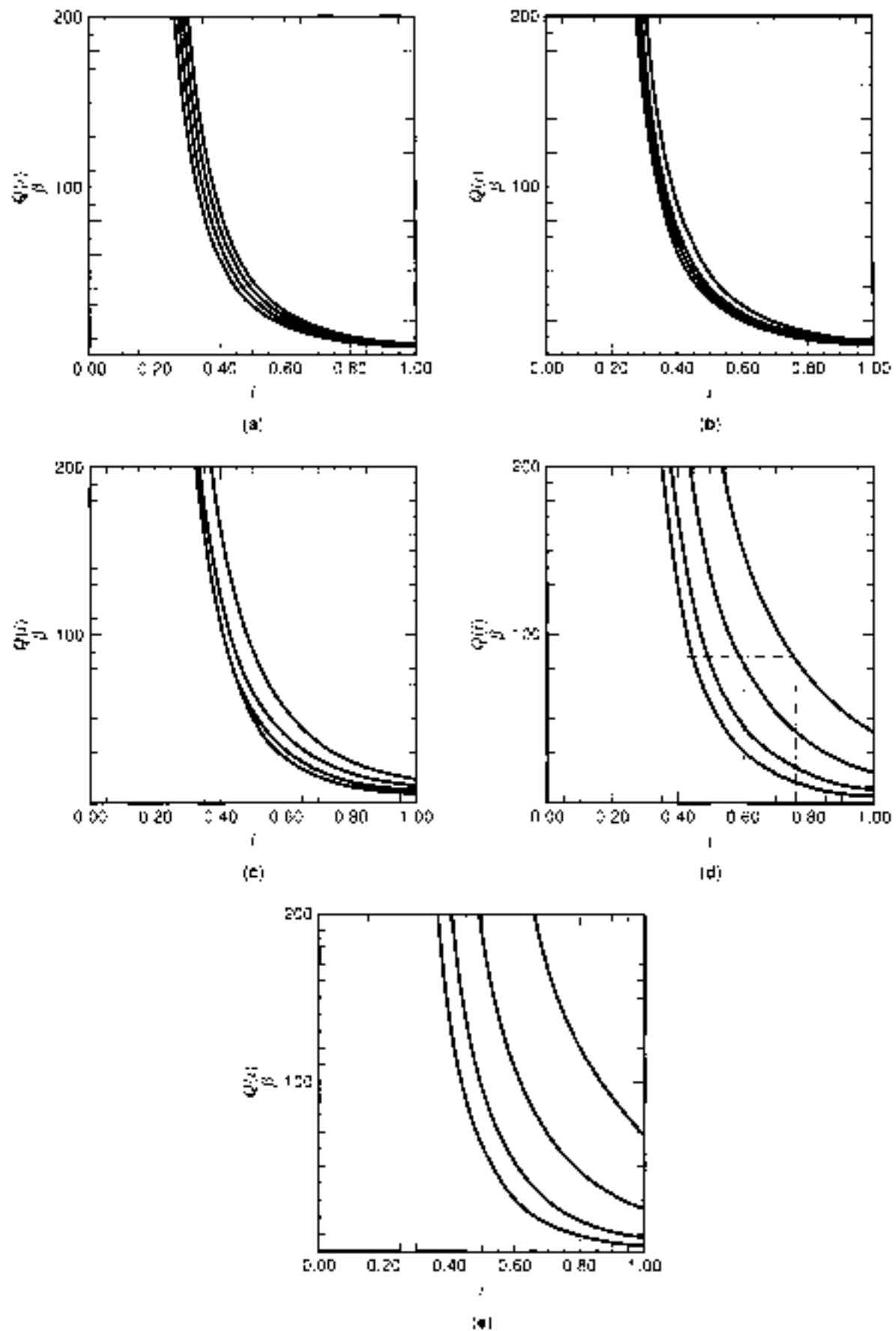
The differential resistivity is

$$\frac{dE}{dJ} = n \frac{E_0}{J_c} \left( \frac{J}{J_c} \right)^{n-1}$$

The "effective critical" current  $J_c^*$  for which this value equals the copper resistivity can be estimated. For  $nE_c/J_c = \rho_{Cu}$ ,

$$J_c^* = \left( n \rho_{Cu} \frac{J_c}{E_0} \right)^{1/(n-1)}$$

Assuming  $\rho_{Cu} \approx 5 \times 10^{-10} \Omega \cdot \text{m}$  and  $E_0 = 10^{-5} \text{ V} \cdot \text{m}^{-1}$ , for  $n =$



**Figure 6.** Stability curves for a 1.35 mm diameter wire with 50% of superconductor  $k_c = 19(0, 1)$ ,  $10^5$ ;  $\alpha_1 = 10^5$ ;  $\alpha_1(5 \times 10^5)$ ;  $\omega = 19^4 \text{ W} \cdot \text{m}^{-2} \cdot \text{K}^{-1}$ . Lines from left to right:  $\alpha_2 = 0.3, 0.5, 0.6, 0.7$ .

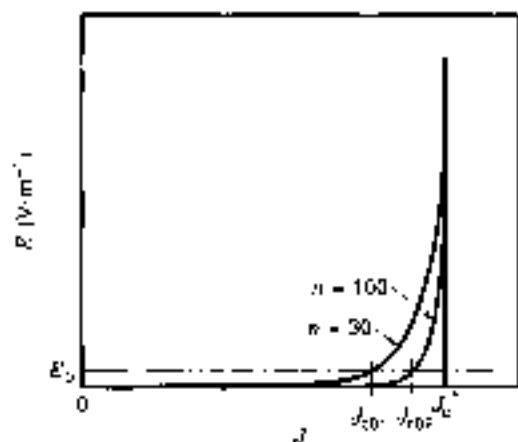


Figure 7. How the determination of  $J_c^*$  is affected by the "n" value.

30

$$J_c^* = 1.32 J_{c0}$$

and for  $n = 100$

$$J_c^* = 1.07 J_{c0}$$

This shows that the critical current density  $J_{c0}$  defined for  $E_0$  has to be noticeably exceeded in order to obtain a stabilizing effect by the copper matrix.

In order to establish the full relevance to our model, the perturbation has to generate enough electric field so that  $J_c$  is locally exceeded. This is hardly to be expected from the ac losses produced by an external changing field or by the self field of the increasing current itself, except in case of very high field or current change rate (in less than 10 ns).

A way to cope with this problem is to evaluate the stability through  $\beta$  and  $i$ , not in using  $J_{c0}$  to define  $\beta$  but with  $J_c$  taken as 10 to 30% more than  $J_{c0}$  depending on the  $n$  value (Fig. 7). This will result in a larger value for  $\beta$  (20 to 60%) and a smaller value for  $i$ . However, the stable transport current  $I_c$  will be larger

$$I_c^* = 1.1 I_c$$

As a result, one can see that a low  $n$  value could suggest that the conductor is more stable, when actually it is only the "true" critical current that has been underestimated.

Some other factors can give rise to a more stable current than given in the calculations. They arise from the existence of a steady-state current distribution (for  $E \approx 0$ ) that is different from the distribution given by the ideal critical state model. The spiral aspect of the filament enables some current to flow in the unsaturated region (5), after a current increase, the current decays in the outer layers of filaments because of the resistive effect ( $\nu$  value), and in a conductor with limited length, the current distribution is mainly imposed by the transfer resistances from layer to layer.

To conclude with these considerations, we can describe a typical perturbation that can obviously give rise to the required electric field for transferring enough current in the copper (Fig. 8). If we consider that the current-sharing temperature

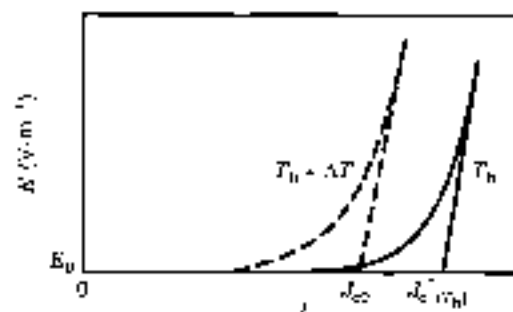


Figure 8. Effect of a small temperature perturbation on the  $E \cdot J_c$  curve.

is given by

$$T_c - T_h = (T_c - T_h)(1 - \epsilon)$$

and if we consider that a shift in critical current of typically 10% is at least necessary, because of the shape of the  $E(J_c)$  curve, a small perturbation  $\Delta T \approx 0.1(T_c - T_h)$  is typical for initiating a flux jump in a wire whose operating parameters ( $B, h, l, \nu$ ) just lie on the theoretical curves between stable and unstable behaviour (in other words,  $J_c, E_0, T_h$ ) can also be considered as  $J_c$  at the temperature  $T_h + \Delta T$ ; see Fig. 7).

CONCLUSION

The general analytical criterion that has been presented concerns the intrinsic stability of superconducting composites with a highly resistive matrix. It has shown consistency and continuity with previous theoretical results.

Apart from the general character of this criterion, the influence of an outer copper shell on intrinsic stability conditions has been clearly emphasized with the help of some analytical expressions.

The "critical" stability curve  $Q_c(i)$  for a given composite under precise operating conditions acts as a useful basis to predict its behavior and to determine a "safe" current for intrinsic stability.

Conversely, our criterion can be used in a straightforward manner for a composite design. It becomes possible to consider how a modification of physical characteristics (ratios, relative positions of the materials, etc.) can improve the stability conditions. The general criterion Eq. (10), given in an analytical form, can be of great help in the optimization of the structure of low-loss composites.

We have also pointed out the great advantage of locating the copper at the periphery of the composite in order to take full advantage of the appearance of a maximum electric field to help current flow in the copper. On the other hand, copper around the axis is of almost no help, as the electric field is always negligible in the unsaturated region. However, the debate is not totally settled. In fact, the self-field effect is reduced and the conditions for a flux jump are less likely when the filaments are distributed in a concentric layer with no superconducting region close to the axis (9). The theory has to be developed for this case. Also undetermined is whether a good compromise exists with some copper both in the center and at the periphery.

## NOMENCLATURE

|          |  |
|----------|--|
| $a$      | Composite external radius (slab model) (m)   |
| $a_0$    | Multifilamentary zone radius (m)   |
| $a_s$    | Saturated zone radius (m)  |
| $B_0$    | Background magnetic field  |
| $C$      | Specific heat of the composite ( $J \cdot m^{-3} \cdot K^{-1}$ )                         |
| $C_1$    | Average specific heat in the filamentary zone ( $J \cdot m^{-3} \cdot K^{-1}$ )          |
| $C_m$    | Average specific heat in homogeneous composites ( $J \cdot m^{-3} \cdot K^{-1}$ )        |
| $C_n$    | Average specific heat in the normal metal shell ( $J \cdot m^{-3} \cdot K^{-1}$ )        |
| $D_m$    | Magnetic diffusivity ( $m^2 \cdot s^{-1}$ )  |
| $D$      | Thermal diffusivity ( $m^2 \cdot s^{-1}$ )   |
| $E$      | Electric field ( $V \cdot m^{-1}$ )  |
| $e_r$    | Normal metal shell thickness (m)   |
| $h$      | $k_m/K_f$ reduced heat transfer coefficient  |
| $h_1$    | Heat transfer coefficient ( $W \cdot m^{-2} \cdot K^{-1}$ )                              |
| $i$      | $I_c/I_c$ , reduced transport current  |
| $I_c$    | Critical current (A)   |
| $I_s$    | Stable transport current (A)   |
| $I_t$    | Transport current (A)  |
| $j$      | Current density in the composite ( $A \cdot m^{-2}$ )                                    |
| $j_c$    | Critical current density at $T$ for the given $B$ ( $A \cdot m^{-2}$ )                   |
| $j_{c0}$ | Critical current density at $T_0$ and $B$ ( $A \cdot m^{-2}$ )                           |
| $j_n$    | Current density in a normal metal region   |
| $K$      | Thermal conductivity in the composite ( $W \cdot m^{-1} \cdot K^{-1}$ )                  |
| $K_f$    | Average thermal conductivity in filamentary zone ( $W \cdot m^{-1} \cdot K^{-1}$ )       |
| $K_m$    | Average thermal conductivity in homogeneous composites ( $W \cdot m^{-1} \cdot K^{-1}$ ) |
| $K_n$    | Average thermal conductivity in normal metal shell ( $W \cdot m^{-1} \cdot K^{-1}$ )     |
| $T_0$    | Helium bath temperature (K)  |
| $T_c$    | Critical temperature for a given field (K)   |

## Greek Letters

|               |   |
|---------------|---|
| $\nu$         | $K_f/K_n$ , dimensionless ratio                                       |
| $\beta$       | Dimensionless stability parameter                                     |
| $\beta_c$     | Critical value of $\beta$ bounding stable and unstable domains        |
| $\gamma$      | $(K_f/K_n)(C_f/C_n)$ , dimensionless ratio                            |
| $\delta$      | $(\rho_n/\rho_0)(C_f/K_f)$ , dimensionless ratio                      |
| $\epsilon$    | $e_r/a_0$ , reduced thickness of normal metal shell                   |
| $\eta_f$      | Superconductor ratio in the filamentary zone                          |
| $\eta_c$      | Superconductor ratio in the composite                                 |
| $\eta_{Cu}$   | Copper ratio in the composite   |
| $\eta_{NiCu}$ | Cupronickel ratio in the composite                                    |
| $\lambda$     | Eigenvalue of the system  |
| $\mu_0$       | Magnetic permeability of vacuum                                       |
| $\rho$        | $(K_f/C_f)/(\mu_0/\rho_f)$ , dimensionless ratio of the diffusivities |
| $\rho_0$      | Matrix resistivity ( $\Omega \cdot m$ )                               |
| $\rho_m$      | Average composite normal resistivity ( $\Omega \cdot m$ )             |
| $\rho_n$      | Normal metal shell resistivity ( $\Omega \cdot m$ )                   |

## BIBLIOGRAPHY

1. P. S. Swartz and C. P. Bean, A model for magnetic instabilities in hard superconductors: adiabatic critical state. *J. Appl. Phys.*, **46**: 4991-4996, 1980.
2. M. N. Wilson et al., Experimental and theoretical studies of superconducting composites. *J. Phys. D*, **3**: 1517-1532, 1970.
3. M. G. Kromley, Stability of critical states in type II superconductors. *JETP Lett.*, **17**: 312-318, 1973.
4. M. G. Kromley, Damping of flux jumping by flux flow resistance. *Cryogenics*, **13**: 132-137, 1974.
5. J. J. Duchateau and B. Turck, Self field degradation effect in adiabatic conditions. *Cryogenics*, **14**: 491-496, 1974.
6. J. J. Duchateau and B. Turck, Dynamic stability and quenching currents of superconducting multifilamentary composites under normal cooling conditions. *J. Appl. Phys.*, **46**: 4989-4995, 1975.
7. R. G. Mears and A. L. Bokhriev, On the theory of flux jumps in hard superconductors. *J. Phys. D*, **9**: 2281-2287, 1976.
8. M. N. Wilson, *Superconducting Magnets*, Oxford, UK: Clarendon Press, 1983.
9. B. Turck, Self field effect in round and rectangular multifilament composites and stability of superconducting coils, *Proc. Conf. ICEC*, **8**, IPC Science and Technology Press, 497-503, 1976.

H. THURR  
CERN  
P. MARIGNI  
Euratom-CEA Association

## INTRODUCTION

Today superconductors can be found in applications as commonplace as MRI systems installed in thousands of hospitals and cellular-telephone base stations, as well as nonmedical applications such as particle accelerators. In Fig. 1 we illustrate two major commercial applications, magnetic resonance imaging (MRI) and high-field nuclear magnetic resonance (NMR). The difference in bore size is quite substantial.

This section is the largest section in the book and contains 26 articles. Many of the applications covered have perfectly viable conventional competitors that offer lower cost and higher consumer acceptance. A feature of this book is that we fully examine the conventional alternative. As superconductor properties continue to improve, we can expect that more and more applications will find in superconductivity a technological or economic edge.

An excellent case in point would be the application of high-temperature superconductor (HTS) microwave filters in cellular telephone base stations. Superconductor is almost lossless at microwave frequencies, and extremely sharp microwave filters can be made, using large numbers of poles, without incurring significant insertion losses. This makes them ideal for use as preselect filters in wireless base stations (1). The combination of high  $T_c$  and improved advances in cryocooling means that microwave filter systems can be supplied for base stations in compact, hot-swappable rack-mount units (2). Because they produce clearer signals, they not only improve service quality but can increase the service area or allow base stations to be built further apart. The accelerated pace at which these units are being installed reflects the competitive economics of the new technology.

### Superconducting Current Leads

The large currents the superconductors can carry must somehow be transferred to the cooled superconductor from ambient temperature. Prior to HTS, this has usually been the primary loss of heat for superconducting devices, because metals

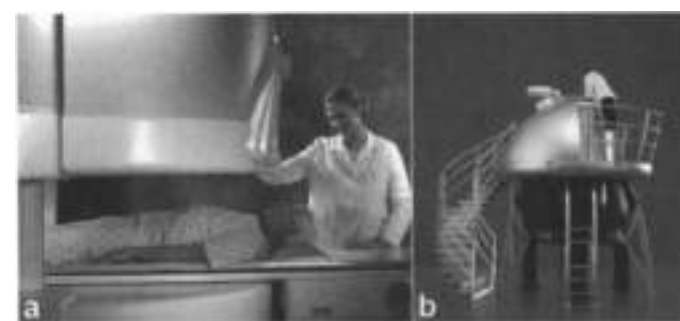
with high enough electrical conductivity to carry the huge currents also have high thermal conductivity. This heat loss has, in the past, significantly reduced the cost effectiveness of superconductor applications. The requirements for current leads for magnets are described in the articles **POWER QUALITY** under the section "What is  $\mu$ SMES?" and **SUPERCONDUCTING TRANSDUCERS**, under the section "Transformer Design and Analysis." HTS represents a breakthrough in this area because ceramics are not good thermal conductors. Of course, they cannot take supercurrent all the way to room temperatures (as of this writing), and account must be taken of their sensitivity to magnetic field; but they can be used as high-current thermal blocks. Combination normal-superconducting current leads are commercially available using BSCCO 2212 and 2223 and YBCO HTSs. HTS leads make both LTS and HTS superconductor applications economically more viable.

### And What of the Future?

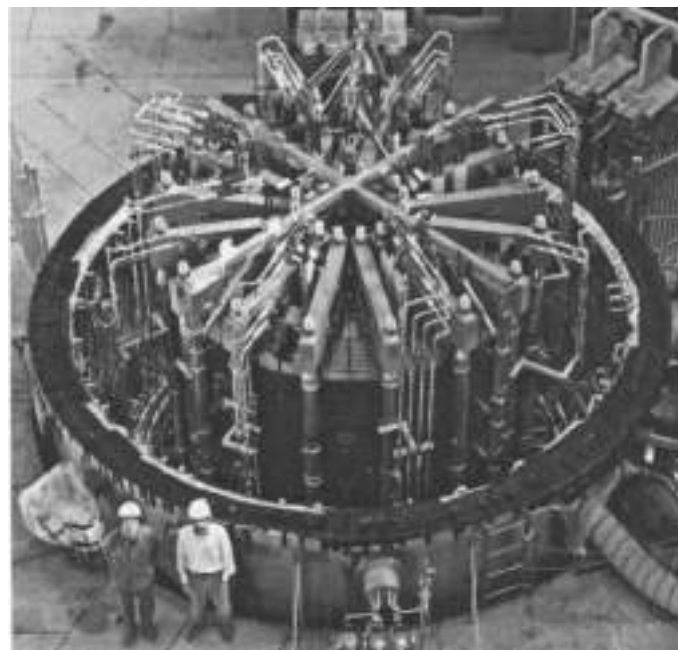
We have come to expect a continual improvement in superconductor properties, and nothing that we have seen over the last twenty years leads us to expect otherwise. This certainly is promising for power and other applications where the competition is copper, which has seen very little improvement over the last fifty years. Increasing energy costs (and the effects of deregulation in the U.S.) will also benefit the competitiveness of superconductors in power applications such as generators (3). Environmental issues should also encourage the application of superconducting power transmission cables. One such cable is already being installed at the Detroit Edison Frisbee substation in downtown Detroit and is expected to triple power throughput in an existing facility. For that project, American Superconductor shipped 4.25 mi of HTS wire to Pirelli for cable assembly in August 2000. SMES (see **POWER QUALITY**, section "MicroSMES") is already commercially available today and is competitive with conventional systems for some applications. A 1000 hp superconducting motor is being developed using HTS technology (4). The development of 10 kWh superconducting flywheel system was initiated by Boeing and ANL in 1999. There are several investigations of HTS-based fault-current limiter technologies that look promising (see **Superconducting Fault Current Limiters**). HTS transformers are being developed, including one by IGC, Waukesha Electric, and CRNL, which has so far resulted in a three-phase, 1 MVA unit (5).

Improved properties developed in  $Nb_3Sn$  strand during R&D production phases for the CS model coil of the International Thermonuclear Experimental Reactor (ITER) have made possible much lower-cost alternatives to the original design that could make fusion power a reality. The success of the ITER CS model coil shows that brittle superconductors, such as  $A15 Nb_3Sn$ , can be made into very large-scale magnets that can be used at high fields in demanding applications (see Fig. 2).

In 2005 the Large Hadron Collider (LHC) ring at CERN in Switzerland will contain over 1000 superconducting bending magnets, each 13 m long, manufactured using  $Nb-Ti$  superconductor. The magnets will be cooled down to 1.8 K and provide a field of 8.65 T. The designed collision energy will be 14 TeV. Already superconductors are being developed that can make a 100 TeV collider possible.



**Figure 1.** Low-field MRI and high-field NMR: (a) GE Signa Open-Speed MRI. This open MRI uses superconducting strand to create a highly stable 0.2 T field. (Courtesy of GE Medical Systems.) (b) 23.3 T  $Nb_3Sn$ -based superconducting magnet manufactured by Oxford Instruments, plc. (c) A high-resolution 900 MHz NMR system by Varian, Inc. (Courtesy of Oxford Instruments, plc.)



**Figure 2.** The world's most powerful pulsed superconducting magnet, the 150 ton 13 T (approximately 260 thousand times more powerful than the earth's magnetic field) ITER CS model. The magnet consists of two modules: the inner module fabricated in the U.S., and the outer fabricated in Japan. The two coils were combined at the Naka Fusion Research Establishment test facility of the Japan Atomic Energy Research Institute (JAERI). (Photo courtesy of and copyright retained by JAERI.)

## BIBLIOGRAPHY

1. M. J. Scheeren et al., Filter subsystems for wireless communications, *IEEE Trans. Appl. Supercond.*, **7**(2.3): 3744–3749, 1997.
2. Y. Ueno et al., High-temperature superconducting receiving filter subsystem for mobile telecommunication base station, *IEICE Trans. Electronics*, **ES8-C-57**: 1172–1176, 1999.
3. W. V. Hassenzahl, Applications of superconductivity to electric power systems, *IEEE Proc. Eng. Rev.*, **20**(5): 4–7, 2000. W. V. Hassenzahl, More applications of superconductivity to electric power systems, *IEEE Proc. Eng. Rev.*, **20**(6): 4–6, 2000.
4. D. Driscoll, V. Tombrassi, and B. Zhang, Development status of superconducting motors, *IEEE Assoc. Eng. Rev.*, **20**(5): 12–15, 2000.
5. R. W. McCunnell, S. Mehta, and M. S. Walker, HTS transformers, *IEEE Proc. Eng. Rev.*, **20**(6): 7–11, 2000.

## CAPACITOR STORAGE

### ENERGY STORAGE IN ELECTRICAL POWER SUPPLY SYSTEMS

Energy storage can be used to solve problems of mismatch between supply and demand. The availability of energy storage is particularly beneficial in electrical power supply systems which suffer disturbances or fluctuations. On the other hand, intrinsic energy storage in the power system can also cause problems during fault conditions. In transport applications, energy recovered during braking can be used during acceleration to supplement the primary energy source; for example, the overhead

supply system for an electric locomotive or the internal combustion engine for a hybrid/electric vehicle. In this way, the primary power source can be buffered from the peak power requirement.

In power engineering, the three main objectives of energy storage are:

- To reduce the overall cost of delivered energy in transmission and distribution systems,
- To reduce the overall cost of fuel input, by charging from less expensive base load generators at night, and during weekends and discharging such inexpensive energy daily instead of using more expensive oil, gas, or coal, and
- To improve the operation of a utility, by following the instantaneous variation in the demand for system regulation.

Surges in demand for electricity at certain times of the day can put a significant strain on the power generation system. Combined-cycle gas turbine (CCGT) plants, which can come on-line rapidly and operate efficiently for short periods, are used to cope with increases in demand such as in the early evening when people return home and start to prepare a meal. One could argue that the energy stored in the gas is released and converted into electricity for distribution. Pump storage systems are used for similar purposes. In this case, water is stored in a large reservoir, often high up in mountainous areas. When demand rises, the water is allowed to fall down a large shaft to a turbine which drives the generator. The water is later pumped back to the reservoir during off-peak periods ready for the next surge in demand. Again, very rapid but short-term power generation is available through such environmentally friendly schemes.

Other disturbances experienced by electrical power supply systems include short periods of local disconnection during faults. Examples on an even shorter time scale (up to a few seconds) include dips in system voltage due to sudden increases in demand or again during faults and increases in the local system voltage due to a large regenerative load such as a container crane lowering a container. Alternative energy sources, particularly wind turbines, also create disturbances in the electric power supply. The fact that over a period of days and sometimes hours wind conditions can vary from light or no wind to gales generally causes few problems unless the only available source of power is a wind turbine (see, for example, Somerville (1)). The gusting effect of the wind, however, results in large power fluctuation over periods of perhaps 2 s to 6 s. A range of problems can arise due to these fluctuations, particularly where many wind turbines are concentrated in a small area (2–4).

CCGT and pump storage schemes are not able to deal with such short-term problems. A variety of energy storage technologies have been proposed to absorb excess energy during periods where supply exceeds or is inadequate to satisfy demand for release during periods of excess demand over supply. Pneumatic systems, in which excess energy is used to compress air, have been reported (see, for example, Musgrove and Slack (5)). Such schemes involve large numbers of high-pressure gas cylinders to store significant quantities of energy. Flywheel systems, in which energy is stored in a rotating mass, have been the subject of considerable interest (6–11). Excess energy in the system is stored as kinetic energy by accelerating the flywheel and



released by decelerating the flywheel. The key to success in such schemes is the ability to control the flow of energy into and out of the flywheel efficiently. Other issues include the safety aspects of containing a high-speed rotating mass (in some cases, rotating at over 50,000 rpm), the need to maintain a vacuum within the containment housing to reduce losses, and the electrical power conversion equipment. As a result, the flywheel energy storage systems which have been developed have struggled to achieve the cost targets at which they become viable.

Electrochemical batteries are also used to store energy in electric power supply applications. Of course, electrochemical batteries are direct current devices and power conversion equipment is required to control the flow of energy between a battery pack and a power distribution grid system (12). Electrochemical batteries are widely used as the power source in a range of other applications. The lead acid battery still commands a major position in the market, particularly for vehicle starter batteries, in uninterruptible power supply units, in electric vehicles (e.g. fork lift trucks, milk floats, and golf buggies), as well as in evaluation electric cars and vans and electric, hybrid vehicles whilst new forms, particularly lithium batteries, have taken a major share of the portable equipment market. However, all forms of an electrochemical battery generally suffer from the disadvantage of poor cycle life and low specific power, particularly on charge. They suffer when subjected to cycling about a partially charged state or deep discharge.

Energy is also stored whenever magnetic or electric fields are produced. A magnetic field is created by flow of current in an electric circuit. The strength of the magnetic field, and hence the energy stored in the field, is determined by the dimensions of the field and the material in which it exists. The energy stored in the magnetic field is proportional to the square of the current flowing in the electric circuit. For efficient energy storage based on this principle, large quantities of current must flow in the circuit with minimal loss. This requires a circuit with very low resistance. To achieve low resistance, superconducting materials must be used to form the circuit. The resistance of superconducting materials becomes very small at very low temperatures within a few degrees of absolute zero ( $-273^{\circ}\text{C}$  or  $0\text{ K}$  (kelvin)). Superconducting magnetic energy storage (SMES) has been developed (see, for example, De Winkel and Lanzaire (13)). Materials used are based on niobium which becomes superconducting at 9.2 K. The viability of such technology depends to a great extent on the outcome of research to develop materials which become superconducting at higher, more easily maintainable temperatures. About 1987, it was discovered that oxides, such as yttrium barium copper oxide ( $\text{YBaCu}_3\text{O}_7$ ), become superconducting at temperatures as high as 80 K or 90 K, and thus, can be cooled by liquid nitrogen. (see Matthias (14) and Tanaka (15)).

An electric field is created and energy is stored when a potential difference exists across a region. The relationships between the potential difference, the electric charge, and energy stored in the field are functions of the dimensions of the field and the material in which the electric field exists. The ratio of electric charge to potential difference is referred to as capacitance. As will be shown in this article, the energy stored in an electric field is proportional to the square of the potential difference or the square of the electric charge in the region of the field. This is analogous to energy storage in a magnetic field, however, since the energy stored in the electric field is a function of potential difference (voltage) not current, superconducting materials and

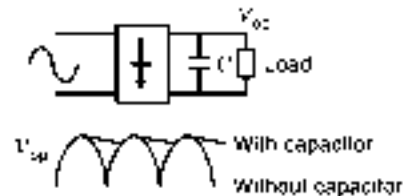


Figure 1. Use of capacitive energy storage for voltage waveform smoothing.

their associated refrigeration systems are not required in this case. Capacitive energy storage systems, however, require the development of structures (capacitors) which possess high values of capacitance but also techniques for efficiently charging and discharging the system.

A simple approach to the application of capacitive energy storage is found in electronic power supplies. A capacitor is placed after the rectifier in the power supply circuit (see Fig. 1) to smooth out the direct but fluctuating rectified voltage. The capacitor stores energy during the period when the rectifier output voltage,  $V_{in}$ , is high and supplies the load during the period when  $V_{in}$  is low.

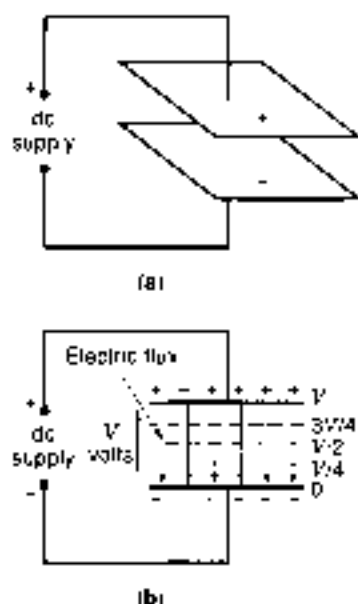
The voltage across the capacitor varies as energy is stored or released. In the previous example, a small variation in voltage is tolerated in a trade-off between regulation performance and cost. Only a small portion of the energy stored is used since the voltage across the capacitor varies by only a small amount. More of the stored energy can be used if the capacitor voltage is allowed to change over a greater range. Whilst some systems will tolerate a wide variation in voltage, it is more common for the system voltage to be fixed within a narrow band. In such cases, a power electronic converter may be used in combination with the capacitor to control the flow of charge (current) into and out of the capacitor whilst maintaining the voltage presented to the system within which the energy store operates. This topic is expanded later in the article.

## THEORY OF CAPACITOR ENERGY STORAGE

### Fundamental Relationships

Conceptually a capacitor consists of two adjacent parallel conducting plates, as shown in Fig. 2(a), which are separated by an insulator called a dielectric. As the plates are electrical conductors, each contains a large number of mobile electrons, some of which are bouncing freely on the surface, and have a small negative charge. These plates can be charged either negatively or positively. For a plate to become charged negatively, extra electrons must be drawn from a source of negative charge to create a surplus of electrons. During this process, those electrons that have already entered the plate oppose or repel other electrons which try to join them. The flow of electrons will stop when the negative repelling force equals the charging force. Similarly, for a plate to become positively charged, electrons must be withdrawn toward a source of positive charge, to create a deficiency of electrons. The flow of electrons ceases when the positive attracting force equals the charging force.

The plates are then said to be charged. The difference in charge across the gap between the plates results in a potential difference and hence an electric field. Figure 2(b) shows lines



**Figure 2.** (a) Parallel plate capacitor. (b) Charge storage and electric field in a parallel plate capacitor

of electrical flux between the plates which represent the field. Charges on the plates, such as those shown in Fig. 2(b), are associated with the electric field which appears as a potential gradient across the dielectric and potential difference between the two plates. The stored charge  $q$  (in coulombs) on a capacitor is proportional to the potential difference or applied voltage  $v$  (in volts), that is

$$q = Cv \quad (1)$$

The constant of proportionality  $C$  in Eq. (1) is referred to as the capacitance of the capacitor. The units of capacitance are farads (F); although  $\mu\text{F}$  ( $10^{-6}$  F) and  $\text{pF}$  ( $10^{-12}$  F) are more commonly used in practice.

As the capacitor is charged, a current  $i$  (in amperes) flows where

$$i = dq/dt = C dv/dt \quad (2)$$

The current is proportional to the rate of change of voltage (in volts per second,  $V/s$ ). For example, if the voltage across 1 F changed by 1  $V/s$ , the current is 1 A. To supply a current of 1 mA to 1 F, the voltage will rise at 1000  $V/s$ .

The electric field strength  $E$  (alternatively referred to as the electric force, electric stress, or voltage gradient with units of volts per meter,  $V/m$ ) is given by

$$E = v/d \quad (3)$$

where  $v$  is the potential difference across the dielectric or between the two plates and  $d$  is the dielectric thickness or spacing between the two electrodes.

The electric flux density  $D$  (also known as the total charge density with units of Coulombs per square meter,  $C/m^2$ ) between the two electrodes, is expressed as

$$D = q/A \quad (4)$$

where  $q$  is the stored charge on a capacitor and  $A$  is the area of one plate.

In electrostatics, the ratio of total electric flux density  $D$  to electric field strength  $E$  is called absolute permittivity,  $\epsilon$ , or dielectric constant,  $K$ , of a dielectric.

$$K = \epsilon = D/E = Cd/A \quad (5)$$

$$C = KA/d = \epsilon A/d \quad (6)$$

Permittivity of free space, sometimes referred to as the dielectric constant of a vacuum, is a constant  $\epsilon_0 = 8.85 \times 10^{-12}$  F/m.

#### Energy Storage, Work Done, and Energy Density

If the voltage across a capacitor of capacitance  $C$  is raised, the charging current  $i$  is given by Eq. (2) and the instantaneous power  $p$  (in watt) received by the capacitor is

$$p = vi = vC dv/dt$$

The change in energy stored by the capacitor (in joules) in time  $dt$  is

$$dW = vC dv/dt \cdot dt = Cv dv$$

The total energy stored by a capacitor when potential difference is increased from  $v = 0$  to  $v = V$  is given by

$$W = \int_0^V Cv dv = \frac{1}{2} Cv^2 \Big|_0^V = \frac{1}{2} CV^2 \quad (7)$$

$$W = \frac{1}{2} CQ; CQ^2 = \frac{1}{2} Q^2/C \quad (8)$$

where  $Q$  is the total charge stored in the capacitor when the applied voltage is  $V$ .

The energy is supplied by the source and corresponds to work done by charges moving through distance  $d$  in the electric field with force (in newtons) of attraction or repulsion  $F$ . This force of attraction between oppositely charged plates is written as

$$F = W/d = \frac{1}{2} Q^2/Cd = \frac{1}{2} Q^2/\epsilon A = \frac{1}{2} E^2 A$$

Volumetric energy density  $W^0$  (in units of joules per cubic meter,  $J/m^3$ ) for a capacitor with a dielectric thickness  $d$  and plate area  $A$  will be

$$W^0 = W/Ad = \left(\frac{1}{2} CV^2\right)/Ad$$

and since  $C = A/d$ ,  $E = V/d$ , and  $D/E = \epsilon$ , it follows that

$$W^0 = \frac{1}{2} \epsilon E^2 = \frac{1}{2} DE = \frac{1}{2} D^2/\epsilon \quad (9)$$

Energy storage capacity is often expressed in watt-hours (Wh) rather than joules and relates to the ampere-hour (Ah) capacity of the cell for a fixed system voltage such as a battery. Energy density or specific energy is defined as the energy that can be stored in a given weight (Wh/kg).

#### Properties of Dielectric Materials in Relation to Energy Storage Capacity

From Eq. (6), three parameters determine the storage capacity of a capacitor. These are

- The thickness of the dielectric material  $d$  that is, the closer the plates are, the greater the capacitance (and the stronger the electric force that exists between the plates),
- The area  $A$  of the plates since a larger area of plate accumulates more charges than a small plate, and
- The dielectric material as defined by the absolute permittivity  $\epsilon$  or dielectric constant  $K$ . Dielectric materials commonly used are air, vacuum, mica, paper, ceramic, plastic of certain types (polystyrene, polycarbonate, etc.), and metallic oxides. A good dielectric has
  - A very low electrical conductivity, hence a high insulating resistivity, to avoid leakage conduction which causes the dissipation of the stored energy in loss,
  - A high permittivity, and
  - A high electric strength to withstand large voltage gradient.

For determination of dielectric constant, all dielectric materials are compared to that of air as a reference. For example if the gap between two electrodes is filled with any other dielectric material, the capacitance is multiplied by a factor, known as the dielectric constant.

The absolute permittivity of any medium may be expressed in terms of  $\epsilon_r$  as:

$$\epsilon = \epsilon_r \epsilon_0$$

where  $\epsilon_r$  is the dimensionless relative permittivity of the medium. It is given by

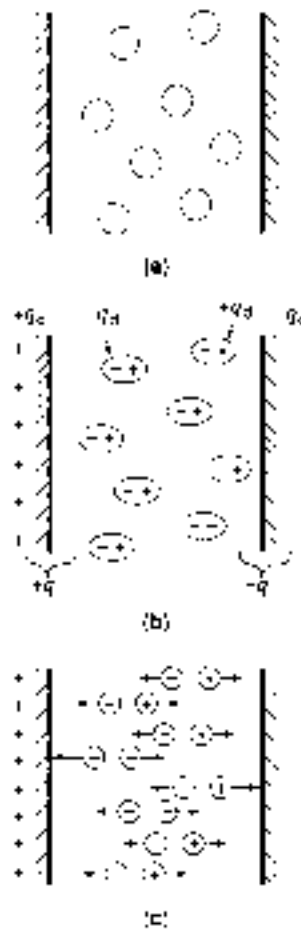
$$\epsilon_r = \frac{\text{(flux density of the field in the dielectric)}}{\text{(flux density of the field in the vacuum)}}$$

**Dielectric Strength.** The thickness of the dielectric in a capacitor determines its voltage rating. Before the voltage is applied to the electrodes, the molecules of the dielectric material are neutral and unstrained. As the voltage is raised from zero, these molecules under the influence of electric field, rotate, stretch, and separate from each other and subsequently orientate their negative and positive charges in opposite directions. This is referred to as polarization of the dielectric. If the voltage is increased beyond the maximum voltage gradient of a dielectric, the dielectric breaks down (fails to insulate). The result of break down is a crack or puncture in the material which causes a flow of current in the dielectric. These processes are shown in Fig 3.

The maximum voltage gradient is known as the dielectric strength (in V/m). This is given by

$$E_{max} = V_{max}/d$$

In high voltage engineering where capacitors are subject to high pulse powers, the problems associated with areas that are under high stress can be overcome by using a high permittivity material and avoiding sharp corners in conductors where the gradient tends to be high. In capacitor bushing, the use of an intersheath maintains the voltage at a suitable level. These improvements in the dielectric performance are called grading. Table 1 shows the dielectric constant and dielectric strength of various insulating materials.



**Figure 3.** Polarization and breakdown in a dielectric material. (a) No p.d.; dielectric molecules unstrained. (b) p.d. applied; molecules polarized. (c) p.d. increased; molecules disrupted.

**Dielectric Absorption and Hysteresis.** In the previous section it was mentioned that the dielectric molecules under the influence of an electric field become polarized and the effect of this is that some apparent charge is stored within the dielectric of the capacitor. This phenomenon is called absorptive capacitance of the dielectric. If the power supply is alternating, each reversal of voltage will cause a reversal of polarity and electrical energy from the supply will be converted into heat within the dielectric. This loss is referred to as dielectric hysteresis.

**Table 1. Properties of Dielectric Materials**

| Materials                | Dielectric Constant (K) | Dielectric Strength (V/cm) |
|--------------------------|-------------------------|----------------------------|
| Air                      | 1                       | 32,000                     |
| Paper (oil)              | 3-4                     | 600,000                    |
| Paper (pressboard)       | 2                       | -                          |
| Cotton tape (rubbered)   | 2                       | -                          |
| Polyethylene             | 2-3                     | -                          |
| Mica                     | 4-8                     | 720,000                    |
| Glass                    | 5-10                    | 80,000                     |
| Porcelain                | 7                       | 380,000                    |
| Titanates                | 100-200                 | 40,000                     |
| Ceramics                 | 6-1000                  | -                          |
| Polyvinyl chloride (PVC) | 3-8                     | -                          |
| Ethylene propylene       | 2.8-3.5                 | -                          |

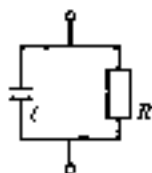


Figure 4. Equivalent circuit of capacitor with leakage.

**Dielectric Leakage and Conduction Currents.** In general an dielectric is a pure insulator and every dielectric possesses a very high resistance in excess of 100 MΩ. Hence, the number of free electrons is low. If a voltage is applied between the plates, those free electrons in the insulator drift from cathode to anode. This is known as conduction current, which produces power loss within the dielectric. The presence of conduction current will be evident, if the applied voltage across a charged capacitor is maintained and the current settles at a small constant value, instead of falling to zero.

Another current which flows between the two plates is the leakage current. This current flows over the surface of the dielectric. The magnitude of the current depends on the applied voltage, the dielectric material, but not its resistivity, and the moisture content of the dielectric and the air between the two plates. An equivalent circuit of such a capacitor is shown in Fig. 4.

Although the effect of leakage current within a circuit is very small, if the voltage across a charged capacitor is measured regularly, decrease in voltage can be noted although with a very large time constant.

The dielectric within a capacitor has a resistance  $R$ , which is given by

$$R = d\rho/A \quad (10)$$

where  $\rho$  is the resistivity of the leakage path (in ohm-meters),  $d$  is the dielectric thickness (in meters), and  $A$  is the area of one plate (in square meters).

Combining Eq. (6) for capacitance with Eq. (10) gives

$$R = \epsilon/\epsilon_0 C$$

where  $\epsilon$  is the dielectric constant.

If a voltage  $V$  is applied across a capacitor which has a resistance  $R$  between its two plates, a leakage current  $I$ , will flow. The value of this current is determined from the following relation

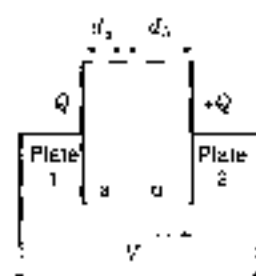
$$I = V/R$$

or

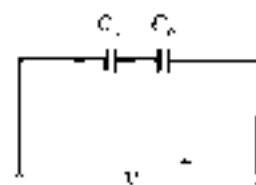
$$I = CV/\epsilon_0$$

This indicates that the leakage current of a capacitor is proportional to its capacitance value. The power loss due to this current is  $I^2R$ , which is dissipated as heat in the capacitor.

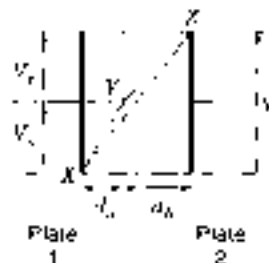
**Composite Dielectric Capacitors.** Figure 5(a) shows that the space between plates 1 and 2 is filled by dielectric  $a$  and dielectric  $b$  with a thickness of  $d_a$  and  $d_b$ , respectively. Assume that



(a)



(b)



(c)

Figure 5. (a) Vector diagram of pure capacitor. (b) Vector diagram of imperfect capacitor.

$E_a$  and  $E_b$  are electric field strengths in dielectrics  $a$  and  $b$ . If the relative permittivities of dielectrics  $a$  and  $b$  are  $\epsilon_a$  and  $\epsilon_b$ , respectively, electric field strength in dielectric  $a$  is

$$E_a = V_a/d_a = D/\epsilon_a\epsilon_0$$

and electric field strength in dielectric  $b$  is

$$E_b = V_b/d_b = D/\epsilon_b\epsilon_0$$

This capacitor, as shown in Fig. 5(b), may be regarded as equivalent to two capacitances,  $C_a$  and  $C_b$ , connected in series. The electric field strengths  $E_a$  and  $E_b$  are represented in Fig. 5(c) by the lines  $XY$  and  $YZ$ , respectively. If the dielectric between plates 1 and 2 is homogeneous, the electric field strength is the  $YZ$  line.

#### Connection of Capacitors

An individual capacitor may not be sufficient for applications requiring large amounts of energy storage. In other cases, the voltage rating of the individual capacitor may not be large enough to cope with the system voltage. In such examples, capacitors can be connected in parallel and series combinations.

**Capacitors in Parallel.** If a bank of capacitors of capacitance  $C_1, C_2, C_3, \dots, C_n$  are connected in parallel and the potential

difference of each is raised to  $V$  volts, the total charge is the sum of the individual charges.

$$Q_{total} = Q_1 + Q_2 + \dots + Q_n$$

or from Eq. (1)

$$C_{total}V = C_1V + C_2V + \dots + C_nV$$

The equivalent capacitance is therefore

$$C_{total} = C_1 + C_2 + \dots + C_n$$

The capacitance of  $n$  capacitors in parallel is the sum of their individual capacitances.

**Capacitors in Series.** For  $n$  capacitors connected in series, the total potential difference is the sum of their individual voltages.

$$V_{total} = V_1 + V_2 + \dots + V_n$$

The charge  $Q$  on each capacitor is the same when connected in series and again using Eq. (1)

$$Q = C_{total}V = Q/C_1 = Q/C_2 = \dots = Q/C_n$$

The equivalent capacitance is therefore

$$1/C_{total} = 1/C_1 + 1/C_2 + \dots + 1/C_n$$

#### Graphical Representation of Voltage, Charge, and Current

**In a Direct Current Circuit.** If a direct voltage  $V$  is applied to a capacitive circuit which contains a very small resistance, the capacitance  $C$  must accept a charge  $Q = CV$  immediately, resulting in a very large current flowing for a very short period. If the voltage across the capacitor is increased or decreased uniformly the charge  $Q$  will follow these changes with a constant charging or discharging current. Figure 6(a) shows these changes graphically.

**In an Alternating Current Circuit.** If an alternating voltage  $V = V_{rms} \cos 2\pi ft$  with a constant amplitude and frequency is applied to a circuit containing a capacitance  $C$  and a very small resistance, the steady-state relations for charge and current are expressed as

$$Q = Q_{rms} \cos 2\pi ft$$

where, from Eq. (1),  $Q_{rms} = CV_{rms}$  and

$$I = dQ/dt = d(CV_{rms} \cos 2\pi ft)/dt = -2\pi f(CV_{rms} \sin 2\pi ft)$$

respectively. All quantities are sinusoidal in shape. The maximum current is given by

$$I_{rms} = 2\pi f(CV_{rms}) = \omega CV_{rms}$$

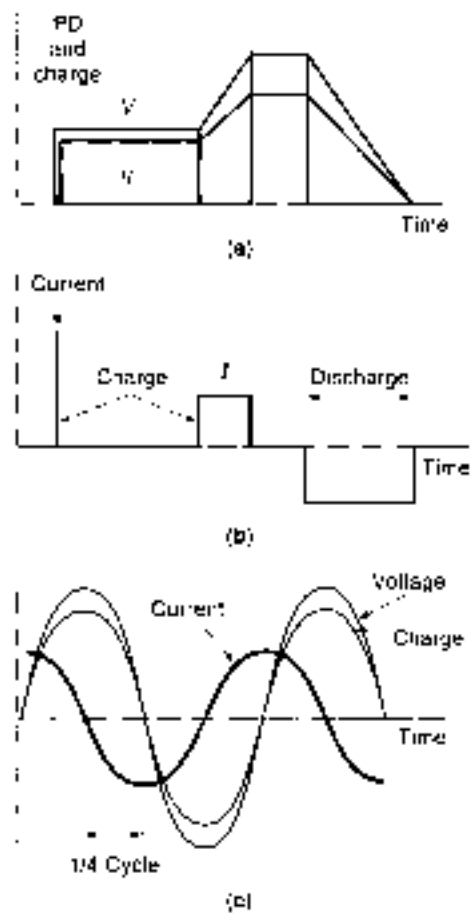


Figure 6. (a) Composite dielectric capacitor. (b) Equivalent circuit. (c) Potential distribution.

and the root mean square (rms) current is

$$I_{rms} = \omega CV_{rms} \quad \text{or} \quad I_{rms} = V_{rms} / (1/\omega C) = V_{rms} / X_c$$

where  $X_c = 1/\omega C$  is the capacitive reactance.

The graphical representation of voltage, charge, and current is shown in Fig. 6(b) while Fig. 6(c) shows the vector diagram of a pure capacitor.

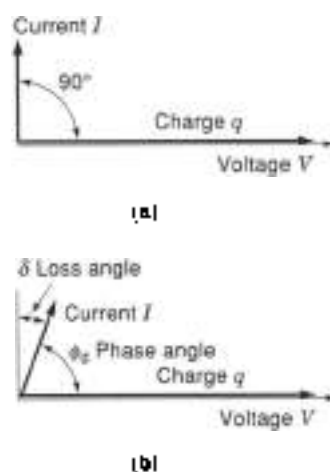
**Losses Across a Capacitor.** If a voltage is applied across a perfect dielectric, no dielectric loss can be detected and the induced capacitance current  $I_c$  leads the voltage by the phase angle  $\phi = 90^\circ$ , as shown in Fig. 7(a). As a perfect dielectric cannot be achieved in practice, a small current  $I_d$  that is in phase with voltage  $V$  will exist. The summation of these two current vectors, as shown in Fig. 7(b), gives the current vector  $I$  that leads the voltage by the phase angle  $\phi_d = 90^\circ - \delta$ . The cosine of the phase angle  $\phi_d$  is the power factor of the dielectric, and  $\delta$  is the dielectric loss angle. For a good dielectric,  $\phi_d$  is close to  $90^\circ$ .

With an ideal capacitor, the current is given by

$$I = I_c = \omega CV$$

With a nonideal capacitor, the loss is

$$P_{loss} = VI \cos \phi_d = VI \cos(90^\circ - \delta) = VI \sin \delta = \omega CV^2 \sin \delta$$



**Figure 7.** Relation between capacitor voltage, current and charge for (a) variable dc conditions, (b) ac conditions.

When  $\delta$  is small, it is expressed in radians and the power loss approximation is

$$P_{\text{loss}} = CV^2\delta$$

The losses of the capacitor are also expressed by the dissipation factor, which is defined as

$$\tan \delta = \omega R_C C$$

where  $R_C$  is the dielectric resistance or the resistance between the two plates.

#### Classification of Conventional Capacitors

Capacitors are often classified according to the material used for their dielectric. The main types include air, mica, paper, plastic, and ceramic and are used in electronic circuits where small value capacitors (farads, nanofarads, or picofarads) are normally required. Electrolytic capacitors have traditionally been used for storing large amounts of energy, particularly in electronic power supplies. In power engineering, where more energy storage is required, capacitor values are often quoted in farads. More recently, double layer capacitors (supercapacitors) have become available offering the potential for storing significantly larger quantities of energy than previous types although they are currently limited to low voltage applications.

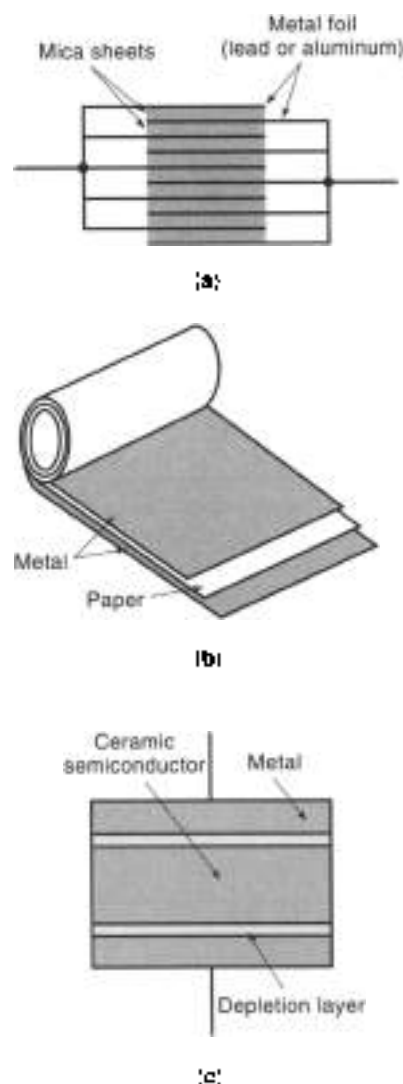
Capacitors are generally manufactured to have a fixed value of capacitance but variable capacitors are used in some specific application. Producing capacitance values within close tolerance is quite difficult particularly for larger values.

**Air Capacitors.** These consist of two metal plates (i.e., aluminum). The capacitance of such capacitors is generally low, typically 1 pF to 500 pF. This is because their plates are spaced far enough apart to prevent arcing. In variable air capacitors one of the plates is fixed, the other variable. Variable air capacitors are used where a variable capacitance and low losses are needed; typical examples are in radios and other electronic circuits.

**Mica Capacitors.** These capacitors consist of two metal foil plates with a sheet of thin mica between them. In their con-

struction, alternate layers of metal foil and mica are clamped tightly together. Usually the whole capacitor is impregnated with wax, to exclude moisture, and placed in a bakelite case. For larger values, several layers of plates and thin sheets of mica dielectric are used. Mica is a good insulator but is expensive and is not used in capacitors above 0.1 F. In newer types of mica capacitors, a thin layer of silver is sputtered on both sides of the mica dielectric to form the plates. The capacitance of these capacitors is stable and does not change with temperature or age. They have high working voltage ratings. These capacitors can be manufactured as fixed or variable capacitors and are normally used in high-frequency circuits. A typical construction is shown in Fig. 8(a).

**Paper Capacitors.** A typical paper capacitor is shown in Fig. 8(b). The electrodes are layers of metal foil interleaved with paper. The length of the roll corresponds to its capacitance. The paper is usually impregnated with oil or wax and is placed in a plastic or aluminum container for protection. These capacitors are commonly used in the power circuits of household appliances. Paper capacitors up to 1 F are made in various working voltages. The capacitance of these capacitors changes



**Figure 8.** (a) Mica capacitor; (b) Paper capacitor; (c) Ceramic capacitor

with temperature and they deteriorate faster than most other types of capacitors.

**Plastic Capacitors.** Plastic capacitors such as polyester, polystyrene, and Teflon are relatively new as capacitor dielectrics. They are manufactured in very thin films, and metallized on one side. Two films are then rolled together, similar to the construction of paper capacitors. These capacitors can operate well under conditions of high temperature, have high working voltages (rating i.e., a few thousand volts), and their leakage resistance is high, around 100 M $\Omega$ .

**Ceramic Capacitors.** These are made in various shapes and sizes. All have basically the same construction. For example, for smaller values, a thin ceramic dielectric is coated on both sides with a metal and they are usually cup or disc shaped (see Fig. 8(c)). Larger values are obtained by stacking up these ceramic layers. Each layer is separated from the next by more ceramic; these are normally a tube shape. In both arrangements, the plates are connected by electrodes and a final coating of ceramic is then applied to the outside to form a solid device for protection.

Certain ceramic materials such as compounds of barium titanate have a very high permittivity, thus enabling a very small separation between the plates and capacitors of high capacitance to be made from relatively small physical size. Ceramic capacitors are available in the range 1 pF to 1 F with a high working voltage rating up to a few thousand volts. The leakage resistance is typically 1000 M $\Omega$ . Ceramic capacitors are used in high-temperature situations and in high-frequency applications.

**Polarized or Electrolytic Capacitors.** According to the dielectric used, these capacitors are characterised as wet electrolytic capacitors or solid electrolytic capacitors. Wet electrolytic capacitors are generally made of two metal foil sheets (usually of aluminum) separated by a layer of paper saturated with a chemical liquid, such as ammonium borate, called the electrolyte as shown in Fig. 9(a). The foils and the paper are rolled up together and sealed in a container. To determine the polarity of the capacitor, a dc voltage is applied between the two foils. The current flow causes a thin layer of aluminum oxide to develop on one foil sheet forming the positive electrode. The absorbent paper between the two metal foils is a conductor and does not act as a dielectric. The oxide layer is the dielectric. As the thickness of an oxide layer is small, for example for a working voltage of 100 V, only about 0.1 $\mu$ m, a high capacitance in the range of many thousand microfarads is achievable in a small space. The typical working voltage range is 6 V to 500 V.

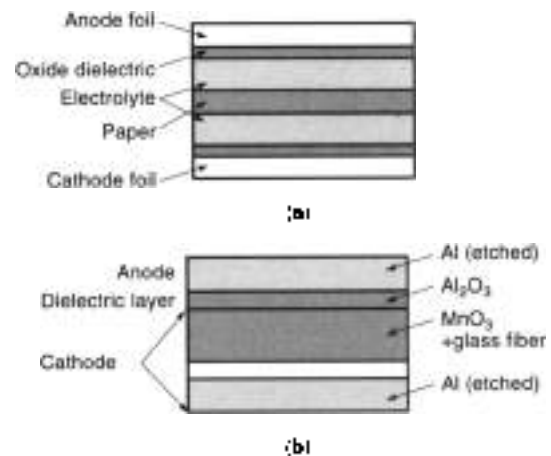


Figure 9. (a) Wet electrolytic capacitor. (b) Solid electrolytic capacitor.

The main disadvantage of this type of capacitor is that the insulation resistance is relatively low and that they must only be used where the applied voltage is direct. The most commonly used type is aluminum electrolytic capacitors. These capacitors are mainly used where a very large capacitance is needed, such as in rectifiers for reducing the fluctuating direct voltages. In ac systems, these can only be safely used if two capacitors (see Fig. 9(b)) are connected back to back, negative to positive and positive to negative.

Presently, solid electrolytic capacitors are much smaller in value but they do not possess some of the disadvantages of the wet electrolytic capacitors. For tantalum electrolytic capacitors, the wet electrolyte is replaced by layers of manganese dioxide and graphite. The anode plate consists of pressed, sintered tantalum powder coated with an oxide layer which forms the dielectric and the cathode is made of silver or copper plate. The layers of manganese dioxide and graphite have electronic conduction rather than the ionic conduction of the liquid electrolyte in the wet polarized capacitors. In these capacitors, the layer of manganese dioxide is coated on the oxide layer and the layer of graphite forms the connection with the cathode. Normally the whole structure is enclosed in a sealed container. Table 2 shows some advantages and disadvantages of traditional dielectric material used in capacitors.

#### Intrinsic Capacitive Energy Storage

Wherever an electric field exists, capacitance is formed. Thus the power distribution system itself possesses capacitance. Although the intrinsic capacitance of a section of the power supply system is relatively small, the energy stored influences the

Table 2. Types of Capacitor and Ratings

| Type         | Capacitance Range | Voltage Range | Comments                |
|--------------|-------------------|---------------|-------------------------|
| Ceramic      | 10 pF–1 F         | 50–1000 V     | Small, cheap            |
| mica         | 1 pF–0.01 F       | 100–600 V     | Good for RF             |
| Glass        | 10 pF–1000 pF     | 100–600 V     | Good for signal filter  |
| Tantalum     | 0.1 F–500 F       | up to 100 V   | Small, polarised        |
| Electrolytic | 0.1 F–0.2 F       | up to 500 V   | Polarised, MV filtering |
| Oil          | 0.1 F–20 F        | 200 V–10 kV   | Large, HV filtering     |

behavior of the system under fault conditions and cannot be ignored.

**Capacitance within Cables.** Cables may have one or more conductors within a protective sheath. The conductors are separated from each other and from the sheath by insulating materials.

For a single conductor cable, assume that a potential difference  $V$  exists between the conductor and the sheath of a cable. If the charge on the conductor and the sheath are  $+Q$  and  $-Q$  C/m of length according to Coulomb's law, the electric flux density at a radius of  $x$  (in meters) is

$$D = Q/2\pi x$$

Since from Eq. (5),  $E = D/\epsilon$

$$E = Q/2\pi\epsilon x$$

If the potential gradient at radius  $x$  is  $dV/dx$ , the potential difference between the conductor with radius  $R$  and the sheath with radius  $r$  is given by

$$V = \int_R^r E dx$$

or

$$V = (Q/2\pi\epsilon) (\ln R/r)$$

where  $\epsilon$  is absolute permittivity of the insulator.

The capacitance between conductor and sheath is

$$C = Q/V \text{ or } C = (2\pi\epsilon/\ln R/r) \quad \text{F/m}$$

**Three Conductor Belted Cable.** As shown in Fig. 10(a), in three conductor belted cables, there are two insulators, one is the conductor insulation with thickness  $T$  and the other is the belt insulation of thickness  $t$ . Because of these two insulations, there are capacitances of  $C_c$  between the conductors and capacitances of  $C_s$  between each conductor and the sheath as shown in Fig. 10(b). Further derivation of appropriate formulae is beyond the scope of this article but can be found in Ref. 16.

**Overhead Lines.** Calculation of all the parameters for determination of capacitance are again beyond the scope of the present discussion. A single-phase overhead line consists of two identical conductors 1 and 2 with radius  $r_1$  and  $r_2$  respectively, separated by a distance  $D$ . With a potential difference  $V_{12}$  between the two conductors, and charges of  $+Q$  and  $-Q$  C/m carried by conductors 1 and 2 respectively, the capacitance (in farads per meter) between conductors is

$$C_{12} = Q/V_{12} \text{ or } C_{12} = 2\pi\epsilon/\ln D_2/r_1 r_2$$

If  $r_1 = r_2 = r$ , then

$$C_{12} = \pi\epsilon/\ln D/r$$

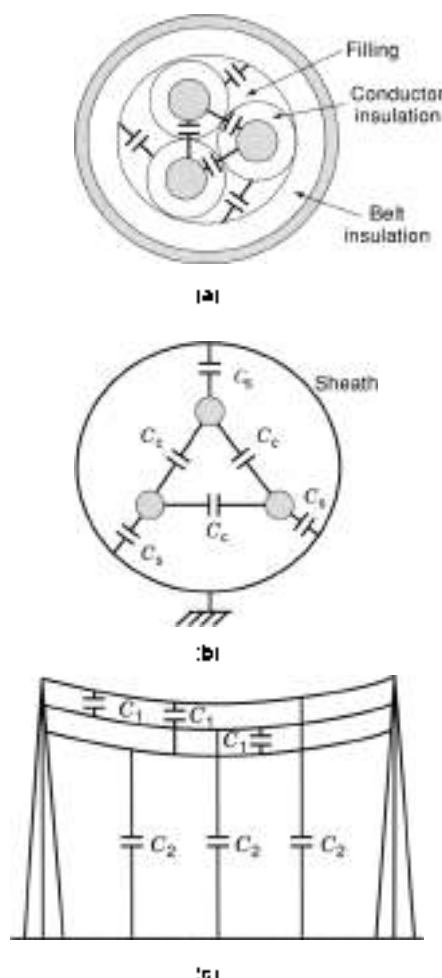


Figure 10. (a) Three conductor belted cable. (b) Equivalent circuit. (c) Representation of three phase overhead transmission line.

Figure 10(c) shows a representation of a three phase line with equal spacing  $D$ . The line to neutral capacitance is given by

$$C_n = 2\pi\epsilon/\ln D/r \quad \text{F/m}$$

If the spacing between the conductors is not the same,  $D$  is replaced by  $D_{eq} = (D_{12} \cdot D_{21} \cdot D_{31})^{1/3}$  in the above equation.

### Power Capacitors

The reactive power of capacitors is used to improve power factor and voltage, thereby reducing the losses, in power supply systems. If capacitors are connected in series with the line, the reactive power is proportional to the square of the load current whilst for shunt (parallel) connected capacitors, it is proportional to the square of the voltage.

The dielectrics of power capacitors are usually made of high purity annealed aluminium foil or metal spray. The dielectric is made of paper, mixed paper-plastic film, typically polypropylene, or plastic film. These dielectrics are designed to have a working voltage gradient of 10 to 50 MV/m.

Normally, these capacitors are in the form of banks and can be connected in a number of configurations such as grounded star, ungrounded star, double star with neutral floating, double star neutral grounded, and delta. For high voltage, star



connections are normally used. In star connection banks, the neutral of the capacitors is grounded only if the system or substation transformer is effectively grounded. The capacitors may be of the switched and nonswitched type, depending on the loading, voltage, etc.

On long and heavily loaded 11 kV distribution systems, the shunt capacitors are usually pole mounted. If a number of capacitor banks are used in parallel, it may be necessary to use series reactors to limit the inrush current.

With 0.415 kV distribution systems, capacitors are installed on individual lines or at consumer loads to reduce system losses and improve the system voltage. These also provide a reduction in kilovolt-ampere demand. For example the capacity (kVAR) required to improve the power factor of an existing kilowatt system, say from  $\cos\phi_1$  to a desired power factor  $\cos\phi_2$ , can be calculated from the following equation:

$$\text{kVAR} = \text{kW}(\tan \phi_1 - \tan \phi_2)$$

Also, where there is variation in load demand, the voltage at various buses must be controlled in order to keep the receiving end voltage ( $V_R$ ) at a specific value within a permissible band of voltage variations. To achieve this, a local VAR generator which consists of a bank of three phase static capacitors, as shown in Fig. 11, must be arranged. The VAR balance equation at the receiving end is

$$Q_R + Q_C = Q_L$$

Where  $Q_L$  is the a fixed amount of VARs drawn from the line by the load,  $Q_L$  is the varying VAR demand of the load, and  $Q_C$  is the varying VAR provided to the line to compensate for  $Q_L$ .

If  $V_R$  is the line voltage in kilovolts and  $X_C$  is the per phase capacitive reactance of the capacitor bank (star connection), the desired reactance of the capacitor bank can be determined by compensating for total inductive reactance of the load and even the transmission line. The expression for the VAR fed into the line can be written as

$$\text{kVAR} = \text{kW}(\tan \phi_r - \tan \phi_d)$$

that is

$$I_C = jV_R/\sqrt{3}X_C$$

or

$$jQ_C(\text{1-phase}) = jV_R^2/X_C(\text{1-phase})$$

Therefore  $Q_C(\text{1-phase}) = V_R^2/X_C$  and  $C = 1/\omega X_C$ .

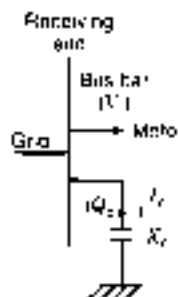


Figure 11. Power capacitor used as VAR generator at distribution busbar.

Local capacitive compensation similar to the one shown previously can be made automatic by using the signal from the VAR meter installed at the receiving end of the line. Nowadays, the capacitor bank switching is achieved using silicon controlled rectifiers (SCR).

### Supercapacitors

The term *supercapacitor* is commonly used to describe double layer capacitors. These are electrochemical energy storage devices that are able to store more energy per unit weight than traditional capacitors. For capacitance values larger than 1000 F, the term *ultracapacitor* is also used. For clarity, the term *supercapacitor* will be used in the following sections.

In supercapacitors, energy is stored by distributing charges across a relatively larger surface area than in conventional devices. They are also capable of delivering and absorbing energy at higher power rating than electrochemical batteries and possess far longer cycle life. Quite early in the development of supercapacitors, in 1994, charge and discharge cycles in excess of 500,000 repetitions were reported by Murphy and Krauer [17], some 500 times greater than a well-developed battery system. This emerging energy storage technology is increasingly being used for pulsed, high-energy, and high-power applications.

In a supercapacitor, an electronic conductor is immersed in an electrolytic solution. Ions from the solution naturally align themselves with electronic charges on the surface of the electronic conductor. This alignment leads to a layer of charges on the surface of the conductor and a layer of charges in the electrolytic solution, hence the term "double layer." In general, positive charges and negative charges are distributed facing each other at an extremely short distance in the boundary between the solid and liquid material [18]. The amount of charge stored in the double layer can be increased if a potential is applied between the solution and the conductive material. This can be achieved by placing a second electrode in the solution. When a potential is applied between the two electrodes, charge is absorbed in the double layer region on both electrodes. The basic configuration of a double layer capacitor, shown diagrammatically in Fig. 12, resembles a battery construction rather than a traditional capacitor. As shown in Fig. 12, this system is equivalent to two capacitors in series separated by an internal resistance.

Activated carbon fiber or powder is the most commonly used electrode material in supercapacitors because it is relatively inert and offers a very high surface area of conductive material. The manufacturing process of activated carbon can easily be modified to control the porosity, surface area, density, pore size, and pore volume of the fiber or powder. Control of the manufacturing process allows activated carbon material to be produced to match particular electrolytic solutions. To prevent short circuiting of the two electrodes, a porous separator which is permeable to ions but provides electrical insulation is placed between the electrodes. Both the separator and the electrodes are impregnated with an electrolytic solution (electrolytic solutions determine the maximum withstanding voltage; that is, the voltage at which electrolysis takes place). The separator allows ionic current to flow through the cell while preventing electrical conduction between the two electrodes. On the back of each active electrode, a current collecting plate is often added to reduce ohmic losses in the capacitor. If these plates

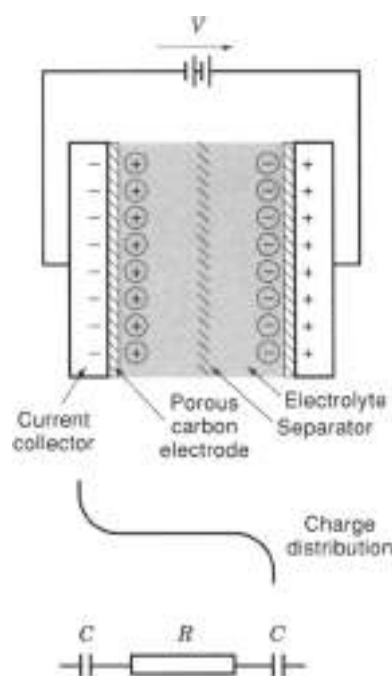


Figure 12. Basic configuration of a double layer capacitor

are amorphous, they can be used as part of the capacitor seal. Voltage is then applied to the polarized electrodes through the collector.

A supercapacitor is very much the same as a battery except that the nature of charge storage in the electrode active material is capacitive; that is, the charge and discharge processes involve only translation of ionic and electronic charges through electronically or ionically conducting domains, respectively. Energy is stored in a supercapacitor by charge separation within the micropores of the high surface area materials. These materials typically do not undergo chemical changes as do the materials in batteries. This storage mechanism is the primary reason supercapacitors are capable of extremely high cycle lives. Energy densities of supercapacitors are much higher than those of ordinary capacitors, but typically lower than those of advanced batteries. Supercapacitors store several hundred times the energy of conventional capacitors.

Supercapacitors can be classified with respect to the material being used in the electrodes (18). Currently, the materials being used include carbon-metal fiber composites; foamed (aerogel) carbon, activated, synthetic, nanofibric carbon; doped conducting polymer films on carbon cloth; and mixed metal oxides coating on metal foil (e.g., Ru/Ta-oxide). All of these materials are used with organic electrolyte. Other materials used are conducting polymers, noble metal oxides, or redox polymers (20).

The energy density of supercapacitors employing aqueous solutions (18) is limited by the low breakdown potential of water. In practice, this limits the voltage rating of such devices to less than 1 V. The low operating voltage also limits the power density of a capacitor stack as more cells must be connected in series to obtain the same operating voltage. Among the other disadvantages of aqueous-based capacitors, the highly corrosive nature of these solutions also necessitates the use of very expensive materials within the structure.

Organic-based electrolyte solutions offer a much higher operating voltage than aqueous-based systems (21). Cell voltages as high as 2.5 V to 3.0 V are achievable, leading to increases in both energy and power capability. These solutions also allow aluminum to be used in the composite electrodes, current collector, and cell casing, resulting in greatly reduced material cost of the capacitor.

Other research has focused on a class of materials that combine the characteristics of a capacitor and a battery (22,23). In redox supercapacitors, faradic charge transfer takes place as in a battery. In these devices, capacitance arises from faradic reaction and is known as *pseudocapacitance*. Redox supercapacitors store more charge than a conventional double layer capacitor and discharge more quickly than a conventional battery. The versatility of the electronically conducting polymers (ECP) enables different configurations to be used. Arhizann et al. (24) report three different redox supercapacitor configurations with an increasing charge storage capacity and operating potential range. These are

- A symmetric supercapacitor based on a *p*-doped ECP.
- An asymmetric supercapacitor based on two *p*-doped ECPs which are dopable at different potential ranges.
- A symmetric supercapacitor with a *p*- and *n*-doped ECP.

In the so-called Evans hybrid capacitor, the cathode in an electrolytic capacitor is replaced with a large capacitance value electrochemical cathode. Because the electrochemical requires little volume, available space is used to increase the size of the anode. The resulting capacitor has several times the energy density of the original.

**Supercapacitor Technology.** In 1991, the US Department of Energy (DOE) established a program to develop and evaluate supercapacitors as an enabling technology for electric/hybrid vehicles. Near-term and advanced goals for energy density (5 Wh/kg and 15 Wh/kg), power density (500 W/kg and 1600 W/kg), and volumetric energy density (11 Wh/L) were set. Development, testing, and evaluation activities have been coordinated by the Idaho National Engineering Laboratory (INEL) from where a substantial collection of publications has emerged, particularly from Burke (now Institute of Transport Studies, Davis University of California) including, for example, Refs. 25 and 26. Progress within the program to 1997 is reviewed by Murphy et al. (27).

A number of industrial groups have supercapacitor development programs including, in Japan, Panasonic (28) and, in the United States, Maxwell Laboratories Inc. (29,30), Argonne National Laboratory (23), and SAFT (31).

Maxwell Laboratories, Inc. has fabricated nonaqueous, carbon-based, 24 V supercapacitors. With eight cell, bipolar stacks they have achieved specific energies of 4.5 Wh/kg and specific powers in excess of 1 kW/kg.

At Idaho National Engineering Laboratory, as reported by Burke (19), research on a 3 V cell has resulted in an energy density of 5 Wh/kg at a constant power density of 2.5 kW/kg. In his paper, Burke (19) suggests that up to an energy density of 200 Wh/kg with low cost and long life cycle can be achieved.

Presently, supercapacitors are manufactured in the form of single cell (spiral wound) and multicell (prismatic and bipolar). The bipolar cell arrangement is used to increase the operating

voltage of supercapacitors, in which individual cells are stacked in series. Because one side acts as the positive electrode and the other side acts as the negative electrode in an adjacent cell, this type of arrangement is called bipolar (H. J. Becker, US Patent No. 2,800,615, 1957).

The optimum design is to reduce ohmic losses. This is usually achieved by separating adjacent cells with only current collector plates. These plates must be nonporous so that no electrolytic solution is shared between cells (18,26).

With the prismatic cell (wound design) arrangement, a thin, large surface area cell forms a continuous winding curve around a central axis in a single cell housing (A. Yoshida and K. Imoto, US Patent No. 5,150,283, 1992). Blank et al. (32) report that bipolar design has increased efficiency over the wound design. This is because, in prismatic cell arrangement, cells are stacked in series and the resistance of the lead wires adds to ohmic losses.

In order to increase the efficiency of supercapacitors, the contact resistance which is the resistance between the particles in the electrode structure must be low. Also, the resistance must also be small between adjacent cells in stack structure. These two resistances are called internal resistance.

For large capacitance, a large surface area is needed, whilst to increase the working voltage, a solution with high potential breakdown is necessary. The organic-based electrolytic solutions provide the best performance. These solutions possess a much larger operating voltage range than aqueous-based systems (33).

Researchers at Lawrence Livermore National Laboratory (LLNL) have applied carbon aerogel material to an aerocapacitor. Carbon aerogel is one of the world's lightest materials and its large surface area can fulfill the need for compact ultracapacitor size. Aerocapacitors are inexpensive and have shown potential to be easily produced. Because of the large surface area provided by an aerogel, it can deliver at least 50 times more power in a given space than batteries. In addition, its conductivity is higher than that of capacitors made from other forms of carbon and carbon powders because of a honeycomb structure. LLNL researchers have developed carbon aerogels the size of grapes with effective surface areas the size of two basketball courts. The aerocapacitors have shown capacities of up to 40 F/cm<sup>3</sup> and excellent performance at temperatures as low as -30 °C. Power densities have been shown to be more than 7 kW/kg. Aerocapacitors also hold auto-stored energy over a period of weeks (21).

Recently, Lassegues and co-workers (21) studied the possibility of replacing the liquid electrolyte by a proton conducting polymer electrolyte. Their main interest was to build an all-solid thin-film device free from problems of leaks and corrosion. Their results so far suggest that the lifetime of such a supercapacitor is unlimited. Rudge and co-workers (20) demonstrated that electrochemical capacitors with a combination of high capacitive energy density and low material cost can be constructed using conducting polymers. They have demonstrated that energy densities of up to 39 Wh/kg can be achieved using a combination of both n- and p-doped conducting polymers for a capacitance ranging up to 1000 F.

Research suggests (Refs. 11–14 of Ref. 34) that for energy densities considerably greater than 5 Wh/kg in combination with low resistance (<0.01 Ω cm<sup>2</sup>) devices may require pseudocapacitance which requires an active electrode material. Materials could be carbon composites, mixed Ru/Ta-oxide, conduct-

ing polymers, redox powders, or fibers or other high surface area materials.

Supercapacitors offer considerable potential for energy storage systems; however, most commercially available supercapacitors, designed for much lower power applications, are small and store only several hundred joules (tens of milliwatt hours). The largest commercially available device, manufactured by Panasonic, stores 1.1 Wh of energy (28). For power supply applications, high voltage stacks (many supercapacitors connected in series) capable of much higher quantities of energy must be developed.

**Basic Equivalent Circuit of a Supercapacitor.** The system of two electrodes in an electrolytic solution represents the basic constituents of a supercapacitor. This system is equivalent to two capacitors in series separated by an internal resistance. The capacitance of each electrode can be represented by the following equation.

$$C = K_r/dA$$

Where  $K_r$  is the effective dielectric constant in the interface region,  $d$  is the distance of charge separation across the double layer, and  $A$  is the active interfacial surface area of the solution and electrode. Typically  $K_r \cdot d$  is in the range of 10 F/cm<sup>2</sup> to 30 F/cm<sup>2</sup>.

The electric double layer capacitor has a very large capacitance. Hence it cannot be measured by using an LCR (inductance/capacitance/resistance) meter as would be the case for conventional capacitors. To determine the capacitance of a supercapacitor with a maximum usable voltage  $E_r$  (in volts), the charging time constant  $\tau$  of capacitor  $C$  is measured and the capacitance found from

$$C = \tau/R_s$$

where  $\tau$  represents the time in seconds elapsed until  $V_r$  reaches a value of  $(1 - 1/e) E_r$ , and  $R_s$  is the series protection resistor which is fixed for a specified type of the supercapacitor.

**Integration of Supercapacitor Energy Storage.** In the case of an electrochemical battery, energy is absorbed and delivered at approximately constant voltage. For a supercapacitor, as for any other capacitor, the voltage across the device varies as energy is stored or released. A power electronic converter must be used in combination with the supercapacitor to control the flow of charge (current) into and out of the supercapacitor whilst maintaining the voltage presented to the system within which the energy store operates (35,36). The form of power electronic converter which can achieve bidirectional energy flow between a supercapacitor and the main power system and load is shown in Fig. 13. Power devices SW1 and SW2 control the energy flow into and out of the supercapacitors. Energy is stored in the supercapacitors by appropriate switching of power device SW1 (forward converter mode (37)) and returned to the link with the main power supply system by appropriate switching of power device SW2 (boost converter mode (37)).

The exact configuration of supercapacitor, power electronics, and load depends on the application. For example, in an electric/hybrid vehicle, by appropriate control of the power electronic converter, energy from the supercapacitors can be used to accelerate the de motor which drives the wheels. The

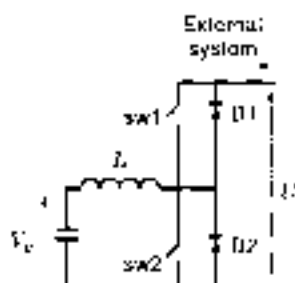


Figure 13. Power electronic converter for bidirectional power control in capacitor energy storage.

motor would be connected to the supercapacitor system by a second power converter. The main system power source would include an internal combustion engine-driven generator in an electric/hybrid vehicle or an electrochemical battery in a purely electric vehicle. In such a system, regenerated energy can be returned to the supercapacitors during braking. For an uninterruptible power supply application, a second power converter would be required to feed an ac load. The discharge of the supercapacitors can be controlled such that the link between the two converters can be maintained at a set voltage, hence mimicking the action of a battery.

Figure 14 shows results from a test in which a supercapacitor converter is controlled to maintain the link voltage at 10 V. The supercapacitors (four 3 V, 60 F devices connected in series) are initially charged to about 8 V (top trace); that is, about 480 J of stored energy. The voltage to the load is maintained at 10 V for about 11 s (trace 2) and the load draws a current of about 2 A during the 11 s period (bottom trace). The energy returned from the supercapacitors during this time is 220 J. The system cannot control the link voltage beyond this point and the remainder of the energy is delivered with a reducing link voltage. With optimization of the system, the load voltage can be

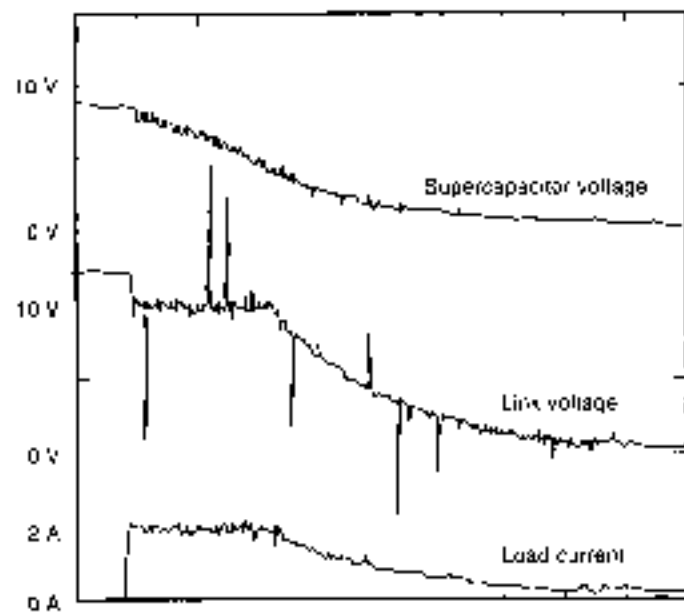


Figure 14. Practical demonstration of energy storage discharge: (a) Top trace: supercapacitor voltage; (b) Middle trace: link voltage; (c) Bottom trace: load current.

maintained at the specified level for longer resulting in a more effective energy storage device. Optimization involves the application of sophisticated control algorithms and reduction in losses in the power converter (38).

## SOME APPLICATIONS

Capacitors can potentially be used as primary energy store in applications from portable equipment through low power lighting to meeting the problems of fluctuating energy demand in electrical power systems and electric/hybrid vehicles. The need now is to develop high-voltage versions capable of storing much larger quantities of energy.

For electric/hybrid vehicles, supercapacitors could accept and store high-energy pulses recovered from regenerative braking. Other applications would include pulse power sources for starting an i.c. engine or energization of a coil for the projection of a weapon. Capacitor storage could find application in most of today's electronic devices for storing electrical energy, for example in communications and medical applications.

When supercapacitors are available which are capable of high-voltage operation and peak power density around 500 W/kg to 700 W/kg, they could be used in power systems. The supercapacitor would receive energy from and deliver energy to a three phase ac electric grid network using a power electronic converter. Overall cost efficiency of the system will be the key to future success of capacitor energy storage.

## Capacitor Storage in Transport Applications

Capacitors are already extensively used for voltage smoothing and power factor correction purposes. There is now widespread interest in their use as intermediate energy storage devices in a variety of heavy duty applications, including electric and hybrid vehicles. The reason is their superior power density and cycle life relative to their nearest counterpart, the electrochemical battery. The hybrid vehicle provides a good example of an application where such properties are required. In a hybrid vehicle, operating under urban conditions, an energy storage device is required to provide for the rapidly fluctuating power demands made by the vehicle traction system and, in particular, to provide the capacity to absorb large transient levels of regenerated power during braking. Such energy storage devices could be subjected to several charge/discharge cycles per minute, which would render electrochemical batteries unsuitable for this purpose.

A disadvantage of capacitors as energy storage devices is their relatively low-energy density. However, in the intermediate storage applications mentioned here, high-energy density may not be a critical requirement, particularly, as in a hybrid scheme, where it can be combined with a high-energy low-power device, such as a generator set, to give the required ratio between energy and power levels. This ratio, which has the units of time, is a useful parameter in the specification of an energy storage system and is often referred to as the time constant, being the time for which it can deliver its rated power.

In hybrid vehicles, energy storage is required to provide for vehicle acceleration and deceleration, which, in urban conditions, takes place over a period of 5 s to 10 s. This suggests the required time constant and also indicates the unsuitability of electrochemical batteries for this application, which generally have charge times of at least 20 min. The suitability of various

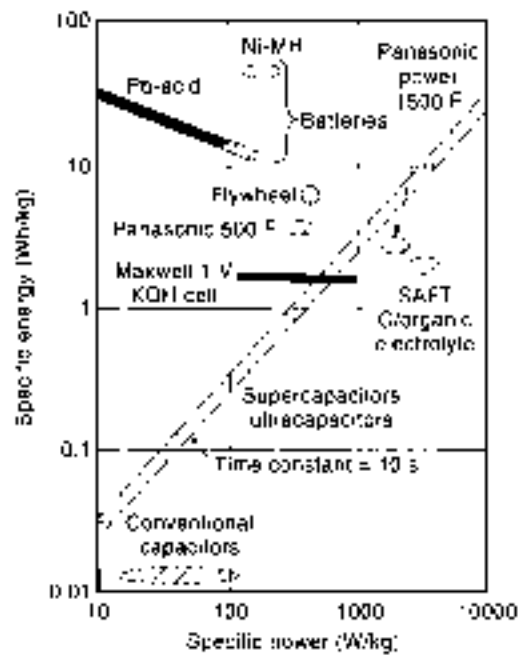


Figure 15. Ragone plot showing comparison of various forms of energy storage.

types of energy storage for a particular application can be illustrated by a Ragone plot of specific energy (Wh/kg) versus specific power (W/kg). Figure 15 shows such a plot for batteries and supercapacitors on logarithmic scales (25).

Projected onto this plot is a locus corresponding to a time constant  $T_c$  of 10 s. The power requirement for a 1 tonne vehicle is estimated to be 50 kW (supplemented by about 25 kW of primary power). In order that the mass of the energy storage should be no more than 10% of the total vehicle mass, the specific power should be at least 500 W/kg for adequate performance. This indicates that most projected ultracapacitors (i.e., in excess of 1000 F) would meet the specific power and energy requirements for hybrid vehicles and could be manufactured to have the required time constant so as to minimize the overall weight. Electrochemical batteries are shown to be lacking the required specific power and this is particularly the case during charging. Battery-based systems, therefore, tend to suffer a weight penalty, which is a disadvantage in mobile applications and can generally increase installation and maintenance costs.

#### Capacitors as Intermediate Storage

Capacitors compare favorably with other forms of energy storage, as is shown on Fig. 16, where similar data for a variety of forms, including i.e. engine/fuel tank systems, are plotted using linear scales. The line joining two energy sources forms a locus on the plot for a variable combination of the two. As an example, in conventional vehicle design, a fuel tank is combined with an i.e. engine to give a propulsion system with a time constant of about 2 h, representing the time taken to use up the fuel running the vehicle continuously at maximum power. Therefore, by use of a hybrid combination of a primary power supply of high specific energy with an intermediate energy storage system of high specific power (possibly based on the supercapacitor), power supply units can be designed with a specified time constant, in which weight is minimized.

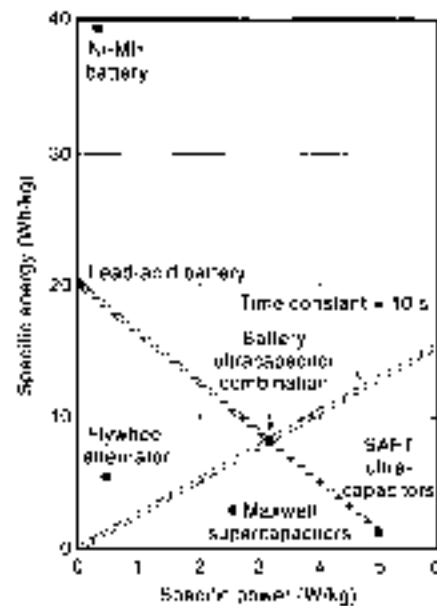


Figure 16. Design of energy storage systems for given time constant.

The incorporation of supercapacitors as intermediate energy storage, in combination with a primary power source, could lead to significant weight reductions, improved performance, and reduced maintenance costs in comparison with other forms of electrical energy storage. Only the flywheel/alternator system can provide comparable power density and cycle life. This has the advantage of a higher output voltage capability and that its ac output permits state of the art power conversion methods. This form of energy storage is already finding applications in hybrid bus propulsion systems. Supercapacitors potentially offer the opportunity for weight reduction but considerable development is required to provide controlled power conversion between the relatively low and varying voltages across the capacitors (typically 1.5 V to 3 V per capacitor) and the prime mover, without unacceptable power losses in conversion.

#### Pulse Power Units

Capacitors are also used for the generation of high-voltage and high-current pulses. In operation these capacitors are under stress of high electromagnetic and electric field strength and therefore need to have the ability to withstand high peak currents and voltages. For this reason they need to possess a high voltage breakdown strength, high dielectric constant, and low dissipation factor.

Controlled current pulse powers are already being used for nuclear and conventional firing sets, weapons, laser excitation, communication radar, welding, and for the actuator of air bags. High-voltage pulses are used in rail guns and particle accelerators. Another application of pulse power is in power engineering, where integrity of equipment such as transformers, switchgears, busbars, and cables which are installed in generation, transmission, or utilization systems are examined. Transient disturbances such as lightning strokes and switching operations are usually followed by a steep fronted travelling wave on the system. For example, if a voltage wave of this type reaches a power transformer, unequal stress distribution could appear along its windings which may lead to break down of the insulation system. The most severe test that equipment used

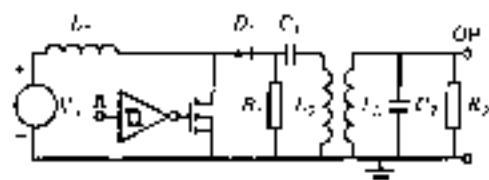


Figure 17. Simple high voltage pulse generator.

in power engineering will be subjected to during its lifetime results from the superimposition on the system of transient voltages.

There are a number of circuits available for the generation of these pulses. Figure 17 shows the basic circuit of a high voltage pulse unit which primarily consists of two coils,  $L_1$  and  $L_2$ , and two capacitors  $C_1$  and  $C_2$ , where  $C_1 \gg C_2$ . The high voltage side coil has a larger number of turns. The capacitance  $C_1$  is charged from a direct current and the frequency of the pulse is controlled by MOSFET and its associated triggering circuits. The pulse shape in terms of rise, decay, and duration is controlled by  $R_1$  and  $R_2$  as well as the electronic circuitry shown in Fig. 17.

The approximate voltage relationship is given by:

$$V_2/V_1 = \sqrt{qC_1/C_2}$$

where  $V_1$  is the maximum voltage to which  $C_1$  is charged,  $V_2$  is the maximum voltage to which  $C_2$  is charged,  $q$  is the efficiency of the energy transfer from the low voltage side to the high voltage side of the circuit, and  $C_1$  and  $C_2$  are the capacitances for  $C_1$  and  $C_2$ , respectively.

## SUMMARY

Capacitor storage is widely used for power conditioning where relatively low levels of energy storage are needed, and in certain applications where high power pulses are required. Electrolytic capacitors are available with capacitance up to the order of 1 F. However, supercapacitors of up to 1500 F are now under development. Once the problem of internal resistance and voltage rating can be overcome, these devices could well become the prime candidate for primary and intermediate energy storage applications in the future.

## BIBLIOGRAPHY

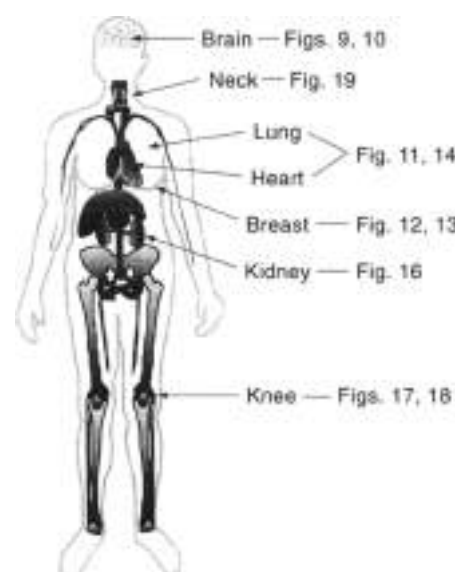
- W. M. Somerville, Operating experience of the Fair Isle wind turbine, *Proc. IEE, Part A*, 134 (5), 412-426, 1987.
- T. S. Davies and R. M. Mayer, Power quality and the interaction of WFGS with the grid, *Proc. Eur. Wind Energy Conf.*, Thessalonika, Greece, 1994.
- T. S. Davies, R. M. Mayer, and H. Nouri, Integration of wind turbines into L.V. networks, *Proc. Univ. Power Eng. Conf.*, University of Greenwich, London, 1995.
- T. S. Davies, H. Nouri, and C. M. Jefferson, Alternative Approaches to Wind Energy Conversion Systems, *Wind Eng. J.*, 10 (5), 265-271, 1995.
- P. Musgrove and G. Slack, Long-term performance modelling of a wind-diesel system with hydraulic accumulator storage, *Proc. 9th British Wind Energy Assoc. Conf.*, 1986, pp. 43-50.
- T. S. Davies, C. M. Jefferson, and N. Larsen, A regenerative flywheel drive for energy storage applications, *Proc. 22nd Inter-Soc. Energy Conversion Eng. Conf.*, "Energy-New Frontiers", Philadelphia, 1987.
- T. S. Davies, C. M. Jefferson, and R. M. Mayer, Use of flywheel storage for wind diesel systems, *J. Wind Indus. Aerodyn.*, 27, 157-165, 1987.
- T. S. Davies et al., Experience with the wind turbine flywheel combination on Fair Isle, *Proc. Eur. Wind Energy Conf.*, Glasgow, 1989.
- P. A. C. Medeiros, Development of a lightweight low cost flywheel energy storage system for regenerative braking applications, *Proc. 20th Inter-Soc. Energy Conversion Eng. Conf.*, Miami Beach, 1985.
- R. Dettmer, Spinning reserve, *IEEE Rev.*, 42 (1), 36-37, 1997.
- P. P. Acarley et al., Design principles for a flywheel energy store for road vehicles, *IEEE Trans. Ind. Appl.*, IA-32, 1402-1406, 1996.
- S. J. Chiang, S. C. Huang, and C. M. Liaw, Three phase multi functional battery energy storage system, *IEEE Proc. Electr. Power Appl.*, 142 (4), 275-284, 1995.
- C. DeWinkel and J. H. Lammie, Starting power for critical loads, *IEEE Spectrum*, 30 (6), 39-42, 1993.
- H. T. Matthias, The search for high-temperature semiconductors, *Physics Today*, 24, 23-26, 1971.
- S. Tanaka, Research on high- $T_c$  superconductivity research in Japan, *Physics Today*, 40 (12), 53-57, 1987.
- Z. A. Yamney and J. L. Bala, *Electromechanical Energy Devices and Power Systems*, New York: Wiley, 1994.
- T. C. Murphy and W. E. Kramer, US Dept. of Energy ultracapacitor development program for load leveling electric vehicle propulsion systems, *Proc. 16th Int. Seminar Double Layer Capacitors Similar Energy Storage Devices*, Finland Educational Seminars, 1994.
- M. Aoki, K. Sato, and Y. Kobayashi, Long-term reliability testing of electric double-layer capacitors, *IEICE Trans. Fundam.*, E77-A (1), 208-212, 1994.
- A. F. Burke, Characterisation of electrochemical capacitors for electric hybrid vehicle applications, *AIAA-94-375*, 1994, pp. 764-769.
- A. Kudge et al., Conducting polymers as active materials in electrochemical capacitors, *J. Power Sources*, 47, 69-107, 1994.
- J.-C. Lassegues et al., Supercapacitor using a proton conducting polymer electrolyte, *Solid State Ionics*, 77 (1), 311-317, 1995.
- C. Arduini, M. Mastrogatto, and L. Moneghelli, Performance of polymer-based supercapacitors, *Proc. Mat. Res. Soc. Symp.*, 369, 1995.
- D. R. Vissers et al., Materials for electro-chemical energy storage, *Electr. Hybrid Veh. Technol.*, 96.
- C. Arduini, M. Mastrogatto, and L. Moneghelli, Performance of polymer-based supercapacitors, *Proceed. Mat. Res. Soc. Symp.*, 369, 605-612, 1995.
- A. F. Burke and E. J. Donaghy, Ultracapacitors for electric and hybrid vehicles: A technology update, *Proc. 11th Int. Electr. Veh. Symp.*, Florence, 1992.
- A. F. Burke, Electric/hybrid super car designs using ultracapacitors, *Proc. Inter-Soc. Energy Conversion Eng. Conf.*, 1996.
- T. C. Murphy, W. E. Kramer, and R. B. Wright, Electrochemical capacitor developments, *Electr. Hybrid Veh. Technol.*, 96.
- Passante, Tentative specification, December, 1993.
- E. Blauk, C. J. Farabonchi, and A. P. Tripp, Ultracapacitors: The much needed boost, *Electr. Hybrid Veh. Technol.*, 95.
- E. Blauk, C. J. Farabonchi, and J. Diepernolle, Material and processes of ultracapacitors for high energy and high power applications, *Proc. 27th Int. AIAAPE Tech. Conf.*, 1995.

31. G. Chagnon, SAFT supercapacitors: Design and Performance. SAFT presentation, 37th Power Source Conf., 1996.
32. E. Blank, C. J. Farahmandi, and J. Dispennette, Materials and processes of ultracapacitors for high energy and high power applications, 27th Int. SAMPE Tech. Conf., 1995, pp. 769-780.
33. J. C. Laspegnes et al., Supercapacitor using a proton conducting polymer electrolyte, *Proc. Solid State Ionics Conf.*, VII: 1994, 311-317.
34. E. J. Dawydnlo, Ultracapacitors for electric and hybrid vehicles—A technology update, *Proc. 11th Int. Electric Veh. Symp.*, Florence, 1992.
35. M. Schmid, Electric double-layer capacitors in a hybrid vehicle, *Proc. 10th Int. Symp. Autom. Tech. Automation*, Florence, 1997.
36. T. S. Davies et al., DC-to-DC power conversion for supercapacitor energy storage system, *Proc. Universities Power Eng. Conf.*, 1996.
37. N. Mohan, T. M. Undeland, and W. P. Robbins, *Power Electronics—Converters, Applications and Design*, New York: Wiley, 1995.
38. T. Coetier and M. D. Mayer, Autonomy enhancement of a lead-acid battery EV—ultracapacitors contribution: A simulation study, *Proc. 10th Int. Symp. Autom. Tech. Automation*, Florence, 1997.

HASSAN NOURI  
 TERRY S. DAVIES  
 COLIN M. JEFFERSON  
 University of the West of England,  
 Bristol

## DIAGNOSTIC IMAGING

Radiology is divided into diagnostic radiology, interventional radiology, and radiotherapy. Diagnostic radiology encompasses not only X-ray imaging but all modes of noninvasive human imaging, which are listed in Table 1. Of these, the principal methods used since 1970 are planar X-ray methods (e.g. fluoroscopy, angiography, gastrointestinal contrast series, urography, myelography), X-ray computed tomography (CT),



**Figure 1.** This, a guide to the second half of this article, gives selected imaging results to demonstrate applications of the major diagnostic methods.

ultrasound (US), techniques of positron emission tomography (positron emission tomography (PET) and single photon emission computed tomography (SPECT)), nuclear magnetic resonance methods of magnetic resonance imaging (MRI), magnetic resonance angiography (MRA), and magnetic resonance spectroscopy (MRS) also known as chemical shift imaging. Currently 70% of all the conventional radiological procedures use X-rays. About 60% of the initial diagnostic procedures are X-ray based with only 8% nuclear medicine emission, 8% ultrasound, and 8% MRI. Advances in imaging speed, resolution, and data processing are rapidly changing the relative importance of these modalities particularly in areas such as MRI, fast X-ray CT, ultrasound, and emission tomography. These main methods are described in general terms below with more technical details presented in (1,2) and elsewhere in this encyclopedia. Following a description of these methods, major medical applications are highlighted for all of the major organ systems along with selected radiologic images (Fig. 1).

### X RAY

Presently the most widely used imaging modality is the X-ray transmission intensity projection, most commonly known as "the X-ray." The difference between the X-ray, a simple 2-dimensional projection image, and the X-ray computed tomograph, "the CT," is that the latter is the result of mathematically reconstructing an image of a slice through the body from multiple X-ray transmission projection images taken at multiple angles (usually equal) around the body.

X-rays are generated by the interaction of accelerated electrons with a target material such as tungsten. The electrons are produced by a heated cathode and accelerated by applying a voltage of about 100 keV between the cathode and a tungsten or molybdenum anode contained in a vacuum container called the X-ray tube or X-ray gun. The electron beam colliding with the anode releases X-rays characteristic of the target materials.

**Table 1. Diagnostic Imaging Methods Summary**

| Method  | Information   |
|---|---|
| (1) X-ray (contrast, X-ray CT, electron beam CT, spiral CT)                 | Electron density<br>Anatomical composition<br>Vascular lumens |
| (2) Ultrasound (US)<br>Doppler Ultrasound                                   | Acoustic impedance mismatches, Motion                         |
| (3) Emission Tomography (PET, SPECT)  | Radionuclide concentration, metabolism, receptor densities    |
| (4) Magnetic Resonance Imaging (MRI), Spectroscopy (MRS), Angiography (MRA) | Spin density, relaxation, diffusion, Chemical composition     |
| (5) Electrical Source Imaging (ESI)   | Brain and heart current sources                               |
| (6) Magnetic Source Imaging (MSI), Magnetoencephalography (MEG)             | Brain current sources   |
| (7) Electrical Impedance Tomography (EIT)                                   | Electrical conductivity                                       |
| (8) Optical Imaging (Spectroscopy)  | Attenuation, scattering, molecular state                      |

— Same as ESI

K-shell X rays from tungsten are about 70 keV and these as well as other X rays or photons are emitted from the X-ray tube placed one or more meters from the patient. The X-ray radiograph is usually a film, not unlike a photographic negative, which is darkened due to the interaction of the photons with the silver halide granules of the film. To enhance efficiency, a screen containing a phosphor is juxtaposed with the film. The electrons which interact with the phosphor release thousands of photons in the visible wavelength. The image reflects the number or intensity of photons reaching the film, and as is the case for a conventional negative, the greater the intensity, the darker the X ray. The intensity of photons transmitted through the body is modulated by the processes of Compton scattering and photoelectric absorption in tissue which are dependent on electron density and tissue elemental composition, respectively. These modulation processes are lumped into a simple attenuation coefficient. The intensity (or number) of photons arriving at a particular position,  $(x, y)$ , on the X-ray film is given as

$$I(x, y) = I_0 e^{-\mu z} \quad (1)$$

where  $I(x, y)$  is the photon intensity at position  $(x, y)$ ,  $I_0$  is the intensity from the X-ray tube,  $\mu$  is the attenuation coefficient (units of length<sup>-1</sup>), and  $z$  is the path length through the patient to the image position  $(x, y)$ . This equation applies to a situation of constant attenuation along  $z$ . The attenuation coefficient of lung, water, tissue, and bone differ. Thus the intensity arriving at the film is more generally

$$I(x, y) = I_0 e^{-\sum \mu(x, y, z) \Delta z} \quad (2)$$

where we divided the path  $z$  into intervals  $\Delta z$ . Here we assume parallel X-ray paths and no magnification; otherwise the image coordinates would be different from the object coordinates. The contrast or intensity difference between a lung tumor region and the surrounding normal tissue, as recorded on a conventional projection X ray, is related to the difference in number of photons projected through the tumor and through the parallel paths surrounding the tumor. This difference in number of photons would be about 23% for a 3-cm tumor, but the contrast visualized will depend on the image detector (e.g., silver halide film, solid state detector).

Recall the fact that as the number of photons decreases, the less the exposure of the X-ray film and the "whiter" the image in that region. X-ray imaging is mainly an anatomical procedure. The absorption of X rays due to differences in elemental composition has an important effect for imaging bone or calcium deposits because this absorption process is proportional to the atomic number. To provide image contrast between the blood vasculature and surrounding tissue, a dense fluid with elements of high atomic number (e.g., iodine, barium) can be injected or swallowed during the X-ray exposures. The movement through the body vasculature of a "contrast agent" such as an iodinated compound can be visualized by acquiring a sequence of X rays. The iodine or barium agent absorbs photons more than blood and tissue because the density is higher and the elements iodine and barium have a high atomic number giving rise to more photoelectric absorption.

## X-RAY CT

X-ray transmission computed tomography (X-ray CT) gives anatomical information of the tissues mainly based on the density and elemental composition in specific regions as derived by manipulation of the projections of X rays through the body at multiple angles. It is useful to examine the log of the ratio of incoming intensity to the exiting intensity to parameterize these projections

$$P(x, y) = \ln \frac{I_0}{I(x, y)} = \sum_{z=0}^{\text{Detector}} \mu(x, y, z) \Delta z \quad (3)$$

Note the logarithm of the intensity ratio which we designate a projection  $P(x, y)$  is simply the line integral of attenuation coefficients along the path orthogonal to  $x, y$  if the  $\Delta z$ 's approach zero. Thus for each angle:

$$P_{\theta}(x, y) = \int_{\text{Detector}} \mu(x, y, z) dz \quad (4)$$

This summation is called a ray sum and the projection position is called a bin. The usefulness of this manipulation is that we can work with projections in a linear fashion to reconstruct the 3-dimensional distribution of linear attenuation coefficients and thus create an X-ray computed tomograph using linear operations discussed below. The denser the electrons, the more the X-ray beams are scattered or attenuated. In addition, the absorption of X rays due to the photoelectric effect has important effects. Indeed the reason there are small differences between gray from white matter in the brain in CT images is that the H, C, N, O, P content for gray and white matter differ slightly resulting in about 4% changes between gray and white matter.

The 3-D reconstruction in X-ray CT is the composite of a stack of 2-D transverse planes. A single 2-D transverse plane or tomograph consists of pixels whose values are proportional to the attenuation coefficients (i.e., scattering from electron density and photoelectric absorption). Individual solid state detectors are used for X-ray CT wherein the photon intensity is recorded as current generated in the detector for a given time interval. To perform the reconstruction, many one-dimensional projections are acquired from multiple angles. The idea is shown as the "1st generation CT" in Fig. 2. From these multiple projections the map or image of the anatomy in a particular slice is mathematically reconstructed using a computer, thus the word CAT for *computer assisted tomography*. The main method of reconstruction involves Fourier transformation of each projection, application of a ramp filter, inverse Fourier transformation, and back projection. This process is equivalent to back projection of each projection after each has been convolved with a kernel which is the Fourier transform of a ramp in spatial frequency space, thus the terminology "convolution method" (3). The contemporary trend for X-ray CT is to acquire finer resolution at high speed. These innovations have application to moving organs wherein data can be collected during 10 s to 16 s of breath holding.

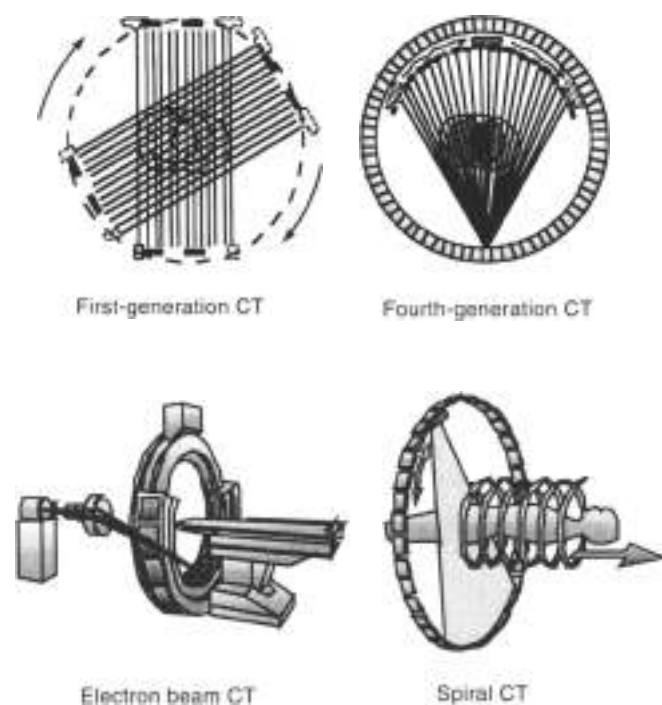
In spiral CT slip-ring technology provides continuous 360° rotation of the X-ray gantry for up to 40 s (4). Because the table continuously moves the patient through the gantry during scanning, the projection data are collected along a continuous spiral path instead of as parallel transaxial slices in



conventional X-ray CT. The simultaneous table movement and continuous gantry rotation (gantry rotation period is 1 s) without an interscan delay allow data to be collected very rapidly. When using intravenously administered iodinated contrast, images of the abdomen can be captured when the contrast is in the arterial or venous phase of the circulation. A typical high resolution (1-3 mm collimated) spiral examination of the abdominal aorta can be performed in 30 s covering 8 to 12 cm of volume. A conventional CT scan may take several minutes to cover the same volume and therefore cannot precisely image during the arterial contrast phase. This mode has brought X-ray CT applications to lung and abdominal scanning including data acquisition for virtual bronchoscopy and virtual colonoscopy.

#### Electron Beam Computed Tomography (EBCT)

Another mode of fast X-ray scanning with or without contrast injection employs an accelerated electron beam that strikes an anode target consisting of a large arc of tungsten encircling about 225 degrees transaxial to the patient (Fig. 2). The purpose is to avoid motion blurring from heart motion and to achieve an image sampling fidelity sufficient to evaluate the flow of contrast through the vascular and extracellular spaces mainly for heart, lung, aorta, kidney, and sometimes brain studies. EBCT applications are mainly for evaluation of coronary calcium and usually involve acquisition of 40 contiguous 3-mm thick sections extending from the right pulmonary artery to the heart apex. Each section is timed for acquisition during late diastole using ECG triggering and breath holding to minimize heart motion effects.



**Figure 2.** X-ray computed tomography has evolved to rapid scanning systems over the last 25 years. Electron beam techniques can scan a few transverse sections without distortion from heart motion, and spiral X-ray CT can cover the abdomen with 2 breath-holds at 1 s per each transverse section.

#### Digital X-Ray Systems and Digital Radiology

Though the present standard detector is X-ray film wherein the efficiency is enhanced by a phosphor which converts X-rays to light photons for subsequent film exposure, there is a major technological development to deploy flat-panel imaging systems using solid-state detectors for direct electronic readout to supplant the screen-film based systems. This is part of the conversion of diagnostic imaging centers to an all digital electronic-based imaging and information communication center. All of the radiological imaging methods with the exception of the most commonly used X-ray are digitally based. It is estimated that approximately 70,000 radiographic units in the United States will be retrofitted by digital detectors in the next few years. The present number of mammography units in the United States is approximately 11,000 and there is currently a major effort to create a digital system capable of providing a resolution at least as good as 50  $\mu\text{m}$  (10 line pairs per mm) which is one-half as good as the contemporary screen film analog mammography systems. The benefits anticipated for digital mammography systems are great if the systems can meet the design criteria including resolution, image acquisition time, and heat loading of the X-ray tube (5). The technologies now being evaluated range from miniature 200  $\mu\text{m}$  or less pixel-based, direct X-ray electronic detectors. A proposed mode uses complementary metal oxide semiconductor (CMOS) electronics coated with selenium for real-time readout of electrons generated by the X-rays striking the detector. Another mode uses cesium iodide scintillators coupled to silicon photodiodes or amorphous silicon solid state two-dimensional arrays to allow direct electronic readout into computer systems. These are developments which go beyond the charged coupled devices (CCDs) which have been limited to wafers (12-cm diameter) which are too small for general radiographic replacement without elaborate methods to expand the field of view using optical systems and multiple wafers.

A new concept in X-ray imaging embodied in the digital detector is known as reverse geometry imaging which matches a large X-ray source to a small solid state detector. The X-rays are created using an electron beam which scans a 2-dimensional anode through magnetic deflection. A cone of X-rays are focused onto a small solid state detector which records the attenuation through each patient position. The ideal detector for this application is cadmium zinc telluride (CZT) which is too expensive to make into panels for large field of view digital radiography. The concept of reverse geometry overcomes the size limitations and though designed for fluoroscopy, this innovation has other applications including breast mammography. Digital radiology systems facilitate image enhancement, calibration, distortion removal, and communications including storage, dissemination, consultation, and teleradiology.

#### ULTRASOUND

Ultrasound, as currently practiced in medicine, is a real-time tomographic imaging modality. Not only does it produce real-time tomograms of the position of reflecting surfaces (internal organs and structures), but it can be used to produce real-time images of tissue and blood motion due to the Doppler effect.

Ultrasound uses longitudinal compression waves generated by the voltage-induced oscillations of a piezoelectric crystal

which is typically a ceramic disk consisting of lead zirconium titanate (PZT). The oscillations are at frequencies of 0.5 MHz to 20 MHz. The ultrasound is a coherent pressure wave which is reflected from surfaces of varying acoustic impedances.  $Z$  ( $Z = \text{tissue density} \times \text{velocity of sound in tissue}$ ). Loss of ultrasound energy dissipated as heat is at about 2 dB per cm for 2 MHz. This dissipation is due to loss of coherence. The ultrasound scatters from surfaces and this scattering is proportional to frequency. The attenuation amounts to 1 dB per cm tissue depth for each MHz. Thus at 3 MHz the loss at 2 cm of tissue penetration is 6 dB or a halving of the signal strength. Thus for imaging the heart in children 5 MHz can be used but for imaging an adult abdomen frequencies are usually 1 MHz.

The major uses of ultrasound are in the examination of the carotid arteries, heart valve function, heart wall motion, gall bladder, and examination of the pregnant abdomen including the anatomy of the fetus.

The technique is to sonicate tissues with a transducer coupled by Vaseline to the skin surface over the area of study (e.g., carotid arteries, heart, gall bladder, uterus in pregnant woman). The location of surfaces within tissues is determined by measuring the time interval between the production of an ultrasonic pulse and the detection of its echo resulting from the pulse reflected from those surfaces. By measuring the time interval between the transmitted and detected pulse, we can calculate the distance between the transmitter and the object.

$$d = \frac{1}{2} t c \quad (5)$$

where  $c$  is the speed of sound in tissue, ca.  $1450 \text{ m s}^{-1}$ . The ultrasound pulses are both produced and detected by the piezo electric crystal or transducer. The reflected ultrasound imposes a distortion on the crystal, which in turn produces an oscillating voltage in the crystal. The same crystal is used for both transmission and reception. Types of ultrasound instruments are shown in Fig. 3. In addition to imaging (discussed under organs

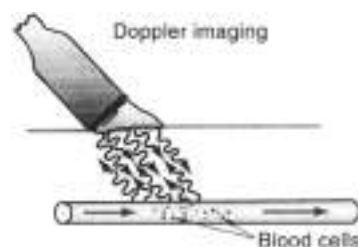


Figure 4. A major use of ultrasound is evaluation of blood flow in particular in the carotid arteries of the neck. The frequency changes associated with moving blood cells can give speed and direction of moving blood cells toward or away from the transducer.

below), ultrasound is being used for monitoring therapy methods such as hyperthermia, cryosurgery, drug injections, and as a guide during biopsies and catheter placements. Tissue acoustic impedance changes with temperature, the material being injected, and the metal or plastic surfaces of the catheters.

### Doppler Ultrasound

If a structure is stationary, the frequency of the reflected wave will be identical to that of the impinging wave. A moving structure will cause a backscattered signal to be frequency shifted higher or lower ( $\pm \Delta f$ ) depending on the structure's velocity toward or away from the sound generator or transducer. The Doppler equation is

$$\Delta f = \frac{-2f_0 \cos \theta v}{c + v} \approx -\frac{f_0 \cos \theta v}{c} \quad (6)$$

where  $f_0$  is the transmission frequency,  $v$  is the velocity of the moving cells,  $c$  is the velocity of sound, and  $\theta$  the angle between the transducer axis and the flow axis.

For example, when an impinging sound pulse passes through a blood vessel, scattering and reflection occurs from the moving red cells (Fig. 4). In this process, small amounts of sound energy are absorbed by each red cell, then reradiated in all directions. If the cell is moving with respect to the source, the backscattered energy returning to the source will be shifted in frequency, with the magnitude and direction proportional to the velocity of the respective blood cell. Thus, if we use ultrasound to image the cross-sectional area of the blood vessel, the volume of blood flow can be calculated from the area of the vessel and the average velocities of the blood cells. The frequency shift data are color coded (e.g., carotid arteries in red and veins in blue) to form 2-D images. Image positional data are provided by the range (distance) of the signal whose frequency shift gives the speed at that range. Obstructions to blood flow are readily evaluated by this method using hand-held scanning devices.

Power Doppler is a relatively new method of imaging the partial volume of moving blood. The sum of the square power of all frequency shifts gives information about the number of moving surfaces independent of direction. A local increase in vascularity as seen in breast cancer and lymph node metastases will give a relatively high power doppler signal.

### MAGNETIC RESONANCE IMAGING

Magnetic resonance imaging (MRI) has evolved into one of the most powerful noninvasive techniques in diagnostic imaging

Ultrasound

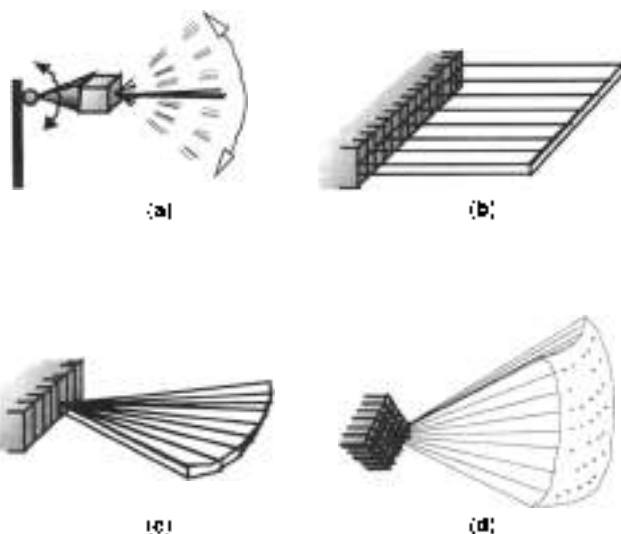
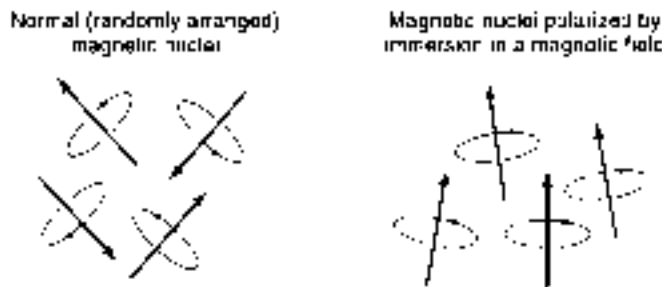


Figure 3. Modes of ultrasound imaging which allow direct 2-D evaluation by mechanical scanning (a) or electronic scanning at frequencies between 1 to 5 MHz (b-d). Catheter-tip transducers are also available for vascular work (c-d) using 10 to 20 MHz frequencies.



**Figure 5.** Magnetic resonance imaging is based on detection of the net magnetism of nuclei which have magnetic spin such as the hydrogen nuclei (protons) of tissue water. The detection is facilitated by RF pulses which stimulate signals from the oriented nuclei.

and biomedical research. MRI uses the principles of a well-known analytical method of chemistry, physics, and molecular structural biology. However, this basic method has been slightly modified by applying relatively small spatial magnetic field gradients of varying amplitudes and directions to achieve spatial information of the density and characteristics of nuclei with magnetic spins. MRI is primarily used as a technique for producing anatomical images, but as described below, MRI also gives information on the physical-chemical state of tissues, flow, diffusion, and motion information. Magnetic resonance spectroscopy (MRS) gives chemical composition information.

Most elements have at least one reasonably abundant isotope whose nucleus is magnetic. In biological materials, the magnetic nuclei of  $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{23}\text{Na}$ ,  $^{31}\text{P}$ , and  $^{39}\text{K}$  are all abundant. The hydrogen nucleus (a single proton) is abundant in the body due to the high water content of many tissues. When the body is immersed in a static magnetic field, slightly more protons become aligned with the magnetic field than against the static field (Fig. 5). At 1 T (10,000 gauss) and 25 °C the difference between these aligned populations of about one proton in a million produces a net magnetization. The net magnetization precesses around the static field at a frequency

$$\omega = 2\pi\nu = \gamma B \quad (7)$$

where  $\omega$  is frequency,  $\gamma$  is the gyromagnetic ratio particular to each nuclear species, and  $B$  is the field.

A rapidly alternating magnetic field at the resonant frequency  $\omega$ , applied by a coil near the subject or specimen in the static magnetic field, changes the orientation of the net magnetization relative to the direction of the static magnetic field.

These changes are accompanied by the absorption of energy (from the alternating magnetic field) by nuclei which undergo the transition between energy states. When the alternating field is turned off, the nuclei return to the equilibrium state, emitting energy at the same frequency as was previously absorbed. The nuclei of different elements, and even of different isotopes of the same element, have very different resonance frequencies. For a field of 1 T (10,000 gauss), the resonance frequency of protons is 42 MHz and that of phosphorus is 17 MHz. Thus, the magnetic nuclei in the body, when placed in a static magnetic field, can be thought of as tuned receivers and transmitters of RF energy. The governing relation is Eq. (7). Unlike its X-ray counterparts, MRI is not a transmission technique. Rather, similar to PET and SPECT, the material imaged is in itself the signal source (i.e., the polarized nuclear spins).

The motion of the macroscopic nuclear spin magnetization,  $\mathbf{M}$  is conveniently described in terms of the phenomenological Bloch equation.

$$\frac{d\mathbf{M}}{dt} = \gamma \mathbf{M} \times \mathbf{B} - \frac{M_x \hat{x} + M_y \hat{y}}{T_2} - \frac{M_z - M_0}{T_1} \hat{z} \quad (8)$$

where  $M_0$  is the equilibrium magnetization, and  $T_1$  and  $T_2$  are relaxation times.  $T_1$  is the characteristic relaxation time for longitudinal magnetization to align with the magnetic field; following a perturbation such as an RF pulse or a change in magnetic field, the longitudinal magnetization typically returns to its equilibrium value,  $M_0$ , with a time constant  $T_1$ .  $T_2$  is the characteristic time for decay of coherent magnetization in the transverse plane: the transverse magnetization decays exponentially with time constant  $T_2$  to its equilibrium value,  $M_{xy} = 0$ . For proton MRI, both relaxation times are determined by interaction of water with macromolecules in tissues. The dependence of image contrast on these parameters is appreciated by the equation components of the signal from a typical spin echo experiment:

$$S(x) = f(x)\rho(1 - \exp(-TR/T_1))\exp(-TE/T_2) \quad (9)$$

where  $f(x)$  is a flow function,  $\rho$  is the density of  $^1\text{H}$  (protons), TR the pulse sequence time interval, and TE the time to detect the echo from a 180° pulse applied TR/2 after the 90° RF pulse. Tissues have  $T_1$  values which vary from 600 ms to 3000 ms (cerebral spinal fluid) and  $T_2$  values of 20 ms to 40 ms, except tumors whose values can be greater than 60 ms. The time needed to acquire MR images was thought to be prohibitive for functional studies such as done with X-ray angiography; however, when it was discovered that flip angles less than 90° for the RF pulse could be used for MRI with acceptable signal to noise ratios and tissue contrast, a major horizon for contemporary studies was discovered (6). Now single slices can be obtained in less than 40 ms.

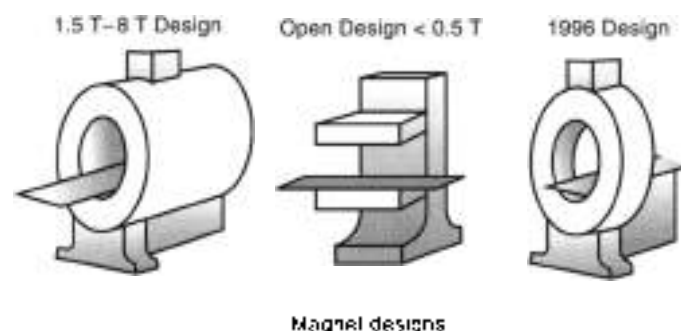
The principal components of the MRI machine are the magnet, radio frequency (RF) coils, and the gradient coils. The gradient coils are used to acquire spatial information. Note that if Eq. (7) is modified by superposing a spatial gradient  $\Delta B/\Delta x$  on the large static field, there will be a frequency shift in accord with

$$\omega + \Delta\omega(x) = \gamma \left( B + \frac{\Delta B}{\Delta x} x \right) \quad (10)$$

It can be seen that the frequency for any nuclei at position  $x$  will be

$$\omega + \Delta\omega(x) \quad (11)$$

The received signal for all the nuclei in the space being imaged will be the linear superposition of the respective frequencies with amplitudes and individual signal decay characteristics (i.e.,  $T_1$  and  $T_2$ ) determined by the local tissue biophysics at each spatial position. Three-dimensional information is gleaned by varying gradients in the  $x$ ,  $y$ ,  $z$  directions. The Fourier transform of this complex signal gives directly the image projections of intensities of nuclear magnetization. Reconstruction is implemented by back projection of these projections with appropriate filtering as is done in X-ray CT. An alternate reconstruction strategy involves use of phase alterations to encode



**Figure 6.** Magnets used for human imaging use superconducting wire in liquid helium systems or permanent magnet elements arranged around the subject for fields of about 0.5 T. Technological advances have allowed open designs for fields up to 1.5 T using superconductor wire.

information related to spatial positions of the nuclei. The RF coils used to excite the nuclei usually are quadrature coils which surround the head or body, but small (e.g., 6 to 10 cm) flat coils placed on the surface of the head or body are also used. Besides being the essential element for spatial encoding, the gradient-coil subsystem of the MRI scanner is responsible for the encoding of specialized contrast such as flow information, diffusion information, and modulation of magnetization for spatial tagging.

Magnet types in current use are of the superconducting, resistive, and permanent magnet designs ranging in strength from 0.05 to 4 T and have evolved from 2-m-long cylinders to more open designs for patient access and acceptability (Fig. 6). The majority of MR systems use superconducting magnets which provide fields of 1.5 T. Most currently produced magnets are based on niobium-titanium (NbTi) alloys, which are remarkably reliable, but require a liquid helium cryogenic system to keep the conductors at approximately 4.2 K (−268.8°C).

#### MRI Contrast Agents

MRI contrast agents are used to demonstrate perfusion and vasculature of organs. The usual contrast agents are water-soluble chelates of gadolinium. The chelates distribute in the vasculature and interstitial spaces much like the iodine contrast agents used for X-ray CT. Gadolinium causes an increase in the relaxation rates of the tissue water, thus, a decrease in  $T_1$  and  $T_2$ . Presence of the gadolinium results in a signal enhancement on the images produced by pulse sequences which are designed to give higher signal for tissues with short  $T_1$  (e.g., short TR in Eq. (9)). Manganese complexed to a chelate accumulates intracellularly in some tissues and also causes  $T_1$  shortening which leads to signal enhancement. Other types of contrast agents used by MRI are colloidal iron oxides (e.g., magnetite) coated with dextran which cause a MRI signal decrease in the region of their accumulation or a transient signal decrease as they perfuse through tissue due to the increase  $T_2$  relaxivity (short  $T_2$  times of Eq. (9)) of local tissue protons.

#### Magnetic Resonance Spectroscopy (MRS)

The nuclei constituting compounds in the human body have a resonance frequency governed by Eq. (7) relative to the imposed magnetic field. A nucleus of hydrogen will be in a slightly different magnetic field than the imposed magnetic field from the magnetic resonance magnet because adjacent spinning

nuclei generate a small local field that shifts the frequency a few parts per million. Thus the protons of lactate are shifted in frequency by a few hundred hertz from the resonance frequency of water protons of 64 MHz at 1.5 T. This shift is known as the chemical shift and the resulting magnetic resonance spectrum gives specific information about the concentration of compounds in tissue noninvasively. Thus by manipulating the MRI gradients, selected regions can be sampled to determine the concentrations of tissue constituents containing hydrogen-1, carbon-13, phosphorus-31, and other NMR nuclei. As the spectra change with disease states, MRS has an important role in diagnoses of diseases, particularly brain cancer, prostate cancer, epilepsy, and brain trauma.

The sensitivity of MRS is very low, thus only tissue constituents with concentrations in the millimolar range are evaluated in selected volumes usually greater than 0.2 ml (0.6 × 6 × 6 mm) for proton spectroscopy and greater than 1 ml for phosphorus spectroscopy. The sensitivity increases with field strength and MRS imaging, though feasible for low-resolution imaging of protons for maps of choline-containing compounds, N-acetylaspartate, lactate, and lipids.

#### Functional Magnetic Resonance Imaging (fMRI)

Though 25 years ago nuclear medicine methods demonstrated flow changes associated with mental activity in humans, it was not until the less invasive (i.e., no radioactivity) methods of MRI showed a magnetic resonance signal associated with stimulation of the brain that the medical scientists embraced fMRI as a method for mapping brain function. Methods that reflect brain activity currently use the fact that in response to a stimulus (e.g., light flash, sound, touch, physical movement), there is an increase in local blood volume and blood flow in order to supply that region of the brain with oxygen and nutrients. The local blood flow in areas of 5 mm or more around the part of the brain being stimulated is in response to an increase in metabolism and electrical activity. Radioactive tracers have shown this functional response and MRI measurements can also show flow changes using injected contrast material such as Gd-EDTA or Gd-albumin or other methods of detecting flow (see above). However, the principal method now employed by fMRI relies on an intrinsic change in the local magnetic field associated with a change in the partial volume of oxygenated hemoglobin (diamagnetic) and deoxygenated hemoglobin (paramagnetic). The process is known as the BOLD (blood oxygen level dependent) effect.

The presence of paramagnetism from deoxyhemoglobin in capillaries, arterioles, and venules causes a magnetic field gradient across the imaging pixel sufficient to lower the MR signal. The positive BOLD signal is due to the effective decrease in paramagnetism and associated magnetic gradients over ranges less than 1 mm due to an increase in the partial volume of diamagnetism associated with more oxyhemoglobin in the region of activation. The fMRI is a map of the result of subtraction of the MRI before the activation from the MRI obtained by fast imaging methods performed during the activation. The more paramagnetism or deoxyhemoglobin the lower the signal and when one subtracts the MRI image taken before the activation from that taken after the activation the region of activation will appear relatively darker (a negative BOLD signal).

The almost universal findings from studies done at 1.5 T show a positive BOLD signal, (e.g., a 2% change), and until

recently, this was thought to represent an in-flow of diamagnetic oxyhemoglobin because the brain was not extracting metabolic oxygen, however, using fields of 4 T it has been shown that the positive BOLD signal is preceded by a negative BOLD signal which corresponds to an initial decrease in blood oxygen (more paramagnetism) within 0.5 s of the activation followed by an over-compensation of blood rich in oxygen to the activated region of the brain. Thus there is initially an increase in oxygen extraction, local blood volume possibly first of the capillary bed and then a compensatory increase in the regional flow of oxygenated blood leading to less local magnetic field distortion and thus a positive signal when the before activation image is subtracted from the activation image. The positive BOLD signal, though not precisely positioned at the area of activation due to the overflow phenomenon, is used in clinical medicine to aid in epilepsy and tumor surgery by showing the neurosurgeon which vital areas of the brain to avoid in the surgical procedure. Magnetoencephalography is also used for this purpose and the combination of fMRI and MEG can give improved functional brain activity maps.

### RADIONUCLIDE EMISSION IMAGING (NUCLEAR MEDICINE IMAGING)

Nuclear medicine imaging, also known by the terms emission imaging, radio-isotope imaging, scintigraphy (i.e., making pictures of the scintillations from crystals), is the technique of making images of the distribution of radioactive compounds. The two major elements of this method are the type of radiopharmaceutical which is injected into the patient and the type of scanner or camera (e.g., Anger camera, PET, SPECT) used for detection of gamma radiations from the radioactive isotope in the injected compound (Fig. 7). The particular diagnostic application relies on the fact that radiopharmaceuticals go to different organs in varying amounts depending on the disease state and the type of radiopharmaceutical. An example of this technique is the use of radioactive iodine to map the metabolic activity of the human thyroid gland in the neck. As about 25% of the ingested or injected iodine accumulates in the thyroid

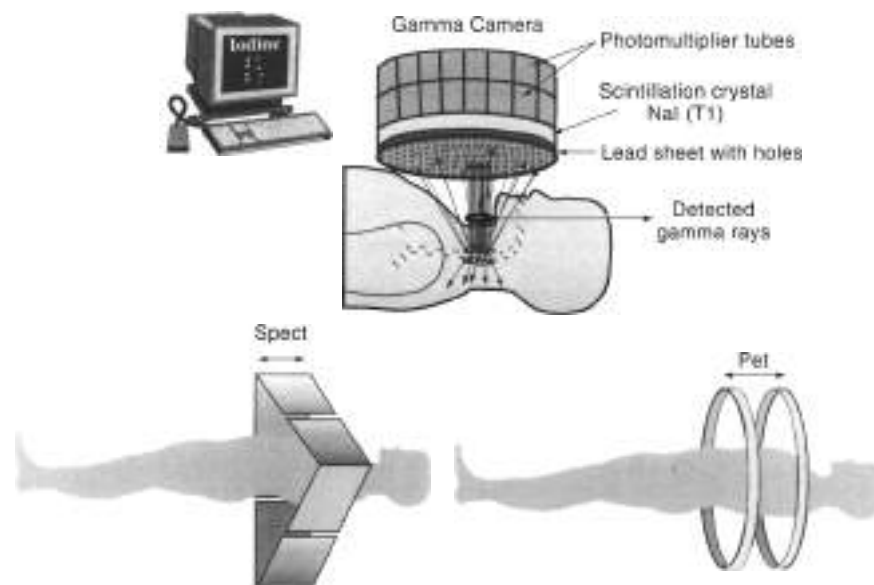
**Table 2. Selected Radionuclides and Associated Imaging Targets**

| Radioactive Compound   | Organ                           |
|--|---------------------------------|
| $^{131}\text{I}$ Iodine  | Thyroid                         |
| Iodine compounds   | Tumors                          |
|  | Brain neuroreceptors            |
|  | Heart muscle                    |
| $^{99\text{m}}\text{Tc}$ $^{99\text{m}}\text{Tc}$ $^{99\text{m}}\text{Tc}$ |                                 |
| Pyrophosphate  | Bone                            |
| DPHA   | Kidney                          |
| HMPAO  | Brain                           |
| Peptides   | White blood cells and platelets |
| Sestamibi  | Heart                           |
| Sulfur colloid   | Liver                           |
| $\text{Tl}^+$ compounds  | Tumors                          |
| $^{18}\text{F}$  | Heart                           |
| $^{15}\text{O}$  | Bone                            |
| $^{18}\text{F}$ - deoxyglucose   | Brain, heart, tumor             |
| - neuroreceptor ligands  | Brain neuroreceptors            |
| $^{15}\text{O}$ - $\text{H}_2\text{O}$                                     | Brain blood flow                |
| $\text{O}_2$   | Brain, heart, tumor             |

(Reprinted with permission)

within 24 h as part of the process of making thyroid hormone, it is possible to detect trace amounts of radioactive iodine by use of nuclear detectors, and to make images of the thyroid. These are not anatomic images but are images reflecting the spatial distribution of metabolic activity.

The majority of contemporary nuclear medicine studies are for brain blood flow, heart flow with and without induced stress by exercise or a pharmacologic agent, bone tumor scanning, thyroid scanning, and cancer detection. Specialized studies such as the detection of internal bleeding by injection of the patient's red blood cells after labeling with  $\text{Tc-99m}$  and a host of other applications are enabled by use of specific tracers. The radiation dose to the patient is about equivalent to that from background radiation of one year (i.e., 0.1 mSv) with a maximum to any organ equivalent to a back X-ray examination (i.e., 1 mSv). Table 2 lists a few of the commonly used agents and their applications.



**Figure 7.** The three commonly used instruments for imaging radionuclides (nuclear medicine imaging). Relative sensitivity is determined by the area of detector material to which a source is exposed. For conventional gamma camera imaging the sensitivity is related to the solid angle ( $\text{sr}^2$ ,  $4\pi\text{sr}^2$ ) provided by the detector or proportional the resolution divided by the distance squared between the radiation and the detector. For PET the sensitivity is related to area (not resolution) of [detector] material divided by  $4\pi\text{r}^2$ .

### Positron Emission Tomography (PET)

Whereas most radioactive isotopes decay by release of a gamma ray and electrons, some decay by the release of a positron. A positron is similar to an electron but has a positive charge. PET imaging begins with the injection of a biological molecule that carries with it a positron-emitting isotope (for example,  $^{14}\text{C}$ ,  $^{15}\text{N}$ ,  $^{13}\text{C}$ , or  $^{18}\text{F}$ ). Within minutes, the isotope accumulates in an area of the body for which the molecule has an affinity. As an example, glucose labeled with  $^{13}\text{C}$  (half-life, 20 min), or a glucose analog labeled with  $^{18}\text{F}$  (half-life, 1.8 hr), accumulates in the brain, where glucose is used as the primary source of energy. The radioactive nuclei then decay by positron emission. The emitted positron collides with a free electron usually within 1 mm, from the point of emission. The interaction of the two subatomic particles results in a conversion of matter to energy in the form of two gamma rays, each with an energy of 511 keV (note:  $E = mc^2 = 511 \text{ keV}$  for the mass of each particle). These high-energy gamma rays emerge from the collision point in opposite directions ( $180^\circ \pm 0.25^\circ$ ), and are detected by an array of detectors which surround the patient (Fig. 8).

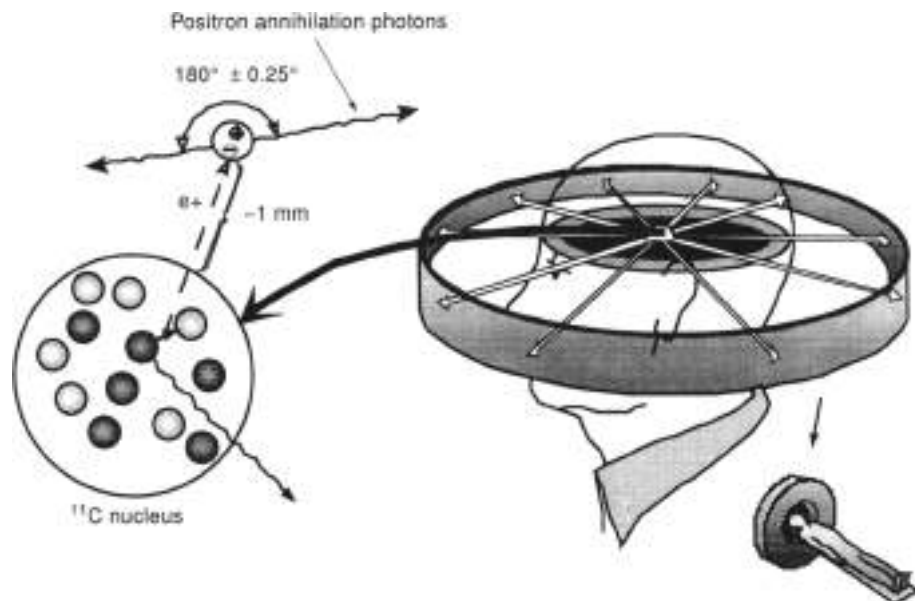
When the two photons are recorded simultaneously by a pair of detectors, the nuclear decay that gave rise to them must have occurred somewhere along the line connecting the detectors. If one of the photons is scattered then the line of coincidence will be incorrect. After 500,000 or more annihilation events are detected, the distribution of the positron emitting tracer is calculated by tomographic reconstruction procedures. PET usually reconstructs a two-dimensional image. Three-dimensional reconstructions can also be done using 2-D projections from multiple angles (7).

The sensitivity of PET to detect nanomolar concentrations of tracers distinguishes this modality and its applications from those of MRI and MRS. MRI has exquisite resolution for anatomic studies and for flow or angiographic studies. In addition, magnetic resonance spectroscopy (MRS) has the unique attribute of evaluating chemical composition of tissue but in the millimolar range rather than the nanomolar range. Since the nanomolar range is the concentration range of most receptor proteins in the body, positron emission tomography is ideal for this type of imaging.

### Single-Photon Computed Tomography (SPECT)

SPECT, like PET, acquires information on the spatial distribution of radionuclides injected into the patient. As in X-ray CT, SPECT imaging involves the rotation of a photon detector array around the body to acquire data from multiple angles. Because the emission sources (injected radionuclides) are inside the body cavity, the reconstruction tomography task is far more difficult than for X-ray CT, where the source position and strength (outside the body) are known at all times. In X-ray CT the attenuation is measured, not the source strength. In PET and SPECT the reconstruction problem requires determination of the source position and strength (concentration) which is a much more difficult problem than simply determining the attenuation coefficient distribution. To compensate for the attenuation experienced by emitted photons from injected tracers in the body, contemporary SPECT machines use mathematical reconstruction algorithms more complex than those needed for X-ray CT (5).

SPECT sensitivity is inferior to PET thus the attainable resolution of SPECT is limited due to a statistical limitation associated with acceptable amounts of injected radionuclides. Radionuclides used for SPECT imaging emit a single photon (e.g., 140 keV for  $^{99\text{m}}\text{Tc}$ , 160 keV for  $^{131}\text{I}$ ) but the positron emitters (e.g.,  $^{14}\text{C}$ ,  $^{15}\text{N}$ ,  $^{13}\text{C}$ ,  $^{18}\text{F}$ ) result in two 511 keV photons for PET detection. Because only a single photon is emitted from the radionuclides used for SPECT, a special lens known as a collimator is used to acquire the image data from multiple views around the body. The use of a collimator results in a significant decrease in detection efficiency as compared with PET. In PET, collimation is achieved naturally by the fact that a pair of detected photons (gamma rays) can be traced to a line after being produced. In PET, there might be as many as 500 detectors that could "see" a PET isotope at any one time where as in SPECT, there may be only 1, 2, or 3 collimators; thus, only 1, 2, or 3 detectors are available to each resolution volume in the subject. The statistics needed for reliable images and practical radioactive dose limitations limit the useful resolution of SPECT. The resulting useful resolution (about 7 mm) for SPECT is inferior to PET resolutions of about 4 mm for contemporary commercial systems and 2 mm for current research instruments.



**Figure 8.** The PET isotopes usually have more protons than neutrons and when a proton becomes a neutron the positron (positive electron) is released which encounters an electron. Positron and electron pair masses annihilate creating two photons that disperse at  $180^\circ$  to be detected by PET sensors.

Although SPECT imaging resolution is not that of PET, the ready availability of SPECT radiopharmaceuticals, particularly for the brain and head, and the practical and economic aspects of SPECT instrumentation make this mode of emission tomography attractive for clinical studies of the brain, heart, bone, and tumors (i.e., tumor metastases).

## ENDOSCOPY

Endoscopy (i.e., visualization of the lumen of human conduits such as blood vessels and intestines) ranks among the most important advances in imaging over the last few decades. Unlike ultrasound, X-ray CT, and MRI, however, endoscopy has not in general been embraced by radiologists, being largely performed by other specialists and often viewed as competitive with well-established radiologic techniques.

The modern endoscope is a complex instrument requiring considerable precision in its construction. The shafts are circular in cross section with diameters dependent on the function of the instrument. Upper gastrointestinal instruments vary in diameter from 6 to 13 mm, side-viewing duodenoscopes are 10 to 13 mm, and colonoscopes can be 15 mm in diameter.

There are two principal forms of the endoscope: the fiberoptic endoscope and the video endoscope. In the former the shaft carries a fiberoptic-optic bundle to transmit light, and a second fiberoptic-optic bundle (the viewing bundle) transmits an image to a lens system in the eyepiece of the endoscope. The fibers in this bundle retain their spatial relationship to each other throughout their length in order to transmit the image. Video endoscopes are now becoming the preferred type. The viewing bundle is replaced by a high-resolution video "chip" at the tip of the endoscope. The image is displayed on a color monitor which is more comfortable for the endoscopist, facilitates image storage and manipulation, and aids training by dynamic image access by multiple viewers. The shaft of both types of endoscopes accommodates a "working channel" (up to 4.2 mm in the largest side-viewing duodenoscopes) for the passage of accessories such as biopsy forceps, baskets, balloons, snares, and stents.

## INTERVENTIONAL RADIOLOGY

For many years a wide range of diagnostic and therapeutic procedures have been performed by fluoroscopic-aided catheter or device insertion. Currently, ultrasound, X-ray CT, and MRI are used in addition to fluoroscopy to guide the placement of probes with minimal disruption of tissue. The procedures include drainage of fluid from the pericardium, lungs, and abdomen; minimally invasive neurosurgical treatment of arteriovenous malformation of cerebral vessels, treatment of vertebral disk pathology; guided placement of radium sources; and image controlled freezing or hyperthermia treatment of cancer. The percutaneous approach to surgical treatment depends on contemporary advances in 3-D imaging methodologies and new methods of visualization including virtual reality. Three technological advances that have enabled interventional radiology to replace many surgical methods include: the 3-D capabilities of MRI, microfabrication methods for developing miniature surgical tools, and image manipulation and visualization methods.

## DIAGNOSTIC RADIOLOGICAL IMAGING

Below the applications of the main methods of diagnostic imaging in human health care are presented under the categories of the body most commonly studied: brain, chest and lung, heart, abdomen, liver, kidneys, female reproductive organs, breast, prostate, and the skeletal system. Some illustrative images are presented as catalogued by Fig. 1. Medical background material and patient-based radiological procedures are found in (9).

### BRAIN

Two of the most widely used imaging modalities in the study of the brain, cranial nerves, and spine are X-ray CT and MRI. As X-ray CT is well suited to imaging of bone, calcifications, and hemorrhage, it is still a mainstay of imaging in the emergency room particularly in the evaluation of head and face trauma and in suspected subarachnoid hemorrhage of the brain. With the advent of X-ray CT and MRI, the uses of plain film X-rays and injection of air for demarcating the ventricles (i.e., pneumocephalography) have disappeared. Though ultrasound study of the infant brain is possible because of less attenuation from the infant's underdeveloped skull bone, even applications of ultrasound have been supplanted by X-ray CT and MRI in infants.

#### Cerebrovascular Diseases (Hemorrhage and Stroke)

A major application of X-ray CT in the emergency room is the evaluation of subarachnoid hemorrhage which is associated with trauma to the head but also occurs spontaneously. Of the spontaneous types 75% occur from the rupture of an arterial aneurysm and 25% occur due to leakage from an arteriovenous malformation. X-ray CT is the diagnostic mode which gives an increased signal from blood in the cerebral spinal fluid spaces. Intracerebral hemorrhage which can occur in hypertension is also best evaluated by X-ray CT which has well known temporal changes in signal intensity with time after the episode.

Stroke is the result of a disruption of the nutrient blood flow to part of the brain and is a major cause of brain malfunction, particularly in the elderly. A transient stroke known as the transient ischemic attack lasts for 24 h or less, the stroke in evolution causes a progressive neurological defect, and the completed stroke is one in which the neurological defect appears fixed with no or very slow return to some function. MRI if available is the procedure of choice at the onset of symptoms using diffusion-weighted imaging methods. Later due to the cytotoxic and then extracellular edema MRI and even X-ray CT can pinpoint the tissues involved. Frequently, the usual MRI study will not reveal the stroke territory until 8 h after the event but the pathology can be detected by diffusion weighted imaging within 15 min of the event. Use of fast imaging methods such as echo-planar MRI will, with a MRI contrast agent, show decrease in regional blood flow.

PET and SPECT techniques can in principle detect a stoppage in flow to regions of the brain at the time of the event because these techniques measure tissue blood flow directly by following the distribution of a flow tracer. The lack of day and night availability of a flow tracer for PET has limited its application to stroke. As PET and SPECT have a major importance in determination of vascular reactivity of compromised regions of the brain they are useful in the staging of patient treatment after the acute phase.

### Brain Angiography

The use of contrast X-ray studies in evaluation of the brain vascular system (i.e., cerebral angiography) has to some extent been replaced by magnetic resonance angiography (MRA) though digital subtraction angiography still provides higher resolution images of the vascular tree when very high resolution is diagnostically important. MRA need not require any contrast medium injection as the method uses intrinsic properties of magnetic resonance of moving protons to attain contrast between flowing blood and stationary protons. The phase-sensitive technique uses a bipolar flow encoding gradient to cause a phase shift for a moving proton but a zero phase shift or canceled phase shift for a stationary spins. By reversing the polarity of the flow encoding gradient on alternate acquisitions and subtracting the data from these acquisitions the movement of the protons provides the signal. The time-of-flight method relies on magnetization of spins flowing into an excited volume wherein the stationary spins have been saturated; thus the unsaturated flowing blood will have a detectably higher signal compared with the surrounding previously saturated parenchyma. In addition, recently developed intravascular contrast agents such as a gadolinium-albumin complex have been used with MRI to improve definition of the cerebral vascular system.

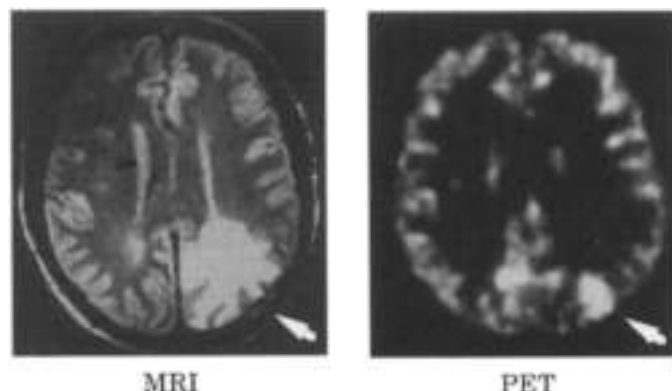
### Brain Tumors

The prevalence of primary brain tumors in the population is lower than most tumors and other diseases (about 2% of autopsies) yet the importance in diagnosis is vital to health care decisions particularly for metastatic tumors, which constitute 20% of brain tumor diagnoses. The majority of primary tumors are glioblastoma multiforme (25%) and the prognosis for these tumors is poor; however, therapy guided by presurgical and postsurgical MRI does aid in prolonging life with good quality for a few years. The prognosis for meningioma, pituitary adenoma, and acoustic neuroma, which together constitute 32% of brain tumors is very good if detected early, and both MRI and CT have played a major role in the presurgical planning for treatment. Tumors are usually hyperintense on  $T_2$  weighted MRI with a large region of vasogenic edema surrounding the main tumor site (Fig. 9). Though MRI studies provide a wealth of data in brain tumor detection, X-ray CT can in some cases (e.g., meningioma) be as valuable. MRI sequences designed to show blood volume have been useful in defining tumor locations and response to therapy.

PET and SPECT have played a major role in the postsurgical followup of treated tumors. PET imaging following glioblastoma surgery for example can distinguish recurrent tumor from radiation necrosis (Fig. 9); however, recently use of magnetic resonance spectroscopy (MRS) has shown that in areas of recurrent tumor there is an elevated choline signal representative of an increase in one or more of the choline-containing compounds in a tumor.

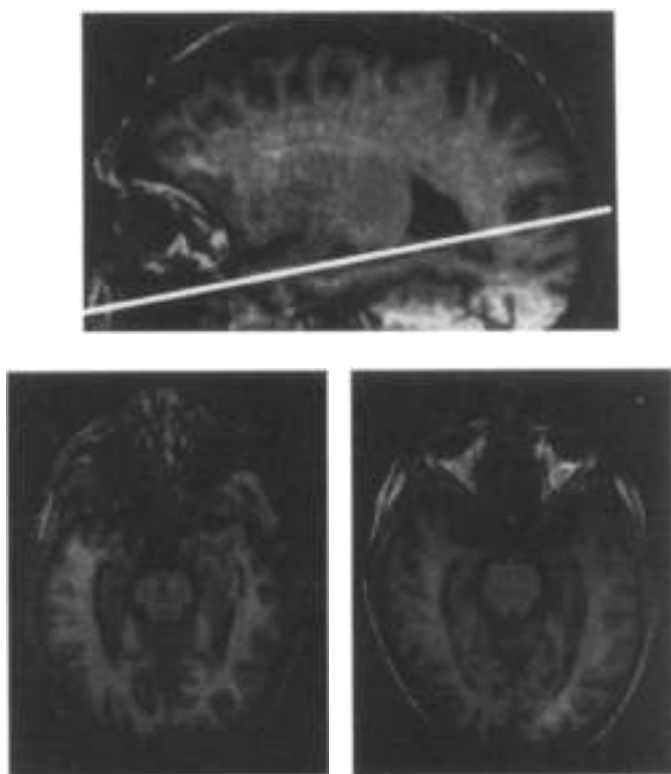
### Brain Neurodegeneration (Alzheimer's Disease and Multiple Sclerosis)

The most common form of neurodegeneration is Alzheimer's disease. As this disease is associated with the atrophy of the hippocampus and the parietal lobe during the early course of the disease, X-ray CT and now more recently MRI have been



**Figure 9.** Comparison between MRI spin echo and PET glucose metabolism of a brain tumor. MRI shows increased water and longer  $T_2$  relaxation in and around (edema) the tumor. PET shows increase in glucose metabolism typical of brain tumors. Detection of probable islands of tumors outside the bright spot (arrow) depends on the resolution of the PET scanner.

used to demonstrate the extent of loss of brain mass. Unfortunately, the atrophy of Alzheimer's disease is not dissimilar to that found in normal aging as shown in Fig. 10. PET patterns of a decrease in the parietal lobe glucose metabolism are very characteristic of Alzheimer's disease as is a decrease in blood flow as revealed by either PET or SPECT. The quantification of



**Figure 10.** MRI using the 3-D method of gradient recall shows brain atrophy during normal aging of an 86-year old athlete (lower left) compared with that of a 77-year-old patient with Alzheimer's disease (lower right). The atrophied areas look like two seahorses facing each other on either side of the central brain stem. These seahorse-like structures are the hippocampi and are responsible for some memory functions.



changes in glucose metabolism do correlate with the progression of the disease and can be used to monitor any efficacy of proposed therapies.

Huntington's and Parkinson's diseases are generally detected by neurological examination but in these two diseases of neurodegeneration of parts of the brain, the X-ray (7) and MRI patterns usually appear normal. However, the metabolic data of PET give specific and diagnostic patterns of decreases of glucose metabolism of the caudate nucleus in Huntington's patients and decrease in the dopamine neurochemical system in the central gray matter of Parkinson's patients.

#### Demyelinating Disorders (Multiple Sclerosis)

MRI is the method of choice in the diagnosis of multiple sclerosis which appears as diffuse lesions of the white matter in the brain of the middle-aged adult. The areas of increased water content are associated with demyelination in regions usually surrounding the ventricles but lesions are found also in the spinal cord and optic nerve. The other demyelinating diseases such as the congenital leukodystrophies have the common characteristic of increased water content of regions of the cerebral white matter.

#### CHEST (LUNG) RADIOLOGY

The major uses of radiological imaging methods has been in the diagnosis of diseases of the chest or thorax. The major conditions for which radiologic imaging is essential is detection and evaluation of pneumonia (bacterial, fungal or viral), cancer, pulmonary embolism, and chest trauma.

Pneumonia (infective consolidation of the lung) is detected as increased density by the standard X-ray radiograph but the specific diagnosis is not usually made (e.g., 30% to 50% of pneumonias have an causative diagnosis). The most commonly identified organism is *Strep. pneumoniae* (30%) followed by *Haemophilus influenzae*, *Legionella*, *Chlamydia*, and *Mycoplasma pneumoniae*. Viral pneumonias are uncommon in adults but do predispose to bacterial pneumonias. The standard radiograph usually shows blotchy densities in localized areas sometimes involving both lungs.

Tuberculosis is a bacterial infection with an insignificant incidence until the early 1990s when an increase in incidence occurred related to HIV infections. Densities around the central thorax (i.e., paratracheal) and tracheobronchial adenopathy, effusions and patchy and multifocal opacities in the apicoposterior segment of an upper lobe or the apical segment of a lower lobe are common X-ray patterns in tuberculosis.

#### Lung Cancer

Lung cancer is usually classified into squamous cell carcinoma, adenocarcinoma, and undifferentiated carcinomas which are further divided into small (not-cell) and large-cell types. Squamous cell carcinoma and adenocarcinoma are found most often. The peak incidence age group is 50 to 60 years old. Once a tumor is suspected, three modalities of imaging are brought to bear: X-ray CT, MRI, and PET (or SPECT). The plain X-ray does show patterns which help distinguish bronchioloalveolar carcinoma from the other cell types, and the pattern of size, localization, and numbers of opacities can dictate treatment. Most patients still require an invasive procedure to distinguish a benign from

a malignant lesion. Bronchoscopy, percutaneous needle biopsy, thoroscopy, and open lung biopsy have associated risks and do not always provide a definitive answer.

A solitary pulmonary mass or nodule is the single most common presenting sign of bronchial cancer in 40% to 60% of lung cancers. But the majority of diagnosed solitary nodules are not malignant and even the malignant solitary nodules less than 20 mm in diameter can be successfully removed with a 5-year survival of 50%. X-ray CT, MRI, and PET can assist in a major way in staging these patients as these modalities allow evaluation of whether there is disease in the mediastinum in which case surgical resection would not be done as there is extended disease.

Unfortunately most lesions remain undiagnosed and patients usually proceed to biopsy or surgery, which frequently reveals benign processes. In a recent study of radiographically indeterminate thoroscopically resected solitary pulmonary nodules, 62% were benign (10). Thus, indeterminate lesions remain a dilemma with significant cost and morbidity. FDG PET imaging of thoracic neoplasms has been successful in distinguishing benign from malignant focal pulmonary abnormalities (11), in staging treatment for known malignancies, and in following patients after treatment for neoplasm (Fig. 11).

Once the diagnosis of malignancy has been established, radiographic staging is usually performed using various modalities such as radionuclide bone scanning thoracic X-ray CT and, if available,  $^{18}\text{F}$ FDG PET. The overall sensitivity and specificity of X-ray CT in detecting intrathoracic lymph node metastases in most series is only about 60%. FDG PET appears to be more sensitive although less specific for malignancy. Surgically proven metastases to hilar and mediastinal lymph nodes as small as 5 mm in diameter have demonstrated increased FDG uptake. After radiotherapy, residual abnormalities noted on chest radiographs can be differentiated as recurrent neoplasm or fibrosis using PET scanning (12).

#### Pulmonary Embolism

Blood clots in the lung circulation represent one of the most serious threats to life in both seriously ill patients, healthy patients in the postsurgical period, and in healthy patients (e.g., travelers in sedentary positions, leg fracture and sprained ankle cases in adults and youth). Postmortem studies have shown that up to 65% of hospitalized patients have emboli lodged in their pulmonary arteries (13) though only 1% have the clinical diagnosis. No single diagnostic test can be regarded as completely reliable in confirming or excluding the



**Figure 11.** Metabolic response of metastatic prostate cancer to Suramin therapy measured by  $^{18}\text{F}$ -fluorodeoxyglucose ( $^{18}\text{F}$ FDG) positron emission tomography (PET). Note also the normal uptake of  $^{18}\text{F}$ FDG in the left heart muscle whose brightness depends on the nutritional status of the patient (courtesy of G. Hoh, UCLA).

diagnosis of pulmonary embolism. In 90% of the cases blood clots arise from thrombosis of the deep veins in the legs. Blood stagnation is the major predisposing cause with only 15% or less of the pulmonary embolism patients having some blood clotting disorder.

Radiological imaging techniques are used in to detect deep vein thrombosis and also to detect the existence of pulmonary embolism by direct radiographic examination of the lungs. Actual detection efforts commence when a patient has some signs of leg thrombosis (e.g., leg pain, swelling). It is important to verify if deep vein thrombosis exists as half of these patients may have had silent pulmonary embolism and a diagnosis leading to therapy (i.e., anticoagulation) is imperative. A second situation is the patient who presents symptoms of pulmonary embolism (e.g., coughing up blood). The deep veins of the legs and possibly pelvis are involved and thrombosis can be detected by injection of X-ray contrast material or a radioactive tracer into the peripheral veins (e.g., foot) and subsequent X-ray or radionuclide imaging can detect blockage. But these techniques are imperfect as there are problems in venous access, incomplete filling, and discomfort to the patient. Indeed the patient can have thrombosis induced by the procedure. Of all the methods explored, magnetic resonance angiography seems now to have the potential for detection of deep vein thrombosis (14). Spiral X-ray CT has shown a 90% sensitivity and similar specificity for detection of pulmonary emboli relative to the gold standard of pulmonary angiography (15). More clots are detected with spiral X-ray CT than with pulmonary angiography. Though now considered to be a potential solution to the accurate, noninvasive diagnosis of pulmonary embolism, spiral X-ray CT shows great variability in detecting subsegmental emboli which have a frequency of 69%. This limitation is shared with the gold standard, however.

## BREAST CANCER

The incidence of breast cancer has been rising at 3% per year over the last 10 years to a prevalence level of about 0.1% of women in the United States. Since the mid-1960s when X-ray mammography became a readily available screening procedure, 44% of adult American women have had at least one procedure. The mortality rate of breast tumors remains high, however, at a rate of 25 per 100,000 per year. There are approximately 11,000 dedicated mammography units in the United States. The consensus is that better methods of verification of suspicious lesions detected on X-ray mammography are needed (only 1 in 4 are biopsy positive). In addition, a method is needed to evaluate (screen) patients with radiographically dense breasts. Even the most advanced methods of digital image processing of high resolution mammograms fail to give specific diagnosis in a large fraction of cases (Fig. 12).

Early breast cancers are often identified by planar X-ray mammography. The cost of this examination is low, the spatial resolution high ( $\sim 100 \mu\text{m}$ ). X-ray mammography is an effective screening technique for detection of cancerous growth in early stages. However, the complexity and heterogeneity of tissue within normal breasts makes unambiguous identification of malignant tumors difficult, as benign tumors and other breast structures can be similar in appearance to cancerous lesions. As a result, a large fraction (50% to 75%) of the suspicious structures identified in mammograms are noncancerous



**Figure 12.** High resolution digitalized mammogram which shows microcalcification which could represent breast tumor but confirmation requires a biopsy (courtesy of D. Liu, University of Pittsburgh)

(16) and, thus, further diagnosis is necessary before determining that the patient should be treated for cancer.

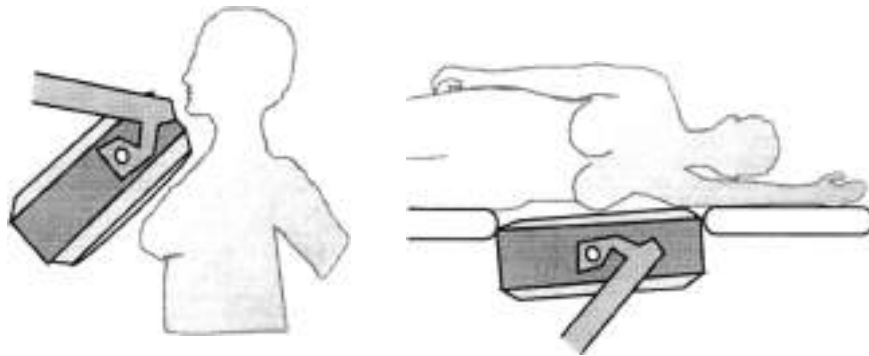
This further diagnosis is often obtained by biopsy. The biopsy can determine whether the structure is cancerous, and if so, can also determine the type of cancer and so suggest a course of treatment. The biopsy is time consuming for the physician, often uncomfortable for the patient, can increase the patient's radiation exposure and, unless done by a cytopathologist at the bedside, takes several days before the results of a full assay for tumor type are available.

Should the tumor prove to be cancerous, it is desirable to know before performing surgery whether the cancer has spread to the axillary nodes, as this affects the treatment selection and prognosis: 60% to 90% of patients without nodal involvement have no further breast cancers detected in their lifetimes (16). Nodal involvement is difficult to determine before surgery from an X-ray CT or MRI scan. Nodes can be enlarged for a variety of benign reasons and small tumors difficult to observe by non-temporary radiologic imaging methods.

Contrast enhanced MRI has demonstrated a high sensitivity to detection of small ( $< 1 \text{ cm}$ ) lesions in the breast; however, the specificity is less than 40%. The specificity improves slightly when dynamic techniques are used to measure the rate of contrast agent uptake during the initial minutes after contrast injection (17), but this specificity increase comes at the expense of spatial resolution or full imaging coverage of the breast. The axillary nodes are poorly evaluated with MRI mainly because of the inability to separate involved and uninvolved nodes based on contrast enhancement, as contrast enhancement occurs for both normal and pathological nodes.

Doppler-ultrasound using color to encode velocity or blood volume (power) is another method of evaluating X-ray mammographically detected suspicious lesions before biopsy (18).

Conventional gamma camera breast imaging



**Figure 18.** A major problem in breast imaging with conventional gamma cameras is to get close to the probable cancer site. Compact solid state imaging systems currently being designed for the breast can overcome this problem.

This method is noninvasive and inexpensive, but it requires a skilled operator and is not envisioned as a screening method. Doppler-ultrasound imaging has been successful in showing which of the suspicious lesions are cancerous and has value in determining lymph node involvement (19) in preliminary detection studies, particularly in those with contrast agents (20).

Positron emission tomography (PET) and the tracer  $^{18}\text{F}$ -fluorodeoxyglucose ( $^{18}\text{F}$ FDG) can provide excellent sensitivity for malignant breast tumors and axillary node involvement. Because of the expense and limited availability of PET, single-photon methods (SPECT, Gamma camera imaging, scintimammography) can play a major role in the differential diagnosis of suspicious X-ray mammography lesions. The specificity of the radiopharmaceutical  $^{99\text{m}}\text{Tc}$ -sestamibi for axillary node involvement with breast cancer ranges from 42% to 88%. This specificity, though better than X-ray mammography, could be increased by an improved detector system. It is believed that the 8% to 20% of tumors that are not revealed by  $^{99\text{m}}\text{Tc}$ -sestamibi and those below 1.5 cm in diameter could be detected if the limitations of contemporary gamma cameras are overcome.

Contemporary scintimammography and SPECT studies use large field of view gamma cameras which use a scintillator block coupled to a bulky array of photomultiplier tubes. By nature of their large size, these instruments are inadequate in most clinical situations because imaging of small organs such as the breast is usually awkward due to the fact that close access to the breast and axilla is prevented by the camera housing imaging system (Fig. 13). Solid state cameras have the potential to overcome these limitations.

## HEART

There are multiple methods used for examination of the human heart: ultrasound, X-ray coronary angiography, magnetic resonance imaging (MRI), magnetic resonance angiography (MRA), and nuclear medicine imaging methods (PET and SPECT). Ultrasound is used in the cardiologist's office for evaluation of valvular function using the M-mode method wherein motion of the valves is recorded as reflected signal distance versus the temporal sequence of the cardiac cycles. Using the 2-D imaging mode, the surface of much of the heart muscle can be imaged on-line, thus abnormalities in the synchronous motion of the contracting and expanding heart can be visualized by the cardiologist and changes of the heart muscle thickening during

infusion of mild pharmacological stress can give evidence of the health of the heart. Ultrasound contrast agents using microbubbles have been introduced recently to aid in imaging the heart chambers and flow dynamics.

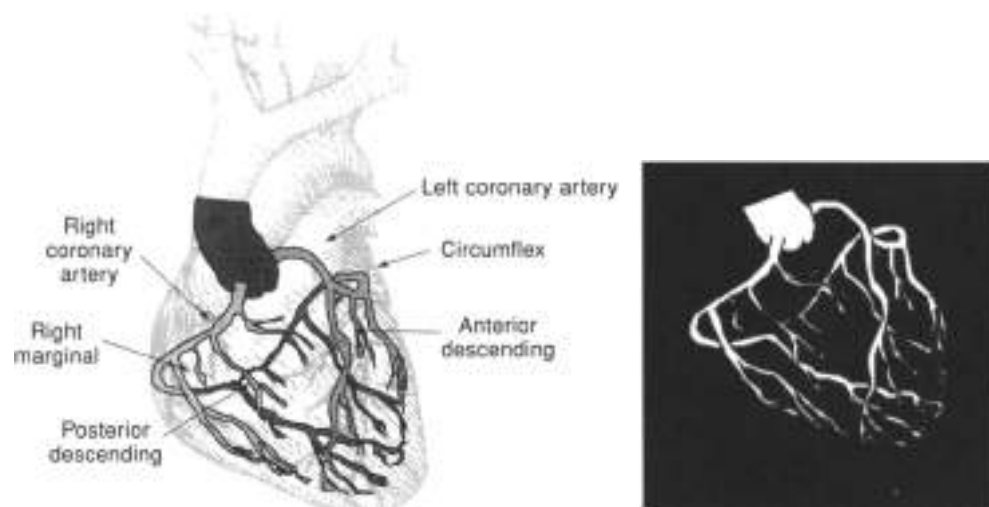
Magnetic resonance imaging can also image the motion of the walls of the heart during each heart beat and though currently more cumbersome to operate, MRI has more complete coverage of the heart than ultrasound particularly on obese patients or patients with large breasts. MRI can also show fine detail of wall motion by specifically showing the muscle motion. A grid of saturated image pixels is placed over the heart slice using a saturating RF pulse sequence. The heart is then imaged during the cardiac cycle and the saturated pixels move with the pixel elements thus allowing determination of muscle motion and calculation of strain. Alternatively, the motion of each pixel can be deduced by phase changes associated with motion and the trajectories of muscle displacement.

Nuclear medicine methods are employed in the evaluation of ischemic heart disease. The most common test is evaluation of the difference in perfusion between rest and stress induced by treadmill exercise or a pharmacological agent. In the absence of flow from an obstructed coronary artery, no tracer will reach part of the heart muscle fed by that artery and a hole will appear on the image. Frequently, when there is some but diminished flow to part of the heart, this difference between normal and low flow can be amplified by stressing the heart which increases the flow to normal tissue by as much as four-fold without any significant increase in flow to the compromised tissue. The difference between rest images and stress images is used to diagnose viable from infarcted tissue.

As emission methods do allow specific chemical and metabolic processes to be detected, PET has been used to show glucose uptake (see heart uptake in Fig. 11), fatty acid metabolism, and aminocyclometal status of the heart. It has been discovered that the viable but jeopardized (low coronary flow) regions of the heart accumulate an analogue of glucose labeled with  $^{18}\text{F}$  and this fact has been used to show that patients with a seeming absence of flow by perfusion methods but an accumulation of FDG do have some viable tissue and are candidates for revascularization surgery.

## Coronary Angiography

The major use of X-ray imaging of the heart other than the routine chest X-ray is coronary catheterization performed for the



**Figure 14.** Coronary angiography using invasive catheters is expected to be replaced by noninvasive MRI using spiral imaging methods and intravascular contrast material. This is a cartoon of the coronary arterial tree.

detection of coronary atherosclerosis. The procedure involves inserting a catheter in the femoral artery of the groin and feeding this catheter into the opening to the left and right coronary arteries which are at the aorta near the base of the heart. Contrast material having a high concentration of iodine is injected selectively into each coronary artery. Separate catheters are usually used for each coronary and the procedure is carried out by the cardiologist or radiologist in a hospital setting. This procedure, though considered the gold standard for definitive diagnosis of coronary atherosclerosis, is uncomfortable and costly, thus alternative methods for evaluation of the heart have been sought.

MR angiography with or without use of MR contrast material is used to show the lumen of vessels just as is done with X-ray arteriography but presently with less resolution. In theory it is possible to image the coronary artery down to about 0.5 mm using intravascular contrast material and fast imaging methods and this modality could replace the commonly performed coronary catheterization (Fig. 14).

#### Coronary Artery Calcification Detection

Calcification of the coronary arteries increases with age in the asymptomatic population, but the quantity of coronary artery calcification is high in patients with clinically symptomatic coronary artery disease. Calcification is detected by intracoronary ultrasound and noninvasively by fast CT or electron beam CT (EBCT) wherein 100 ms duration scan times and EKG gating minimize motion blurring, which have hindered fluoroscopic and conventional CT for detection and quantification of calcium deposits. The fact that 87% of the asymptomatic subjects have calcium quantities above the 75th percentile for age and sex has led to the recommendation that older patients with risk factors (e.g., high cholesterol) for a coronary event have an EBCT (21).

## GASTROINTESTINAL

### Liver

Two major radiological applications of ultrasound have been detection of primary and secondary cancer of the liver and gall-

stones in the gallbladder. The principal radiological diagnostic problem in the evaluation of the liver is the detection of cancer metastases from cancers elsewhere in the body and the differentiation of solid tumor masses from cysts, abscesses, and hemangiomas. Liver imaging is also valuable in evaluation of cirrhosis and fatty infiltration. X-ray CT with contrast agents, ultrasound, radionuclide methods, MRI, and most recently spiral X-ray CT have been compared in their accuracy to detect and differentiate liver lesions (22). The accuracy of US and CT is generally believed to be about 60% to 60%.

A major problem of liver imaging in the past has been blurring and artifacts from normal liver motion during the tomographic scan. Motion, until recently, has rendered MRI less valuable than ultrasound or X-ray CT. But since 1994 methods of compensation for motion and faster imaging strategies of MRI and spiral X-ray CT have resulted in significant improvement in diagnoses. X-ray CT is the most widely used cross-sectional method for liver tumor evaluation in the United States. However, with the introduction of different MRI contrast agents which have specificity for normal parenchymal Kupfer cells, extracellular space or blood pool, the diagnostic potential of MRI might compete with spiral X-ray CT and possibly avoid needle biopsy frequently required to make the diagnosis of liver masses. MRI is the best technique for evaluating fatty infiltration and both MRI and ultrasound are effective in evaluating liver vasculature.

### Colorectal Tumors

Colorectal tumors are usually suspected from symptoms of pain and bleeding with confirmation by conventional barium enema and colonoscopy. Since 1990, spiral X-ray CT (see above) has been utilized for the study of the abdomen and pelvis because the rapid coverage avoids motion artifacts from respiratory-based organ motion. Spiral X-ray CT data from the air-filled colon has been used effectively and is the basis for virtual colonoscopy. A major diagnostic problem is the evaluation of tumor therapy. PET was found to be superior to pelvis X-ray CT in detecting local disease recurrence (23). The accuracy for local disease recurrence was 95% for PET and 65% for pelvis CT. No significant difference was found for the accuracy in detecting lesions in the liver, 98% for PET and 93% for CT or US.

PET whole body scanning detects dissemination of colorectal tumors (24).

## GENITAL AND URINARY

### Fetus and Pregnancy

Ultrasound using either an abdominal probe or a transvaginal probe is used to confirm the presence of intrauterine pregnancy at 4 to 5 weeks of gestation, the heart beat at 6 weeks, and subsequently the normal or abnormal development as well as the gestational age by the length of the embryo/fetus.

Abdominal sonography has a vital role in deciding whether amniocentesis should be performed for genetic purposes and for detection of lung maturity. The electronic control of data from modern ultrasound transducers (Fig. 8) allows three-dimensional data extraction such that the face and other surfaces of the fetus can be visualized *in utero* (Fig. 15).

Other uses of ultrasound relative to pregnancy are evaluation of causes of infertility, facilitation of embryo transfer, and investigation of causes of abdominal pain and masses, because X-ray methods are not used due to concern of exposing the fetus to ionizing radiation. MRI procedures are not in widespread accepted use at least in the first 2 trimesters though there is no known reason not to take advantage of MRI diagnostic capabilities throughout pregnancy.

### Female Reproductive Organ Cancer

Ultrasound of the female pelvis is routinely used to depict the normal pelvic anatomy and to demonstrate both physiological and pathological changes of the ovaries, uterus, and cervix. Whereas ultrasound at 3.5 to 5 MHz provides diagnostic information of clinically suspected disease, X-ray CT and MRI are used to provide accurate localization for both biopsy and radiotherapy.

Carcinoma of the uterine cervix is the second most common malignancy in women and accounts for two-thirds of malignant tumors found in the female genital tract. Neither ultrasound nor X-ray CT are clinically useful in the diagnosis of early

disease, but X-ray CT can show local spread in more advanced disease. MRI is now established as a highly accurate method of demonstrating invasive cervical carcinoma but in early disease MRI underestimates superficial cervical carcinoma. These modalities are also of limited effectiveness in detection of carcinoma of the uterus which is one-third as common as cervical carcinoma. X-ray CT for uterine cancer requires use of intravenous contrast which enters the normal myometrium better than into the cancer. A common benign but clinically important condition of the uterus is fibroids which are usually accurately detected by ultrasound.

Both ultrasound and X-ray CT are used in the definition of ovarian tumors and cysts but the differentiation is frequently difficult. The accuracy of MRI in distinguishing benign from malignant ovarian cancers remains controversial. Overall imaging techniques currently have a limited supportive role in the initial assessment of ovarian cancer but MRI appears to be superior in depicting local tumor infiltration of the bladder, pelvic fat, and other tissues including metastases to the liver and local invasion of the myometrium. Although MRI is superior to X-ray CT, X-ray CT is currently preferred for the assessment of the entire abdominopelvic cavity as it is more widely available, less expensive, and is a method with which there is a vast experience.

### Kidney

The major diseases for which radiological procedures are needed include, kidney injury, kidney artery stenosis (e.g., atherosclerosis), kidney failure, kidney carcinoma, and kidney transplant rejection. X-ray CT, ultrasound, radionuclide procedures, and MRI all have a role in the diagnosis of kidney disease. X-ray CT is the procedure of choice in evaluation of the patient with suspected kidney trauma. Use of spiral X-ray CT with breath holding and contrast agents provides a comprehensive evaluation for many situations, however, equivalent and in some cases supplemental information can be gleaned from use of fast MRI methods with injected contrast material (e.g., Gd-DTPA) as shown by the comparison in Fig. 16. Ultrasound methods provide a convenient approach to determination of dilatation of the collection system in order to determine the presence of obstruction which can be caused by obstruction of an ureter or even benign prostatic hypertrophy.

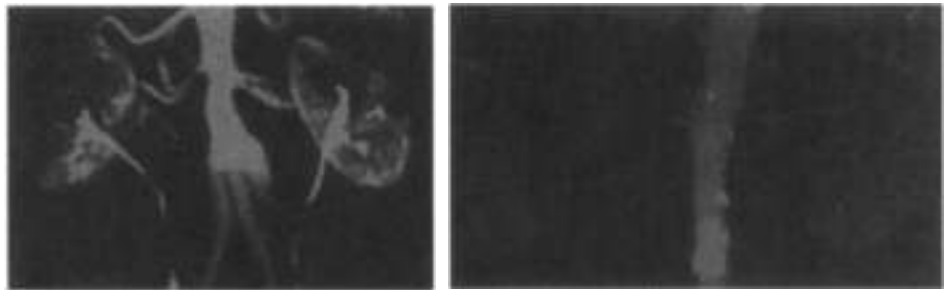
### Prostate

The two major diseases of the prostate gland are benign prostatic hypertrophy and prostate cancer. Prostatitis is an acute or chronic bacterial infection diagnosed and treated medically. Prostate cancer is the second commonest cause of cancer death in American men over age 55. It is the commonest human cancer found at autopsy in 30% of men at age 50 and in 90% at age 90 but usually as a latent disease. The best techniques for early diagnosis are measurement of prostatic serum antigen (PSA) and the rectal exam. The radiological procedures are transrectal ultrasound (TRUS) and transrectal MRI. Though neither technique can definitively separate prostatitis from carcinoma in all cases, both techniques can help define the likelihood of cancer and TRUS as well as X-ray CT are used to guide biopsy needles to confirm cancer and metastases. Both techniques help define whether the cancer has invaded the capsule or tissues outside the prostate and therefore aid in the therapeutic choices. When cancers are confined to the prostate gland they are curable by surgical removal of the prostate gland. X-ray



**Figure 15.** Three-dimensional ultrasound image of the face and hands of a unborn fetus in the mother's uterus (courtesy of Alnka Co. Ltd.)

**Figure 16.** Post MRI with contrast material (left—courtesy of GE Medical Systems) reveals major vessels as well as kidney function through visualization of the ureter because the contrast material is excreted into the ureters by the kidney. X-ray CT (right—courtesy of Siemens Medical Systems) using contrast material shows vessels and the presence of severe atherosclerosis revealed by multiple large calcium deposits in the aorta.



CT is not now used for routine tumor staging but is useful in advanced cancer when lymph nodes are enlarged.

PET and SPECT are not useful in defining the primary disease because radionuclide accumulation in the bladder interferes with imaging the adjacent prostate. The prostate tumor metastases have low avidity for static accumulation of fluorodeoxyglucose (FDG) in bone but soft tissue metastases are detected by PET (Fig. 11). Recently, MRS (magnetic resonance spectroscopy) superimposed on MRI has been effective in detection of prostatic tumors and in evaluation of therapy using an image of the ratio of choline containing compounds to citrate obtained from spatial maps of spectral information or chemical shift imaging. The normal prostate shows a high concentration of citrate on proton MRS and tumors have a relatively high signal from the trimethyl groups of choline-containing compounds.

## SKELETAL SYSTEM

### Joint Disease

The three main categories of joint disease are injury, rheumatoid arthritis, and osteoarthritis. The important tissues of the joint are the synovial membrane, which is like a sleeve around the joint; the cartilage material on the surfaces of the articulating bones, which provides a low coefficient of friction surface allowing smooth joint motion; and the ligaments and tendons, which attach the articulating surfaces or are attached to these surfaces.

Athletic injuries to shoulders and knees are common injuries which require diagnostic methods of X-ray CT and MRI. MRI has provided a major advance in the diagnostic imaging of the shoulder (rotator cuff) and knee (meniscus, cartilage) injuries (Fig. 17). Meniscal MRI, however, has a unique role. In many circumstances it substitutes for other techniques that are either less accurate (physical examination), largely operator dependent, mildly invasive and expensive (i.e., arthrography and diagnostic arthroscopy). MRI of the knee meniscus is one of the most efficient uses of this technology. Although MRI certainly has a role in evaluating muscle, tendon, and ligament pathology in the knee, clinical assessment determines therapy.

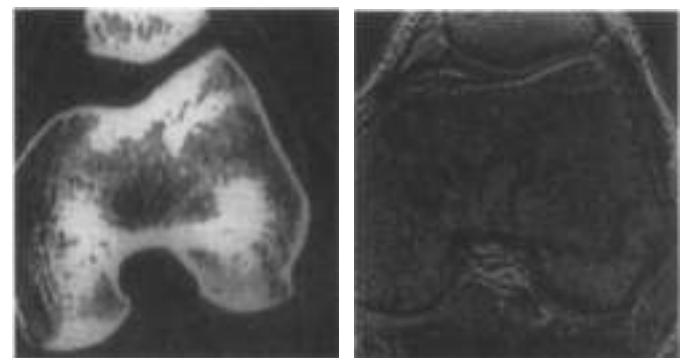
Rheumatoid arthritis is a systemic disease and at the joint level major pathophysiological mechanisms include immune, inflammatory, and healing or remodeling reactions. This disease starts as an inflammatory reaction of the synovium from an unknown cause. This reaction attracts white blood cells with a resulting swelling, edema, and soft tissue laxity, thus weakening the joint and causing dysfunction between opposing bones. The nutrient supply to the cartilage is from nonvascular

diffusion into and out of the synovium; thus, processes which influence the function of the synovium will inevitably affect functioning of cartilage which depends on a balance between collagen, proteoglycan polysaccharide matrix, water, and ions for its health.

Uncontrolled inflammatory responses of the white cells break down collagen and components of the cartilage with eventual destruction of the cartilage, a narrowing of the joint space and destruction of the adjacent surfaces of the bone of the joint. There also occurs a loss of bone early in the disease due to the inflammatory reaction which causes increased local destruction of the bone. The pattern of rheumatoid arthritis is also associated with some general disorders such as systemic lupus erythematosus, ankylosing spondylitis, scleroderma, and juvenile chronic polyarthritis but there is a high prevalence of rheumatoid arthritis of the extremities with no known etiology.

The radiological pattern of rheumatoid arthritis is that of progressive changes in the joint tissue symmetry, bone alignments, and patterns of swelling which can be chronic or intermittent. Osteopenia (local regions of loss of bone mineralization) inevitably occur in spite of the treatment. These patterns are readily observed by plane X-ray examination of the hands and other joint spaces. MRI now can give exquisite details of synovium and cartilage changes (Fig. 17) particularly using sequences such as magnetization transfer.

Osteoarthritis is the most common joint disease and is characterized by noninflammatory cartilage loss accompanied by new bone formation. Joint trauma and joints undergoing repeated stress (knees and ankles of athletes (Fig. 17) and



**Figure 17.** Different types of information provided by X-ray CT and high resolution MRI of the human knee (right—courtesy of S. Myung, University of California, San Francisco). The difference in magnetic susceptibility between trabeculae of bone and marrow allow high resolution and high contrast imaging.



**Figure 18.** MRI of joints can demonstrate defects in cartilage, ligaments, and bone with greater specificity than possible by other methods including arthroscopy. This image shows a lateral meniscus tear and cyst of the knee (courtesy of D. Rubin, University of Pittsburgh).

laborers) are probably the major causes of what is known as secondary osteoarthritis. The primary osteoarthritis currently of unknown etiology is believed to be the result of an intrinsic abnormality of cartilage leading to its degeneration. Both types have the radiological pattern of joint space narrowing and sclerosis of the opposing surfaces of subchondral bone. Deviations and subluxations between articulating bones develop slowly particularly in the fingers and knees. The abnormal bone growth in the fingers is often noted in the elderly as joint prominences known as Heberden's nodes. The radiological pattern of narrowed joint spaces is easily visualized by the standard X-ray, but MRI can give much more detail of the pathology. Figure 15 shows a comparison of spiral X-ray CT and high resolution MRI of the knee, tibia and patella.

#### Vertebral Spine Diseases

Degenerative disk disease is the most common pathology of the vertebrae and may occur anywhere in the spine but is most dominant in the lower cervical and lower lumbar spine. It is characterized by disk-space narrowing, sclerosis of the vertebral body endplates, generation of osteophytes (bone spurs). The basis for much of lumbar spine disease is from the disk cartilage degeneration due to abnormal physical stress-related biomechanical factors but also narrowing of the spinal column (spinal stenosis) by abnormal bone growth during aging plays an important role in progressive low back spinal cord symptoms. Abnormally high or sustained loading on the human disk cartilage results in water absorption from the nucleus pulposus by adjacent vertebrae, osmotic pressure change in the cartilage, stress on the collagen fibers of the disk annulus, bulging of the disk, and subsequent detachment of disk material on the nerve roots. Congenital or acquired narrowing of the spinal canal (spinal stenosis) through abnormal bone growth is frequently associated with disk disease. Spurs of the bony process of the joints of Luschka in the cervical spine result frequently in neurological compressive symptoms. Comprehensive evaluation of debilitating symptoms of spinal column and nerve root compression is now provided by MRI examination of the cervical and lumbar regions (Fig. 19).



**Figure 19.** MRI of sagittal view through the cervical spine shows a disk defect which is impinging on the spinal cord. MRI of the spine is also the optimum diagnostic imaging method for low back pain.

#### Bone Tumors

Bone tumors and tumorlike lesions are most commonly detected by conventional radiography. MRI and X-ray CT are used to determine the extent of the tumor. MRI's capability to distinguish general tissue types has been valuable in differentiating benign from malignant tumors. Radionuclide imaging is used to evaluate metastases of neoplastic tumors (e.g., lung, prostate, and breast) to bone. For this purpose conventional gamma camera planar imaging (scintigraphy) is one of the most often performed nuclear medicine procedure after injection of  $^{99m}\text{Tc}$ -phosphate complex (e.g., pyrophosphate, diphosphate salt). The  $^{99m}\text{Tc}$  complex is readily available at all hospitals with a gamma camera and after injection 50% of these bone-seeking radiotracers is deposited in the body skeleton as a technetium-calcium-phosphate complex one hour after injection. Whole body planar images are obtained. The amount of local uptake reflects the metabolic state independent of the amount of bone mineralization. The increase in blood flow accompanying the presence of increased metabolism is the major cause for increased uptake in tumors and areas of rapid bone turnover (e.g., osteoarthritis).

#### Osteoporosis

Osteoporosis is the most common metabolic bone disorder beginning in the fifth or sixth decade in women and the sixth or seventh decade in men. The loss of bone mineral is 3% to 10% per decade. The prevalence of osteoporosis in the aged is high and usually evades routine diagnosis because as much as 60% of bone mineral loss will occur before the osteopenia is detected on conventional X-ray examination. Postmenopausal osteoporosis is identified by a fracture pattern involving vertebral bodies and frequently the wrist (distal radius). Senile osteoporosis presents by fractures involving the hip and vertebrae. There is bone loss of both trabeculae and cortex but in the main

the loss is of secondary trabeculae. Fractures in the elderly following minimal trauma as well as back pain frequently lead to diagnostic examination of bone density quantitation using two energy bands of photons either from a dual gamma ray emitting radionuclide such as gadolinium-153 or a single X-ray source operated with filters or different energies to provide two energy bands such that the differential absorption between calcium and the elements of soft tissue can be measured. The bone density is interpreted as a percentage change from a normal or from that expected for a given age. Recent advances in MRI high resolution imaging show promise of providing architectural information of trabecular bone undergoing changes as shown by a comparison between high resolution X-ray CT and high resolution MRI of the knee (Fig. 18).

### Vascular System

Radiological methods of imaging the arteries and veins of the body generally have used injection of iodinated contrast material followed by rapid imaging using plane X-ray film or digital subtraction methods. These methods involved inserting catheters in arteries and though precise in diagnostic specificity, they have some morbidity. Yet they remain the major methods for identification of atherosclerosis in brain, carotid, coronary, renal, aorta, and leg arteries. There are three methods of visualizing these arteries using MRI. The phase-sensitive technique uses a bipolar flow encoding gradient to cause a phase shift for a moving proton but a zero phase shift or canceled phase shift for a stationary spin. By reversing the polarity of the flow encoding gradient on alternate acquisitions and subtracting these acquisitions the movement of the protons provides the signal. The time-of-flight method relies on magnetization of spins into an excited volume wherein the spins have been saturated thus the unsaturated flowing blood will have a detectably higher signal compared to the surrounding previously saturated parenchyma.

A third method which will replace much of the conventional X-ray angiography uses injected contrast material which unlike previously used MRA contrast material stays in the blood pool (intravascular contrast). This material is gadolinium complexed to albumin which does not diffuse into tissues due to the large molecular size of albumin. The new contrast material overcomes limitations of the past related to temporal changes of contrast material which diffuses through the extracellular space. Image acquisition from moving organs and an imaging time not dependent on the rapidly changing concentration of the injected bolus allows high-resolution angiography. This method has promise to visualize most of the important elements of the arterial blood pool including the coronary arteries down to less than 1 mm before year 2000 (see Fig. 14).

Flow quantitation particularly in the carotid arteries has been provided by Doppler ultrasound techniques which are in widespread use and when combined with ultrasound imaging the method is known as duplex scanning. Doppler images in color denoting direction (red for artery and blue for venous) with intensity related to speed are superimposed on the gray level 2-D image which for the carotid shows the vessel and surrounding soft tissues. A new technique known as Power Doppler superposes the intensity of the Doppler signal rather than direction and speed thus this image represents the volume of moving red cells. This technique has been applied with success to a number of problems including detection of malignant lymph nodes from breast cancer and evaluation of kidney disease.

### Picture Archiving and Communication System (PACS)

A major component of diagnostic radiology is the digital data-based image storage, transfer, and processing systems embodied in software and hardware of PACS (picture archiving and communications system). PACS, developed many years ago for archiving and digitally transferring conventional radiographs after digitization, has established the methodologies and formats needed for the management of digital data acquired by contemporary radiologic imaging devices (e.g., MRI, PET, SPECT, US) (25). The very high bandwidth data currently being handled by PACS includes routine angiography image sequence involving  $512 \times 512$  images taken at 30 frames per second. For transmission and display, optical fiber-based networks based on asynchronous transfer mode (ATM) are used. Workstations for review of these data can handle a complete or full injection sequence in "loop RAM (random access memory)" using the  $512 \times 512$  format interpolated to  $1024 \times 1024$  while displaying at 30 frames per second.

A common standard for image transfer, storage, and retrieval is the DICOM (digital imaging and communications in medicine) standard. Standards for interfacing image data acquired by different modalities (PACS) with radiology information systems (RIS) and hospital information systems (HIS) will allow integration of patient images with the radiologist's interpretation, clinical data, and demographic information for ready access by authorized health-care personnel. These capabilities enable a new dimension in teleconferencing and local workstation-based image processing for 3-D visualization, segmentation, and contrast enhancement. This technology not only enables rapid access to patient information but expedites consultation, education, and research.

### ACKNOWLEDGMENTS

This work was supported by the Office of Health and Environmental Research of the U.S. Department of Energy under contract DE-AC03-76SF00098 and National Heart, Lung, and Blood Institute. Dr. Kathleen Brennan assisted in the manuscript.

### BIBLIOGRAPHY

1. E. Kreutel (ed.), *Imaging Systems for Medical Diagnostics*, Berlin and Munich: Siemens, 1990.
2. J. D. Bronzino (ed.), *The Biomedical Engineering Handbook*, Boca Raton: CRC Press, IEEE Press, 1995.
3. T. F. Budinger et al. (ed.), *Mathematics and Physics of Emerging Biomedical Imaging*, Washington, DC: National Academy Press, 1996.
4. W. A. Kalender et al., Spiral volumetric CT with single-breath-hold technique, continuous transport, and continuous scanner rotation, *Radiol.*, **176**(1): 181-5, 1991.
5. S. A. Feig and M. J. Yaffe, Digital mammography, computer-aided diagnosis, and tele-mammography, *Radiol. Clinics North Amer.*, **33**(8): 1205-30, 1995.
6. A. Haase et al., FLASH imaging: Rapid NMR imaging using low flip angle pulses, *J. Magn. Resonance*, **67**: 258, 1993.
7. S. R. Cherry and M. E. Phelps, Positron Emission Tomography: Methods and Instrumentation, in M. E. Sandler et al. (eds.), *Diagnostic Nuclear Medicine*, Baltimore: Williams & Wilkins, 1985, pp. 121-133.



8. T. F. Redinger, Single Photon Emission Computed Tomography, in M. P. Sandler et al. (eds.), *Diagnostic Nuclear Medicine*, Baltimore: Williams & Wilkins, 1987, pp. 121-134.
9. J. B. Wyngaarden and L. H. Smith Jr. (eds.), *Textbook of Medicine*, Philadelphia: Saunders, 1984.
10. M. J. Mack et al., Thorascopy for the diagnosis of the indeterminate solitary pulmonary nodule, *Ann. Thoracic Surg.*, **60**(4): 825-30, discussion 830-2, 1993.
11. E. F. J. Paix and P. C. Goodman, Positron emission tomography imaging of the thorax, *Health Care North Amer.*, **32**(4): 811-21, 1994.
12. E. E. Kim et al., Differentiation of residual or recurrent tumors from post-treatment changes with F-18 FDG PET, *Radiographics*, **12**: 269-279, 1992.
13. D. G. Freeman, J. Suyemoto, and S. Weisler, Frequency of coronary atherosclerosis in man, *New England J. Med.*, **272**: 1278-1280, 1965.
14. C. E. Spritzer et al., Detection of deep venous thrombosis by magnetic resonance imaging, *Chest*, **104**(1): 54-60, 1993.
15. M. Bemy-Jordan et al., Diagnosis of pulmonary embolism with spiral CT: Comparison with pulmonary angiography and scintigraphy, *Radiol.*, **200**(3): 699-703, 1993.
16. Institute of Medicine (ed.), *Strategies for Managing the Breast Cancer Research Program*, Washington, DC: National Academy of Sciences Institute of Medicine, 1993.
17. S. H. Heywang et al., Dynamische Kontrastmitteluntersuchungen mit FLASH bei Kernspintomographie des mamma, *Diagn. Radiol.*, **8**: 7-13, 1988.
18. R. P. Kedar et al., Automated quantification of colour Doppler signals: A preliminary study in breast tumours, *Radiol.*, **197**: 33-42, 1995.
19. C. I. Perez, V. C. M. Kim, and P. de Boege, Colour Doppler ultrasonography in the diagnosis of axillary lymph node metastases in breast cancer, *The Breast*, **6**: 10-12, 1993.
20. R. P. Kedar et al., Microbubble contrast agent for colour Doppler US: Effect on breast masses. Work in progress, *Radiol.*, **3**: 679-686, 1991.
21. E. B. Kaufmann et al., Quantity of coronary artery calcium detected by electron beam computed tomography in asymptomatic subjects and angiographically studied patients, *Mayo Clinic Proc.*, **70**: 223-232, 1995.
22. R. E. Larson et al., Hypervascular malignant liver lesions: Comparison of various MR imaging pulse sequences and dynamic CT, *Radiol.*, **182**(2): 389-93, 1994.
23. C. Schiepers et al., Contribution of PET in the diagnosis of recurrent colorectal cancer: Comparison with conventional imaging, *Eur. J. Surg. Oncol.*, **21**(7): 517-522, 1995.
24. G. Beets et al., Clinical value of whole-body positron emission tomography with <sup>18</sup>Ffluorodeoxyglucose in recurrent colorectal cancer, *Br. J. Surg.*, **81**(11): 1666-1670, 1994.
25. H. K. Huang, *IACS, Basic Principles and Applications*, New York: Wiley, 1998.

THOMAS F. REDINGER  
University of California at Berkeley

## FREQUENCY CONVERTERS AND MIXERS

An essential part of most microwave receivers is the frequency converter, a device that converts the frequency of an incoming signal to another frequency. The output frequency may be *downconverted*, *upconverted*, *multiplied*, or *divided*. Important and common is the *mixer*, which downconverts a high-frequency input signal  $f_i$  to a much lower and more man-

ageable signal  $f_{IF}$ , preserving information concerning the amplitude, frequency, and phase of the input signal. The devices used are nonlinear (i.e., the relationship between current and voltage is not linear). The devices may have either two terminals (diodes) or three terminals (transistors).

Mixers are important for very high frequencies, where amplifiers are not available and direct amplitude or phase detection is difficult. A mixer can be used to downconvert, say, a terahertz frequency signal to a microwave frequency one, where electronic methods are readily available for amplification and any kind of demodulation. In fact, in almost any radio set or mobile telephone receiver or base station, there are several mixers and other types of frequency converters.

Both two-terminal devices (diodes) and three-terminal devices (transistors) are used in mixers. The frequency conversion is accomplished by using the *nonlinear* properties of the device. Virtually all semiconductor devices, such as diodes and transistors, show nonlinear properties in certain bias ranges. Common devices for microwave mixer applications are the Schottky diode, the field-effect transistor (FET), and the bipolar transistor. There are many other devices available as well, for example, the superconducting tunneling device (STJ) and the superconducting hot electron device for low-noise (high-sensitivity) millimeter and submillimeter wave receivers. For infrared wavelengths, metal-insulator-metal devices and, for optical frequencies, photoconducting devices have been used. Note that ordinary resistors, capacitances, and inductances are linear components.

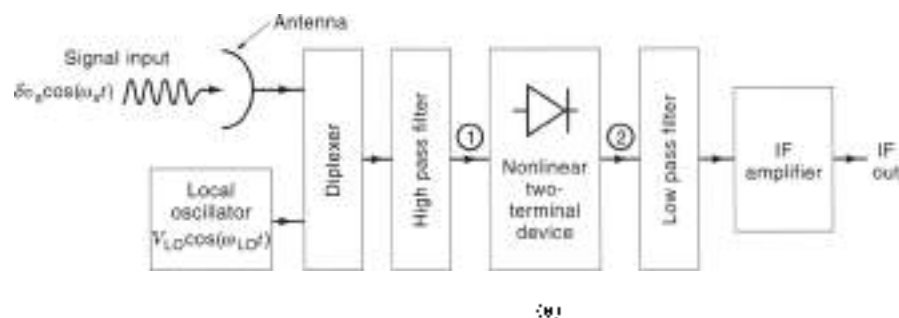
Figure 1 gives an example of how mixers are used in a receiving system. The antenna is connected directly to the mixer. For example, in a TV-satellite receiver, an amplifier is placed just after the antenna to increase the signal amplitude. In reality there are many other systems (e.g., radar, radio, measurement systems) where mixers are used to frequency downconvert the input signal.

To give a simple illustration of how a diode mixer may work, consider the detector circuit shown in Fig. 2 using an ideal diode (zero resistance in the forward direction and infinite resistance in the backward direction). If a sinusoidal voltage  $V_i(t) = V_{LO} \cos(2\pi f_{LO}t)$  is applied, it will be "rectified" (detected) by the diode. The voltage over the resistance  $r_L(t)$  will be a constant dc voltage proportional to  $V_{LO}$ . Next, add a small signal voltage  $\delta v_i \cos(2\pi f_s t)$ , that is,  $V(t) = V_{LO} \cos(2\pi f_{LO}t) + \delta v_i \cos(2\pi f_s t)$ . Assuming that  $\delta v_i/V_{LO} \ll 1$ , the resulting voltage  $V(t)$  will become amplitude modulated, as shown in Fig. 2. The detected voltage  $v_R(t)$  over the load resistance  $R$  will be proportional to the envelope  $V_R(t)$  (assuming  $\tau = RC \ll 1/\omega_{IF}$ ) (see Fig. 3), that is,

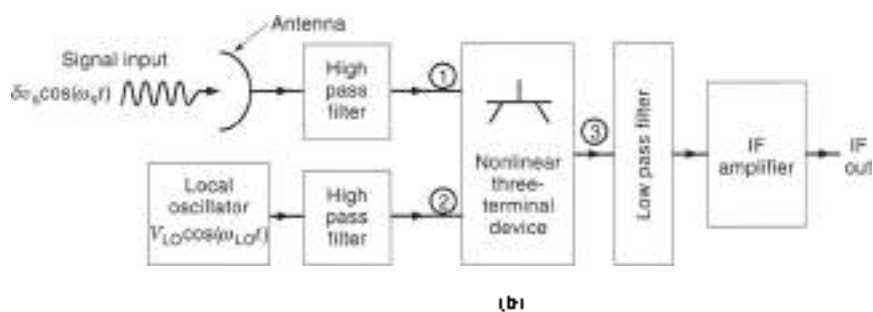
$$v_R(t) = V_{LO} \left( 1 + \frac{\delta v_i}{V_{LO}} \cos(2\pi f_{IF}t) \right) \quad (1)$$

where  $f_{IF} = f_{LO} - f_s$  is the intermediate frequency (IF). The dc part of Eq. (1) is the "detected local oscillator (LO)" and the alternating part is identical to the IF voltage and, in this example, is equal in amplitude to the input signal. This IF signal is fed into an amplifier as described in Fig. 1.

Note that in the simplified example above we have not correctly accounted for a number of parameters, such as the impedances of the LO and signal sources. This means that the IF voltage will not become equal to  $\delta v_i$ . A more detailed description of a more correct calculation is given in the section



**Figure 1.** Typical mixer block diagram. At the input there is a high pass filter that will prevent any low-frequency IF power from escaping to the mixer input (left). The low pass filter will stop any input signal or LO power from going to the IF circuit (right). Part (a) shows a layout for a two-terminal device (diode) and (b) for a three-terminal device (transistor).



entitled “Schottky Diode Mixers.” For details concerning microwave mixers, see Refs. 1 and 2.

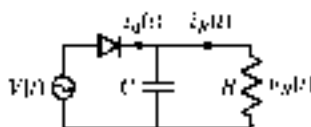
**GENERAL PROPERTIES OF TWO-TERMINAL NON-LINEAR DEVICES**

Below we describe some general results obtained when a two-terminal device such as the Schottky diode is excited with a sinusoidal signal. The impedance of the Schottky diode is voltage dependent. The speed of this device is indeed high. It shows a nonlinear behavior up to several terahertz.

**Frequency Multiplication**

Consider a nonlinear diode device exposed to  $I(t)$  (or pump) power yielding a large voltage swing  $V(t) = V_{LO} \cos(\omega_{LO} t)$  over the diode. (Note. Below we use  $\omega = 2\pi f$ ). Since the relation between current and voltage is not linear (i.e.,  $I \neq \text{const.} \times V$ ), the resulting current will not have a sinusoidal shape like the input voltage. However, the current in this case is still a periodic function versus time with the same periodicity,  $\tau = 2\pi/\omega_{LO}$ , as the LO frequency and can consequently be expressed as a Fourier series with harmonics of the LO:  $\omega_{LO}, n$ .

$$I(t) = \sum_{n=0}^{\infty} i_n \cos(n\omega_{LO} t + \varphi_n) \tag{2}$$



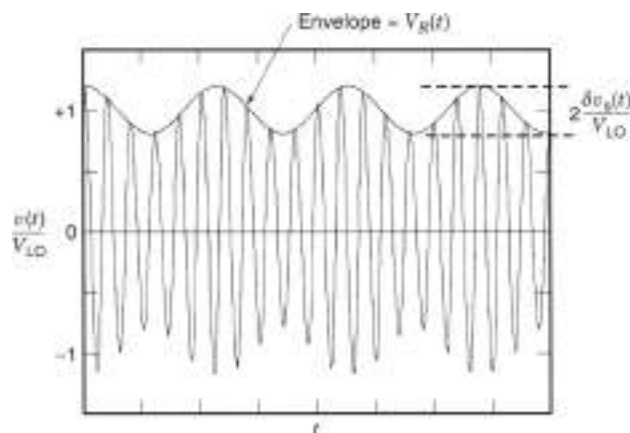
**Figure 2.** A rectifying diode circuit. If the capacitance  $C$  and the resistance  $R$  are large enough,  $i_R(t)$  will follow the envelope of  $V(t)$  (see Fig. 2).

The component  $i_0$  is the dc component. To obtain power at a particular harmonic, for example, the third harmonic, it is required that the current component at  $3\omega_{LO}$  be passing through a resistance  $R_1$  delivering a power of  $1/2 i_3^2 R_1$ . To avoid any power being delivered at other harmonics it is necessary to ensure that the device is reactively terminated at these harmonics.

In reality, the impedance of most nonlinear devices is complex with both the real and the imaginary parts voltage (or current) dependent. For a more detailed theory, see the section entitled “Large Signal Analysis by Harmonic Balance.”

**Frequency Conversion**

If two signal voltages at  $\omega_{LO}$  and  $\omega_s$  are simultaneously interacting with the nonlinear diode impedance, the resulting current



**Figure 3.** An example with a drive voltage  $v(t)/V_{LO} = \cos(\omega_{LO} t) + \delta v_s/V_{LO} \cos(\omega_s t)$ , where  $\omega_s/V_{LO} = 0.2$  and  $\omega_{LO}/\omega_s = 1/16$ . The IF is  $0.18 \omega_{LO}$ .

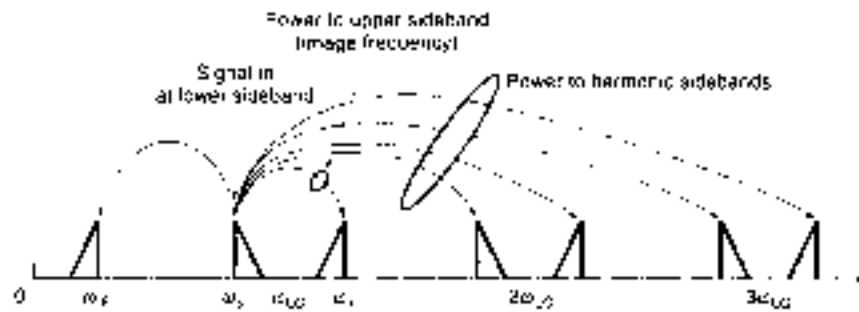


Figure 4. Power flow in a lower sideband mixer. The input signal power is distributed not only to the IF but also to the harmonic sidebands.

can be expressed in a more complex Fourier series,

$$f(t) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} I_{mn} \cos[(m\omega_{LO} + n\omega_s)t + \phi_{mn}] \quad (13)$$

In the case when both the signal voltage and the local oscillator voltage are large, a large number of frequency conversion products are obtained. Obviously it is possible to generate power at any frequency  $m\omega_{LO} + n\omega_s$ . If for  $m = n = 1$  the required output frequency is higher than the signal frequency, one has *frequency upconversion*, and if the output frequency is lower than the signal frequency, one has *frequency downconversion*.

**Linear Mixing**

As already mentioned, a mixer receiver usually handles "small signals." We have a small signal case if the amplitude of the signal at  $\omega_s$  is much smaller than the LO amplitude at  $\omega_{LO}$ . In this case only harmonics of  $\omega_{LO}$  are important and we are left with

$$f(t) \approx \sum_{n=-\infty}^{\infty} I_n \cos[(n\omega_{LO} + \omega_s)t + \phi_n] \quad (14)$$

In most mixer applications, we are interested in the "intermediate frequency,"  $\omega_{IF} = |\omega_{LO} - \omega_s|$ . An IF load resistance in the circuit will allow power at the IF frequency to be extracted (compare Fig. 2). Equation (14) suggests that power may go out only to the IF but also to *harmonic sidebands*, as illustrated in Fig. 4. The only way to prevent this power loss is to make sure that the harmonic sideband current components are facing impedances that are purely reactive.

Note that an IF signal (Fig. 4) will be created if either  $\omega_{LO} = \omega_{IF} + \omega_s$  or  $\omega_s = \omega_{LO} - \omega_{IF}$ . The former frequency  $\omega_{LO}$  is called

the upper sideband, and the latter  $\omega_s$ , the lower sideband. For a *lower sideband mixer* the upper sideband is denoted the *image frequency*, and vice versa for an *upper sideband mixer*. For a lower or upper sideband case, some signal power may go to the image frequency. In an *image reject mixer* a filter prevents the image frequency from entering the mixer. In the *image enhanced mixer*, the image terminal is terminated reactively so that the conversion loss is reduced.

At high signal powers, there may be confusion because the mixer may produce output signals in the IF band (see intermodulation below) for

$$|m\omega_{LO} - n\omega_s| = \omega_{IF} \quad (15)$$

**Harmonic Mixer**

In a harmonic and small signal linear mixer one has  $\omega_{IF} = |m\omega_{LO}| = \omega_s$ . A spectrum analyzer always uses a harmonic mixer to analyze the signal, and the harmonic number  $m$  can be very high ( $> 10$ ). The power flow in a harmonic mixer is shown in Fig. 5.

**Intermodulation**

All mixer products created by two or more signals are called *intermodulation (IM) products*. Most IM products are unwanted.

For example, if two signals at slightly different frequencies  $\omega_{s1}$  and  $\omega_{s2}$  and with a power of the same order of magnitude as the LO are interacting with the nonlinear device, the current will contain frequency products as shown in the following equation:

$$f(t) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} I_{mkn} \cos[(m\omega_{LO} + k\omega_{s2} + n\omega_{s1})t + \phi_{mkn}] \quad (16)$$

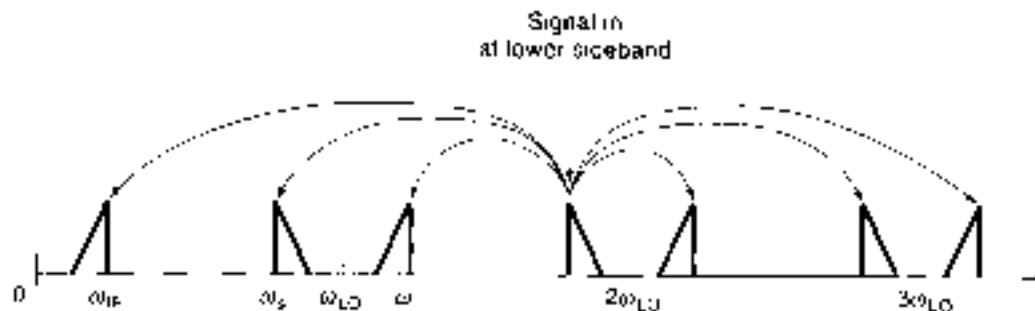


Figure 5. Power flow in a second harmonic lower sideband mixer.

All signals at frequencies fulfilling the requirement

$$\omega_{s1} = (\omega_{s0} \pm \omega_{s2} \pm \omega_{s3}) \quad \infty < \omega_{s1} < \infty \quad (7)$$

may create IM products in the IF band. Indeed, intermodulation must be considered a potentially serious problem in all applications where strong signals may occur.

## SYSTEMS ASPECTS

In a system, the following properties are important: (1) conversion loss, (2) noise properties, and (3) intermodulation properties.

### Conversion Loss

An important property of a mixer is the conversion loss  $L$ , defined as

$$L = \frac{P_s}{P_{IF}} = \frac{\text{Signal power available at the input}}{\text{IF power delivered to the IF load}} \quad (8)$$

In most practical mixers the conversion loss is larger than one. However, it is possible to obtain gain owing to parametric amplification caused by a nonlinear capacitance. In certain configurations using transistors (see Refs. 1 or Ref. 4) conversion gain can be obtained. Superconducting mixers, however, due to quantum phenomena, can show stable conversion gain.

There are several loss mechanisms causing the conversion loss:

1. Losses due to absorption in the nonlinear device
2. Losses due to power lost to harmonic sidebands
3. Losses due to reflection at the input part
4. Losses due to reflection at the IF output part

See also the section entitled "Schottky Diode Mixers."

### Mixer Receiver Noise

The important noise measure in practical applications is always the total receiver noise temperature. The contributions approximately in order of importance are: (1) mixer device noise, (2) the IF amplifier noise, (3) thermal noise from the mixer circuit, and (4) LO noise. There are two different noise measures usually cited in the literature: the *single sideband* (SSB) noise temperature and the *double sideband* (DSB) noise temperature.

When calculating the noise temperature of a mixer, it is advisable to always start adding up all noise contributions at the IF amplifier input (see Fig. 6). The temperature then becomes

$$T_n = \frac{T_1}{L_s} + \frac{T_1}{L_i} + \sum_{n=2}^{\infty} \left( \frac{T_1}{L_{s_n}} + \frac{T_1}{L_{i_n}} \right) + T_{\text{Mixer}} + T_{\text{IF}} \quad (9)$$

where  $L_s$ ,  $L_i$ ,  $L_{s_n}$ , and  $L_{i_n}$  are the conversion losses at the signal, image, and upper harmonic sidebands and lower harmonic sidebands, respectively.  $T_1$  is the noise temperature of the input resistance,  $T_{\text{Mixer}}$  is the noise from the mixer diode entering the

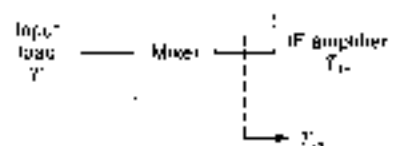


Figure 6. Receiver configuration for calculating the receiver noise. For the calculation it is wise to determine the noise temperature at the input of the IF amplifier,  $T_n$ .

IF amplifier. As the next step, simplify the signal to noise ratio:

$$\frac{P_s}{P_n} = \frac{P_s/L_s}{kT_n \Delta f} = \frac{P_s}{kT_{\text{Mixer,SSB}} \Delta f} \quad (10)$$

where  $T_{\text{Mixer}}$  is by definition the system noise temperature for a *single sideband* receiver. It is assumed that the useful and interesting signal enters only one sideband. Hence

$$T_{\text{Mixer,SSB}} = T_{\text{Mixer}} = T_1 \left[ 1 + \frac{L_s}{L_i} + \sum_{n=2}^{\infty} \left( \frac{L_s}{L_{s_n}} + \frac{L_s}{L_{i_n}} \right) \right] + T_{\text{Mixer,DSB}} + L_s T_{\text{IF}} \quad (11)$$

where we have defined the equivalent noise temperature of the mixer itself,  $T_{\text{Mixer,DSB}} = T_{\text{Mixer}} L_s$ .

For the double sideband case, one assumes a useful signal to enter both the upper and the lower sidebands. Hence the signal to noise ratio for this case should be defined as

$$\frac{P_s}{P_n} = \frac{P_s(1/L_s + 1/L_s)}{kT_n \Delta f} = \frac{P_s}{kT_{\text{DSB}} \Delta f} \quad (12)$$

that is,

$$T_{\text{DSB}} = \frac{T_n}{1/L_s + 1/L_s} = T_1 \left[ 1 + \frac{L_s L_i}{L_s + L_i} + \sum_{n=2}^{\infty} \left( \frac{1}{L_{s_n}} + 1 \right) \right] + T_{\text{Mixer,DSB}} + \frac{L_s L_i}{L_s + L_i} T_{\text{IF}} \quad (13)$$

where the double sideband noise temperature of the mixer itself is defined as

$$T_{\text{Mixer,DSB}} = T_{\text{Mixer}} \frac{L_s L_i}{L_s + L_i} \quad (14)$$

Note that if  $L_s = L_i$ , both the single sideband mixer and system noise temperatures are twice as large as for the double sideband case. The LO noise (if important) can be taken into account by adding a certain amount at the input port. Also note that the noise entering the mixer at the harmonic sidebands may considerably influence the total receiver performance.

## SCHOTTKY DIODE MIXERS

The Schottky diode mixer is the most common type of mixer for frequencies from megahertz to terahertz.

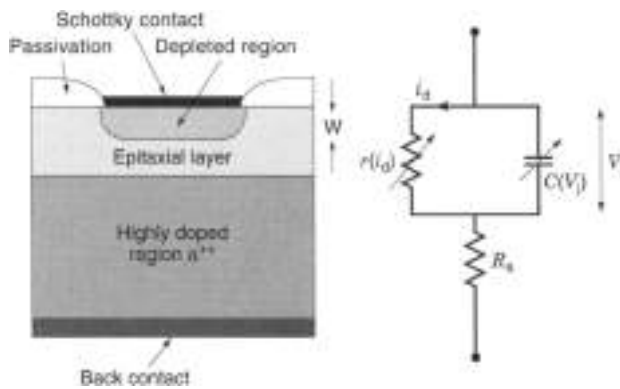


Figure 7. Simple design of a GaAs Schottky barrier diode and the corresponding equivalent circuit. The higher the frequency the smaller the diodes required.

**Schottky Diode for Mixer Applications**

A schematic diagram of a common design of millimeter wave Schottky barrier diode is shown in Fig. 7. Note that the radio frequency (RF) current is flowing from the diode contact at the surface of the diode chip to the back contact. Hence the RF series resistance is slightly larger for RF than for dc. The current-voltage (*I*-*V*) characteristic of the junction itself can be calculated from

$$i_d = I_s \left[ \exp \left( \frac{qV_j}{n k T} \right) - 1 \right] \quad (15)$$

where *I*<sub>s</sub> is the saturation current, *V*<sub>j</sub> the junction voltage, *q* the charge of the electron, *k* Boltzmann's constant, *T* the physical temperature, and *n* the ideality factor, which for good diodes at room temperature is between, say, 1.03 and 1.10.

The small signal RF junction properties of a Schottky diode can be modeled as a nonlinear resistance *r<sub>j</sub>* in parallel with a nonlinear capacitance *C<sub>j</sub>*,

$$r_{j(d)} = \frac{nV_j}{qI_d} = \frac{kT}{qI_d} \quad C(V_j) = \frac{C_0}{1 + \frac{V_j}{V_b}} \quad (16)$$

where *C*<sub>0</sub> is the zero bias capacitance and *V*<sub>b</sub> is the built in voltage of the diode. For more details, see Ref. 3.

A common measure of the high-frequency properties is the cutoff frequency, which is defined for the diode at zero bias,

$$f_c = \frac{1}{2\pi R_s C_0} \quad (17)$$

where *R*<sub>s</sub> is measured using the dc *I*-*V* characteristic. For good mixer performance this cut-off frequency must be much larger than the signal frequency. The series resistance uses up signal power and should for obvious reasons be made as small as possible. For the conversion efficiency the nonlinear capacitance is of much less importance than the nonlinear resistance. The diode is typically designed as a thin *n*-doped active layer of 1000 Å (Mott diode), which leaves one with a diode with reduced capacitance variation and a minimum series resistance.

GaAs is preferred for millimeter wave diodes. The main reason is that the high mobility of GaAs yields a low series resistance and consequently a high cutoff frequency. Another

advantage (to silicon) is that carriers do not freeze out when the diode is cooled to cryogenic temperatures in order to improve the mixer noise properties [see Eq. (7)].

The noise performance and the conversion efficiency are the prime properties of the diode when used in mixer applications. Noise properties are discussed next.

**Shot Noise.** The shot noise is due to fluctuations in the electron particle current between cathode and anode of the diode (see Ref. 8). The root-mean-square (rms) fluctuations in the current are  $\hat{i} = \sqrt{2qI \Delta f}$ , where  $\Delta f$  is a small frequency interval. The noise power of the Schottky diode can then be calculated as

$$P_n = \frac{\hat{i}^2}{4} \frac{R_s}{C(V_j)} = \frac{1}{2} k_B T \Delta f \quad (18)$$

Identifying this equation with the ordinary Johnson noise expression,  $P_n = kT \Delta f$ , it is seen that the equivalent noise temperature of the Schottky diode is

$$T_{sh} = \frac{1}{2} T \quad (19)$$

Note that the noise temperature decreases linearly with decreasing physical temperature. However, the charge transport at room temperature over the Schottky barrier is due to thermionic emission and decreases when the temperature is lowered. Hence for temperatures on the order of 50 K to 100 K and below, temperature-independent tunneling becomes the dominant process for electrons passing the barrier, and the equivalent temperature  $nT$  of Eq. (8) becomes

$$T_{tunn} = \frac{q\hbar}{k} \sqrt{\frac{N_0}{4\epsilon n^*}} \quad (20)$$

where *N*<sub>0</sub> is the doping concentration in the epitaxial layer,  $\epsilon$  the dielectric constant of the semiconductor, and *n*\* the effective mass of the electron. In practice, *T*<sub>tunn</sub> is 50 K for *N*<sub>0</sub> = 3 × 10<sup>16</sup>. This doping concentration is recommended for mixers operated at 15 K or 20 K, a typical temperature for commercial cryogenic cooling machines.

**1/f Noise.** There is excess noise at low frequencies (of the order 100 kHz and lower for a good diode), which is related to surface phenomena at the metal-semiconductor interface. This noise is normally not important in mixers for millimeter wave receivers. However, it is an important limiting factor for certain radar and communications systems.

**Thermal Noise.** The series resistance is essentially an ordinary resistor and consequently causes ordinary thermal (Johnson) noise. The main noise contribution of this type comes from the substrate (corresponding resistance *R*<sub>sub</sub>) and is denoted as *T*<sub>sub</sub>.

**Hot Electron Noise and Intervalley Scattering Noise.** Since the diode area is very small in order to make the capacitance reasonably small, a high current density is required to make *r<sub>j</sub>* *R<sub>s</sub>*. This means that electrons may obtain energies larger than

the energy related to their thermal movement and hot electron noise is obtained. The increase in energy also means that electrons can be transferred from the main  $\Gamma$  valley in the  $E$ - $k$  diagram (for details see Ref. 3) to the upper  $\Gamma$  valley, causing fluctuations in the electron velocity and intervalley scattering noise is obtained. The hot electron and intervalley scattering noises occur essentially in the undepleted part of the epilayer (resistance  $R_{ep}$ ) and are denoted as  $T_{ep}$ .

**Large Signal Analysis by Harmonic Balance**

When the nonlinear device is pumped by the I.O, harmonic currents are created (e.g., see Refs. 4, 5, or 6). The equivalent circuit of a Schottky diode mixer is shown in Fig. 8. Note that the series resistance is assumed to be linear and will be included in the embedding circuit.

The current  $I_d(t)$  contains harmonics of the I.O pump frequency,  $k\omega_{LO}$ . For a Schottky diode, this current consists of two parts: one that is associated with the nonlinear resistance  $i_R(t)$ , and one with the parallel nonlinear capacitance  $i_C(t)$ . We have [using now the complex notation; compare Eq. (2)]

$$I_d(t) = i_R(t) - i_C(t) = \sum_{k=-\infty}^{\infty} I_{dk} e^{jk\omega t} \quad I_{dk} = I_{d,k}^* \quad (21)$$

This current  $I_d(t)$  flows through the embedding circuit, creating voltages at harmonic frequencies at  $k\omega_{LO}$ ; namely,

$$\sum_{k=-\infty}^{\infty} V_k e^{jk\omega t} = V_k = V_k^* \quad (22)$$

The boundary conditions set by the embedding circuit require that

$$V_k = -I_{dk} Z_{Lk}(j\omega_{LO}) + R_k(j\omega_{LO}) \quad k = 1, 2, 3, \dots, \infty \quad (23)$$

$$V_{-1} = -V_{LO} - I_{d,-1} [Z_{L,-1}(j\omega_{LO}) + R_1(j\omega_{LO})] \quad (24)$$

$$V_0 = V_{DC} - I_{d,0} R_0(0) = R_0(0) \quad (25)$$

where  $V_{LO}$  and  $V_{DC}$  are the I.O and dc bias voltages, respectively. The frequency dependence of  $R_k$  is due to the skin effect. If  $V_k(t)$  is known,  $i_R(t)$  and  $i_C(t)$  (see Fig. 8) can be calculated from Eqs. (15) and (16). We now have a nonlinear problem to solve in order to determine  $I_{dk}$  and  $V_{dk}$ . Several iteration type methods have been suggested. See Refs. 4 and 5 for more details.

Having determined the components  $I_{dk}$  and  $V_{dk}$ , we have  $V_d(t)$  and  $I_d(t)$ , and we can determine  $i_d(t)$ ,  $i_C(t)$ ,  $g_d(t)$ , and  $C_d(t)$ .

$$g_d(t) = \sum_{k=-\infty}^{\infty} G_{dk} e^{jk\omega t} \quad G_{dk} = G_{d,k}^* \quad (26)$$

$$C_d(t) = \sum_{k=-\infty}^{\infty} C_{dk} e^{jk\omega t} \quad C_{dk} = C_{d,k}^* \quad (27)$$

These equations together with the embedding impedance  $Z_e(j\omega)$  allow us to determine the small signal properties of the mixer.

**Small Signal Analyses**

The relation between the small signal current and voltage vectors  $\delta I$  and  $\delta V$  can be expressed in a more general form [28]

$$\delta I = Y \delta V \quad (28)$$

where

$$Y_{kk} = G_{d,k} + j(\omega_k C_{d,k} + \omega_{ep} C_{ep,k}) \quad (29)$$

where for convenience we use  $\omega_{ep}$  for  $\omega_{ep}$ . It is convenient to form an augmented  $Y$  matrix,  $Y'$ , as indicated in Fig. 9. This augmented network contains the whole mixer, including diode and embedding network, but does not contain signal sources associated with these terminations. Since we define the signal sources as current sources, the augmented network in this case is open-circuited. For the augmented network

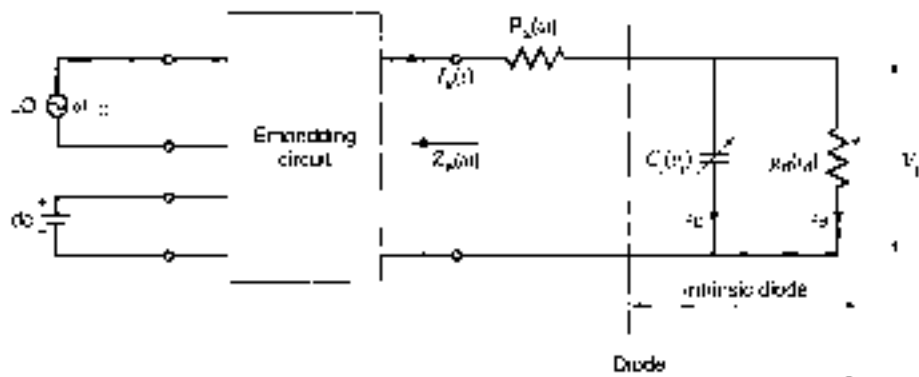
$$\delta I' = Y' \delta V' \quad (30)$$

and

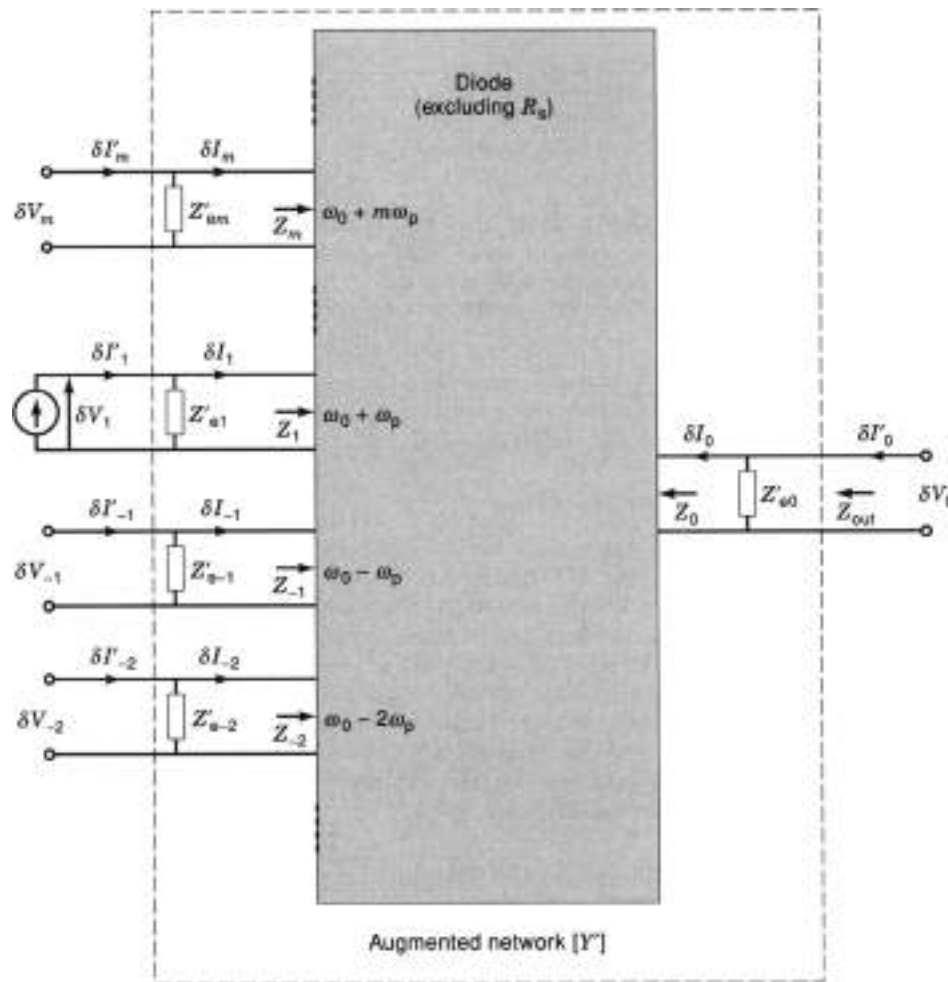
$$Y' = Y + \text{diag} \left[ \begin{matrix} 1 \\ Z_{L,-1} \\ 1 \\ R_{ep} \end{matrix} \right] \quad (31)$$

Inverting Eq. (23) yields

$$\delta I = Z' \delta V \quad (32)$$



**Figure 8.** Equivalent circuit of a mixer. For the intrinsic diode,  $C_d$  and  $g_d$  are non-linear and are characterized in the time domain, while the diode series resistance  $R_s$  and the embedding impedance  $Z_e$  are linear and can be described in the frequency domain.



**Figure 9.** The small signal representation of the mixer as a multifrequency linear multiport network. Notice that  $R_s$  is included in  $Z_{in} = Z_{in} + R_s$ . In a more exact model,  $R_s$  is frequency and bias dependent. In this test we use  $R_{s2}$  for the series resistance at the input frequency (fundamental mixer) and  $R_{s0}$  for the series resistance at the IF. Index 0 indicates parameters at the IF,  $m$  is the IF.

where

$$Z' = \frac{1}{Y} = \begin{bmatrix} \vdots & \vdots & \vdots \\ \dots & Z_{11} & Z_{12} & Z_{13} & \dots \\ \dots & Z_{21} & Z_{22} & Z_{23} & \dots \\ \dots & Z_{-11} & Z_{-12} & Z_{-13} & \dots \\ \vdots & \vdots & \vdots \end{bmatrix} \quad (33)$$

**Mixer Port Impedances**

The port impedance  $Z_{in,m}$  defined in Fig. 9, can be determined if the corresponding embedding impedance is open-circuited, that is,

$$Z_{in,m} = Z_{in,m} \quad (34)$$

where the subscript  $\infty$  means that  $Z_{in,m}$  is evaluated for  $Z_{in,m} = \infty$ . The IF output impedance becomes

$$Z_{out,0} = Z_{00} - R_{s0} = Z_{00} + R_{s1} \quad (35)$$

**Conversion Loss**

The conversion loss of a mixer is defined as [compare Eq. (8)],

$$L = \frac{\text{Power available from source } Z_{s1}}{\text{Power delivered to load } Z_{L0}} \quad (36)$$

yielding the conversion loss

$$L = \frac{1}{4} \frac{|Z_{s1} + R_{s1}|^2}{|Z_{in,1}|^2} \frac{|Z_{01} + R_{s0}|^2}{|Z_{out,0}|^2} \quad (37)$$

where  $Z_{s1}$  is defined in Fig. 9 (see figure caption) and  $Z_{01}$  is the 01 diagonal element of Eq. (33). A more general expression for the conversion loss from any sideband  $j$  to any other sideband  $i$  is

$$L_{ij} = \frac{1}{4} \frac{|Z_{s1} + R_{s1}|^2 |Z_{ij} + R_{s1}|^2}{|Z_{in,j}|^2 |Z_{out,i}|^2} \quad (38)$$

**Equivalent Noise Temperature of the Mixer**

The equivalent input noise temperature of the mixer,  $T_M$ , is defined as the temperature that the real part of the  $Z_{in,1}$  (lower sideband) must have in order to generate the same noise voltage

as the diode itself generates during mixer operation. It is

$$\mathcal{E}_N = \frac{(i_s V_{N0}^2)}{4k\Delta f} \frac{|Z_{s1} - R_{s1}|^2}{|Z_{s1}|^2 |R_{s1}| |Z_{s1}|} \quad (43)$$

where  $(i_s V^2 / v_0 N_0)$  is

$$(i_s V_{N0}^2) = Z_{01} (\Delta I_m^2 \Delta I_m^2 / Z_{01} - Z_{01}^2 \Delta I_m^2 \Delta I_m^2 / Z_{01}) \quad (44)$$

where

$$\Delta I_m^2 \Delta I_m^2 = \frac{4kT_m R_{sm} \Delta f}{|Z_{sm} + R_{sm}|^2} \quad m \neq 0 \quad (45)$$

$$\frac{4kT_m R_{sm} \Delta f}{|Z_{sm}|^2} \quad m = 0$$

and

$$\Delta I_m^2 \Delta I_m^2 = 2e I_m^2 N_m^2 \quad (46)$$

$I_{m,n}$  is the current component at the harmonic  $(m - n)$  and  $Z_{0n}$  is the zero row of the matrix  $Z'$ . Eq. (31)  $Z_{01}$  is defined in Fig. 9.

The noise temperature as defined in Eq. (42) is the *single sideband noise temperature* for the lower sideband.

Knowing the diode parameters, the Schottky mixer can be analyzed with high accuracy using commercial software.

## DIODE MIXER TOPOLOGIES AND MIXER DESIGN

There are a number of different approaches one may consider in designing mixers. Symmetry properties are one way of defining basic types of mixers, namely,

1. The single-ended mixer with one diode and a common input port for the signal and the LO.
2. The single-balanced mixer with two diodes and separate ports for signal and LO.
3. The double-balanced mixer with four diodes and separate ports for signal and LO.
4. The double-double-balanced mixer with eight diodes and separate ports for signal and LO.

For millimeter and submillimeter wave applications, types 1 and 2 have been implemented experimentally and are described in the literature, while all four types are common in microwave frequency applications.

The performance of either mixer depends on the impedances seen at the signal, at the IF, and at the harmonic sidebands. The LO should experience a reasonably good match in order to reduce the LO power requirement. In practice, the impedances at the harmonics of the LO and at the harmonic sidebands are very difficult to control. The exception is the impedance at the image frequency, which can often be controlled. The importance is illustrated by referring to properties of a typical broadband mixer (when  $L_s = L_i$ ) used in a single sideband application. In such a mixer a considerable amount of the signal power ( $\sim 25\%$ ) may end up at the image frequency.

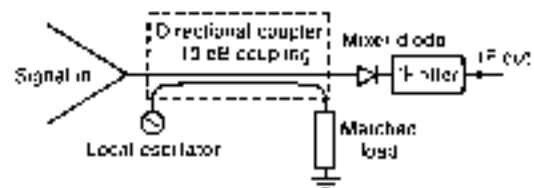


Figure 10. Example for a single-ended mixer configuration. Note that the directional coupler used for injecting the LO couples only 10% of the LO power to the mixer diode, and that the coupler consequently attenuates the signal by a factor of 0.9.

It is also obvious that the noise entering the mixer at the image port will be converted with the same efficiency as the signal at the signal port, adding to the system noise.

### Single-Ended Mixer

A single-ended mixer has one input port, used for both the signal and the LO. Hence it is necessary to incorporate a circuit in front of the mixer itself for injecting the LO. This circuit should not significantly attenuate the signal. For example, using a 10 dB directional coupler will attenuate the signal  $\sim 10\%$  and add noise (see Fig. 10). A common way of introducing the LO in microwave and millimeter wave mixers is to use a narrow band diplexer, for example, a filter structure in the input waveguide or a quasi-optical interferometer in the signal path in front of the mixer (e.g., see Ref. 10).

### Single-Balanced Mixers

In single-balanced mixers, the signal and the LO enter the mixer through different ports, isolated from each other. Either 90° or 180° hybrids or baluns are used (Fig. 11a); (see Ref. 1 for details). Figure 11(b) shows a low-frequency equivalent circuit

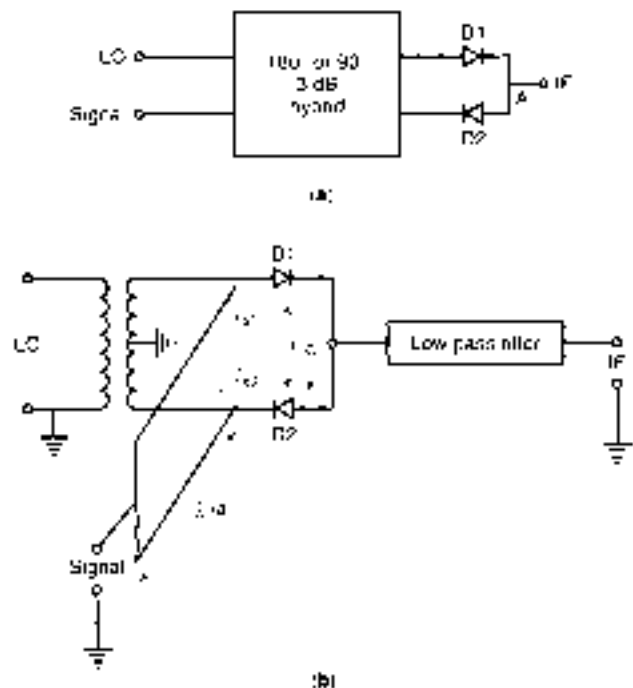


Figure 11. Basic design of the single-balanced mixer (a) and equivalent circuit of the 180° hybrid mixer (b).



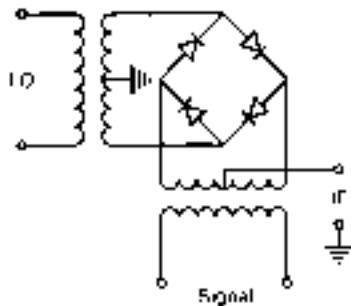


Figure 12. A double balanced mixer configuration—the ring mixer.

of a single balanced mixer. The paths of the signal current  $i_s$  and the LO current  $i_{LO}$  indicate that they add in one diode and subtract in the other. This causes an imbalance in  $A$ , which will slowly cycle at a frequency equal to the IF. Hence the IF power can be subtracted between  $A$  and ground. Note that if the LO is noisy, this will not cause any output noise at the IF port.

The use of two diodes rather than one means that the mixer can handle twice as much power for the same intermodulation as for the single ended mixer. In summary, for the single balanced mixer:

- The signal, LO, and IF ports are isolated from each other.
- The LO noise cancels at the IF port.
- The power handling is superior to the single ended mixer.

**Double Balanced Mixers**

Essentially, a double balanced mixer is constructed from two single balanced mixers, coupled in parallel and 180° out of phase. The diodes can be arranged in either a star or ring configuration (see Fig. 12). The ring can be arranged very compact as a monolithic circuit.

If the diodes are perfectly identical, the symmetry ensures perfect isolation between the signal and the LO ports. The topology also yields cancellation at the IF port of the even harmonics of both the signal and the LO frequencies. This also means that intermodulation is reduced as compared to the mixers mentioned above. Hence the advantages of the double balanced mixer are:

- Excellent isolation between the signal, LO, and IF ports.
- LO noise cancellation at the IF port.
- Superior power handling compared to the double balanced mixer.
- Superior intermodulation properties compared to the double balanced mixer.

**Double-Double Balanced Mixers**

Double-double balanced mixers are constructed using two double balanced mixers. Eight diodes are used, leading to further power handling capacity and still better intermodulation properties.

**Image Rejection and Image Enhancement**

The system properties of a single sideband mixer receiver can be improved by introducing a proper circuit at the image fre-

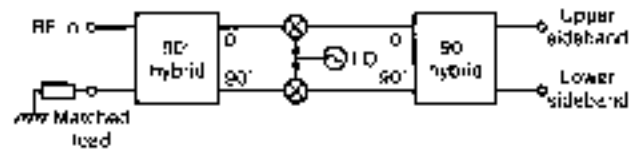


Figure 13. A single sideband mixer using two balanced mixers and two 90° hybrids. Note that at one output port the upper sideband appears, while at the other output port the lower sideband appears.

quency. It is of particular importance to reactively terminate the image frequency so that no signal power is lost at the image frequency, and no noise (or any other unwanted signal) at the image frequency can be converted to the IF frequency. Furthermore, if the reactance at the image frequency is chosen properly, the signal will be reflected back into the mixer, such that the conversion is enhanced and/or the noise properties are improved. A stopband filter can be added in the input transmission line to prevent one sideband from reaching the diode. The distance to the diode is chosen to optimize the mixer conversion loss. The conversion loss becomes several decibels lower than the typical 5 dB for a common broadband microwave mixer.

A most elegant method to realize a single sideband mixer is shown in Fig. 13. By using two balanced mixers and two 90° hybrids, it is possible to arrange that the upper sideband and the lower sideband exit the mixer at different ports (see Ref. 1 or 4 for details).

**HARMONIC MIXERS**

In a harmonic mixer a harmonic of the LO frequency,  $n f_{LO}$ , is used for mixing; that is, the IF is obtained as

$$f_{IF} = |m f_{RF} - n f_{LO}| \quad (13)$$

Harmonic mixers are practical when it is difficult to realize LO power at a frequency near the signal frequency. They are particularly useful at millimeter and submillimeter waves.

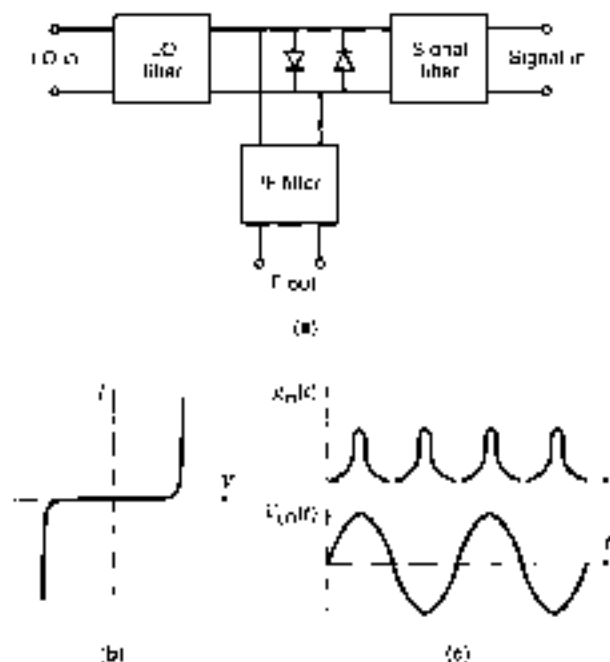
Large harmonic numbers  $n$  are often used when maximum sensitivity is not required. For example, in spectrum analyzers large harmonic numbers may be used.

**Two Diode Subharmonically Pumped Mixers**

If two diodes are used in an antiparallel configuration (see Fig. 14), the small signal conductance will vary with twice the LO frequency. Hence the mixer will convert signals located near  $2 f_{LO}$  and no conversion will occur near  $f_{LO}$ . The advantage of the two diode to single diode subharmonically pumped mixers is that no conversion can occur at the fundamental frequency. Moreover, the LO noise will contribute less, since the frequency difference between the signal and the LO is of the order  $f_{LO}$ . Another advantage is inherent self-protection against large peak reverse voltage burnout.

**PARAMETRIC FREQUENCY CONVERSION**

In a parametric frequency converter, a nonlinear reactance, such as a bias biased Schottky diode, is used. Common parametric components are frequency downconverters, frequency



**Figure 14.** Subharmonically pumped mixer using antiparallel diodes: (a) the mixer circuit, (b) the dc  $I$ - $V$  characteristic, and (c) the resulting waveforms for the LO voltage and the time-dependent small-signal conductance.

upconverters, and frequency multipliers. In a frequency down-converter, a strong pump signal  $f_{LO}$  and signal  $f_s$  (strong or weak depending on application) are applied to the device. The output frequency is

$$f_{out} = |f_{LO} - f_s| \quad \text{frequency downconversion} \quad (44)$$

For a frequency upconversion, we have

$$f_{out} = |f_{LO} + f_s| \quad \text{frequency upconversion} \quad (45)$$

The signal frequency in this case may be much lower than the LO (or pump) frequency, that is, the output frequency is not far from the LO frequency.

Note that if the device has no resistive parasitics, no power is lost in the device itself, and 100% efficiency is theoretically possible. However, power may go to harmonics or harmonic sidebands and there are always some parasitic resistances present, for example, the series resistance in a Schottky diode.

If we select a large ratio  $f_{LO}/f_s$ , the frequency upconverter may have high gain. This is possible since, by proper choice of the circuit parameters, parametric amplification is achieved (e.g., see Ref. 7). In the case of gain, one has negative resistance in the circuit and one may face stability problems. However, it is very difficult to make a broadband parametric upconverter, since proper impedances have to be realized at  $f_s$ ,  $f_{LO}$ , and  $f_{out}$ . Parametric converters are used much less today than a few decades ago. The reason is that the quality of mixer diodes and FETs has improved significantly and it is much easier to make diode (or FET) mixers very broadband. This is the reason why resistive mixers are preferred in most applications. Note

that FET mixers can be designed for a conversion gain greater than one.

A classical reference concerning varactor circuits is the book by Penfield and Rafuse (7). A parametric downconverter, like the Schottky mixer, can be analyzed using commercial software.

### NEGATIVE RESISTANCE DIODE MIXERS

In the current-voltage ( $I$ - $V$ ) characteristic of, say, the Esaki tunnel diode or the resonant tunneling diode, there is a region that has a differential negative resistance. This means that the mixer can have conversion gain. Tunnel diode mixers have been built and tested. However, a large junction capacitance made the frequency range quite limited, which together with poor power handling, stability problems, and less favorable noise properties means that these mixers have very little practical use today.

### SELF-OSCILLATING MIXERS

The negative resistance devices can as well promote an oscillation. Hence it is possible to design circuits where the LO is delivered by the same device that is performing the mixing. Besides the devices mentioned already Gunn diodes have also been used in self-oscillating mixers. The sensitivity of such mixers is limited. The advantage may be in applications where the best performance is less important and the lowest price is required, as, for instance, in low price Doppler radar applications.

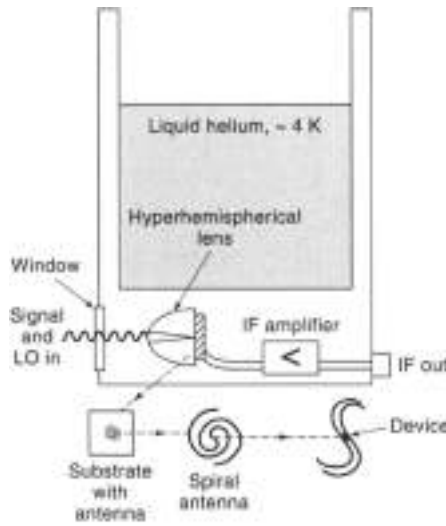
### BOLOMETER MIXERS

Bolometer mixers have been constructed since the 1950s. Since the electromagnetic absorption in bolometer devices can be essentially frequency independent, it should be possible to do mixing to several terahertz. In this type of mixer, one is using the fact that when two signals at slightly different frequencies are superimposed the resulting signal can be described as a signal which is amplitude modulated with the difference frequency. The first useful bolometer mixer was based on InSb devices cooled to temperatures of a few kelvin. When the device absorbs the modulated signal, the electron temperature becomes modulated, leading to a modulation in the device resistance. The theory is described in more detail in the section entitled "The Hot Electron Bolometer Mixer" and by Arps et al. (8).

However, the thermal time constant for the InSb device is long, allowing a maximum IF of only about 2 MHz. A more recent bolometer mixer is based on a two-dimensional electron gas in HEMT materials allowing an IF to about 1 GHz (see Ref. 9). However, the most successful hot electron bolometer mixer so far is the superconducting hot electron bolometer mixer.

### MIXERS BASED ON SUPERCONDUCTING DEVICES

Room temperature mixers for frequencies from about 100 GHz to a few THz frequencies use only Schottky diodes. However, if sensitivity is an issue (e.g., as in radio astronomy), there is a better alternative in mixers based on superconducting devices cooled to a few kelvin. Low noise superconductor-insulator-superconductor (SIS) mixers have excellent performance up to



**Figure 15.** Schematic description of typical dewar setup for a submillimeter wave SIS or HEB mixer. The LO and the signal are entering together through the dewar window. The radiation is focused on the antenna using a hyperhemispherical lens, downconverted to the IF in the nonlinear device and finally amplified by the IF amplifier. In this figure the antenna is illustrated as a spiral antenna. There are many other possible planar antenna structures available (see Ref. 9 for details).

about 1 THz (see Ref. 10). Superconductor hot electron bolometer (HEB) mixers are the best alternative for frequencies above 1 THz.

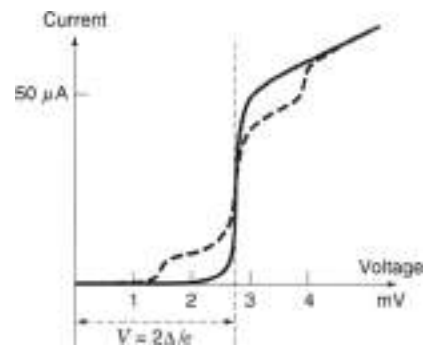
In Fig. 15 is shown a schematic of a receiver (except for the the input quasi-optical) based on superconducting devices.

**The SIS Mixer**

The SIS mixer is also called the "quasiparticle mixer." Due to the extremely strong nonlinearity in the  $I-V$  characteristic of the SIS device, quantum effects are important. Indeed, for certain choices of embedding impedance network, this can result in a conversion gain (see Ref. 11 for details).

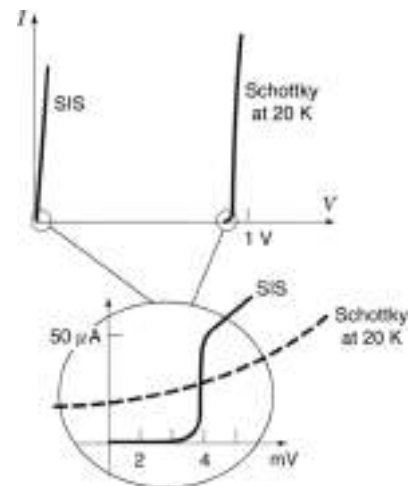
In a superconductor and at a temperature below the superconducting transition temperature, electrons form pairs, called Cooper pairs. When they do so, the energy of the electrons near the Fermi energy is lowered by a certain amount  $\Delta$ . Hence to break up a Cooper pair, an energy of  $2\Delta$  is required. This can be described in terms of a bandgap with the energy  $2\Delta$ , as shown in Fig. 16.

In the SIS mixer, a tunneling phenomenon is used. In the device, two superconducting films are separated by a thin ( $\sim 20$  Å) layer of insulator. Under bias, Cooper pairs on one side of the insulator break up into two electrons (quasiparticles) that



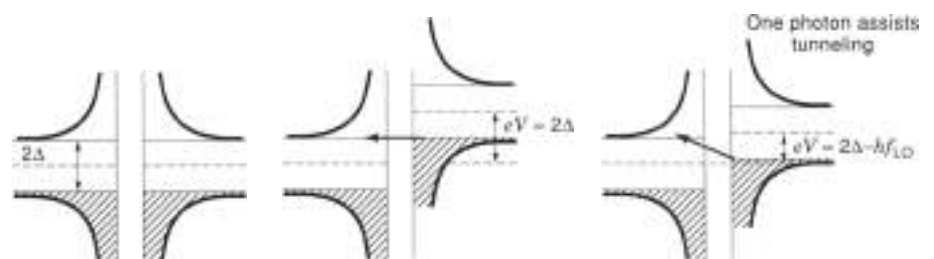
**Figure 17.** Typical  $I-V$  characteristic of a SIS element. The thin line indicates the shape of the pumped  $I-V$  characteristic, where  $f_{LO} = 1.4$  meV ( $f_{LO} = 325$  GHz), while the thick line is for the unpumped device.

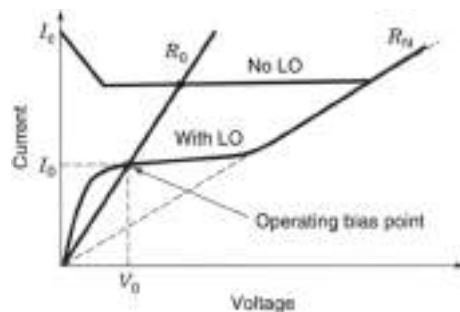
individually tunnel through the insulator and recombine on the other side of the insulator. This is illustrated in Fig. 16. It is interesting to note that the density of states near the band edges becomes "infinite." This is one important reason why there is such a sharp increase in the current when the device is biased to a voltage  $V = 2\Delta/e$  ( $e$  is the charge of the electron). The  $I-V$  characteristic is shown in Fig. 17. Note that the voltage scale is in mV, and that 1 meV corresponds to 240 GHz. The steps in the  $I-V$  curve correspond to dc bias voltages, where exactly  $eV = 2\Delta - hf_{LO} \approx 1.35$  mV. Since the  $I-V$  characteristic is strongly nonlinear within a fraction of a millivolt, the mixer is operating in the quantum regime. Compare also Fig. 18, where the  $I-V$  character of an SIS device is compared with a Schottky diode.



**Figure 18.** Comparing the  $I-V$  characteristic of a SIS element with that of a Schottky diode. Note the enormous difference in nonlinearity.

**Figure 16.** The SIS device under bias: (a) no bias, (b) for a voltage bias  $V = 2\Delta/e$  electrons will tunnel from right to the left; (c) tunneling is assisted by a photon with energy  $hf_{LO}$  for the bias voltage  $V = (2\Delta - hf_{LO})/e$ .





**Figure 19.** Current-voltage ( $I$ - $V$ ) characteristic of a HEB with and without a LO. At the operating point,  $V_0/I_0 = R_0$ . For large bias voltage the whole strip is normal conducting ( $R = R_N$ ).  $I_c$  is the critical current, the maximum current in the completely superconducting state.

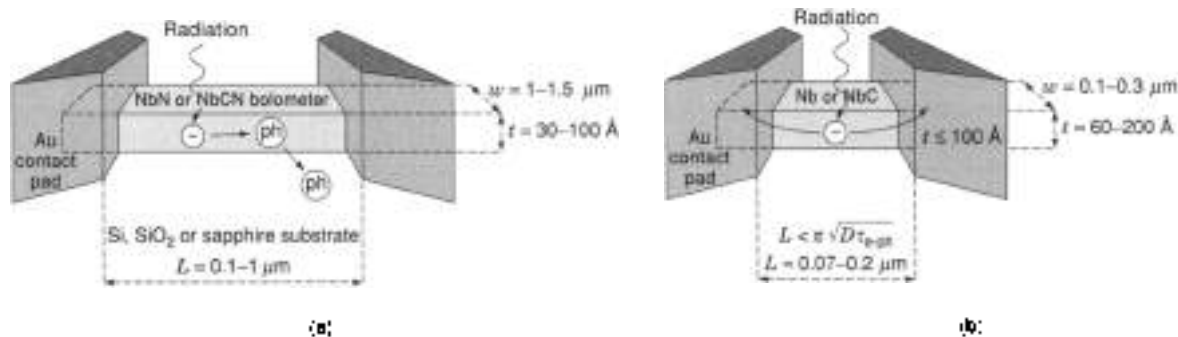
The Schottky diode obviously is not very nonlinear within a voltage interval of  $\approx \pi kT$  and is therefore operating fully as a classical mixer for frequencies up to several terahertz. It should also be mentioned that the required LO power is very low, on the order nanowatts. This is of great importance for submillimeter wave mixers, where substantial LO power is difficult to obtain.

The best SIS devices are realized in so-called Nb trilayer technology (see Ref. 11). The device structure is  $\text{Nb}/\text{Al}_2\text{O}_3/\text{Nb}$ , where the  $\approx 20$  Å thick  $\text{Al}_2\text{O}_3$  serves as the insulator in the SIS device. For frequencies above about 700 GHz, one is trying to develop devices based on  $\text{NbN}$ , which has a higher bandgap ( $\approx 1.2$  THz) than Nb. So far these attempts have not been very successful.

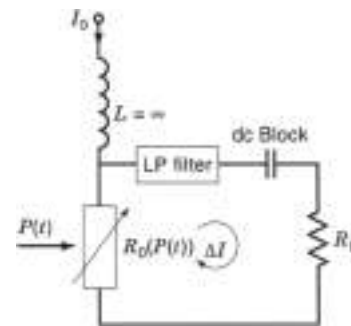
The basic noise in the SIS mixer is shot noise. Comparing the  $I$ - $V$  characteristics of the SIS device and Schottky diode and using the classical theory described in the section entitled "Schottky Diode Mixers," one can see that the mixer noise and the conversion loss are essentially lower for the SIS mixer. However, the theory must include quantum effects (see Ref. 11 for more details). This leads to a conversion gain that is possibly larger than one, a fact that has been demonstrated in practice.

### Hot Electron Bolometer Mixer

A bolometer consists of an absorber that is heated by radiation and a temperature-dependent resistance as a "thermometer." The bolometer has a thermal time constant  $\tau_0$  limiting the maximum detectable modulation frequency of the absorbed power. This means that the maximum feasible IF is  $f_{\text{IF}} = 1/(2\tau_0)$ .



**Figure 20.** Two types of bolometer devices: (a) phonon-cooled and (b) diffusion-cooled.



**Figure 21.** Equivalent circuit of bolometer with load.

When the LO and signal are added together, the instantaneous power variation is described by  $[V_{\text{IF}} \cos(\omega_{\text{LO}}t) + v_s \cos(\omega_s t)]^2 / (1/R_{\text{HP}})$  (see Fig. 3). If the bolometer can respond to  $\omega_{\text{IF}}$  but not to  $(\omega_{\text{LO}} \pm \omega_s)$ ,  $2\omega_{\text{LO}}$ , and  $2\omega_s$ , the bolometer temperature and the resistance will be approximately proportional to  $P_{\text{LO}} + P_s - 2\sqrt{P_{\text{LO}}P_s} \cos(\omega_{\text{IF}}t)$ . Note that the "slow" response of the bolometer device means that there are no harmonics of the LO created and no signal power is transformed to the image frequency, as is the case for both the SIS and Schottky mixers.

A *superconducting hot electron bolometer* (HEB) consists of one or several superconducting thin film strips in parallel, deposited on a substrate, for example, silicon, single crystalline quartz, or sapphire. The strips are cooled to the superconducting state and then heated by dc and microwave power to temperatures near the superconducting to normal transition temperature, where the superconductor will gradually become normal (Fig. 19).

The maximum IF is determined by the electron temperature relaxation time  $\tau_0$ , that is,  $f_{\text{IF}} < 1/(2\pi\tau_0)$ , and a major issue is to find ways of making the time constant  $\tau_0$  short enough. Figure 20 indicates how cooling occurs in the so-called phonon-cooled and the diffusion-cooled bolometer, respectively (compare Refs. 10 and 12).

When operating the mixer, the device is absorbing LO power ( $P_{\text{LO}}$ ) and signal power ( $P_s$ ) as well as power from the dc bias supply ( $P_{\text{dc}} = V_{\text{IF}}I_{\text{dc}}$ ). When the power increases, obviously the electron temperature increases and the resistance of the device increases as  $\Delta R = (dR/dT)\Delta T = C_0\Delta P$ . Figure 21 shows a simple equivalent circuit of the mixer, where the device is biased by a constant dc current. Consequently, the modulation at the IF of the resistance will cause an IF voltage to appear across the device, causing an IF current through the IF load

resistance  $R_L$ . The IF current  $\Delta I$  is superimposed on the dc bias current through the mixer device  $R_D(P)$  and will cause a "modulation" of the dc power  $\Delta P_{dc} \cos(\omega_{IF}t)$ . The total power dissipated in the device is then

$$\begin{aligned} P_{dt} &= P_D + \Delta P_{dc} \\ &= P_{dc} + \Delta P_{dc} \cos(\omega_{IF}t) + P_{LO} + P_s + 2\sqrt{P_{IF}P_s} \cos(\omega_{IF}t) \end{aligned} \quad (46)$$

Assuming that dc and RF power affects the resistance by the same amount, the IF modulation of the device resistance becomes  $C_0 \Delta P(t)$ . There is a resulting bias point of the device  $V = V_0$  and  $I = I_0$ . Defining the device dc resistance  $R_0$  as the time average of  $R_D(P)$ , that is,  $R_0 = V_0/I_0$ , one obtains (8) the conversion gain,

$$G = \frac{P_{IF}}{P_s} = 2C_0^2 \frac{P_{LO}P_s}{R_0 + R_L} \cdot \frac{R_0}{R_L} \left( 1 - C_0 \frac{P_s}{R_0} \cdot \frac{R_0 - R_L}{R_0 - R_0} \right)^2 \quad (47)$$

where  $C_0 = dR_D/dP$ ,  $R_0$  is the IF load resistance.  $P_{IF}$ ,  $P_s$ ,  $P_{LO}$ , and  $P_{dc}$  are the IF, signal, LO, and dc power, respectively, dissipated in the device.

The commonly assumed fundamental limit of  $\sim 6$  dB gain for hot electron mixers is not valid if a negative differential resistance of the unpumped  $I$ - $V$  curve is available (see Ref. 13 for detail). The load resistance for maximum gain is equal to the differential resistance of the  $I$ - $V$  curve at the bias point of the pumped mixer.

## Experimental Results

In experiments on phonon-cooled NbN HEB mixers, noise temperatures of about 400 K (DSB) have been obtained at 800 GHz, and 1000 K (DSB) at 900 GHz. At 2.5 THz a noise temperature of about 1400 K has been obtained. The conversion loss is typically 10 dB, including losses from the optics in front of the mixer.

The noise of these mixers is caused by thermal fluctuations in the balometer device (causing resistance fluctuations) and by Nyquist noise. For diffusion-cooled mixers a noise temperature of 650 K (DSB) at 533 GHz was measured by Skalare et al. (14). These experiments indicate that an IF bandwidth of at least 3 GHz is achievable. Later experiments show that it should be possible to obtain at least 50% higher IF bandwidths in practical mixers. A crucial number is the maximum IF bandwidth. For a phonon-cooled HEB mixer, a 1.7 GHz IF bandwidth (1.3 dB reduction in conversion gain) has been measured and for diffusion-cooled between 2 and 6 GHz. While this bandwidth is defined for a conversion loss increase of 3 dB, the bandwidth defined for when the noise temperature has increased by 3 dB is about 1.5 times larger ( $\sim 5.5$  GHz for the phonon-cooled one).

The LO power needed is less than 100 nW, which is much lower than needed for Schottky diode mixers.

## BIBLIOGRAPHY

- 1 S. A. Maas, *The RF and Microwave Circuit Design Cookbook*, Norwood, MA: Artech House, 1996.
- 2 E. I. Kolberg, Mixers and detectors, in K. Chang (ed.), *Handbook of Microwave and Optical Components*, New York: Wiley, 1990.
- 3 K. E. Yngvesson, *Microwave Semiconductor Devices*, Norwood, MA: Kluwer, 1991.
- 4 S. A. Maas, *Microwave Mixers*, 2nd ed., Norwood, MA: Artech House, 1990.
- 5 S. A. Maas, *Nonlinear Microwave Circuits*, Piscataway, NJ: IEEE Press, 1997.
- 6 D. N. Held, Conversion loss and noise of microwave and millimeter-wave mixers, Part 1—Theory, and Part 2—Experiment, *IEEE Trans. Microw. Theory Tech.*, **26**: 49–61, 1978.
- 7 P. Penfield and E. P. Kalousek, *Varactor Applications*, Cambridge, MA: MIT Press, 1962.
- 8 F. Arams et al., Millimeter mixing and detection in bulk InSb, *Proc. IEEE*, **54**: 308–316, 1966.
- 9 J.-X. Yang et al., Wide-bandwidth electron bolometric mixers: A 2DEG prototype and potential for low noise THz receivers, *IEEE Trans. Microw. Theory Tech.*, **41**: 581–589, 1993.
- 10 R. Blandell and C. E. Tong, Submillimeter receivers for radio astronomy, *Proc. IEEE*, **80**: 1702–1720, 1992.
- 11 M. J. Wengler, Submillimeter wave detection with superconducting tunnel diodes, *Proc. IEEE*, **80**: 1810–1826, 1992.
- 12 J. E. Carlstrom and J. Zmudynas, Millimeter and submillimeter techniques, in W. R. Stone (ed.), *Reviews of Radio Science 1995*, 1995, London: Oxford Univ. Press, 1996.
- 13 H. Ekstrom et al., Conversion gain and noise of ambient superconducting hot-electron mixers, *IEEE Trans. Microw. Theory Tech.*, **43**: 839–847, 1995.
- 14 A. Skalare et al., A heterodyne receiver at 533 GHz using a diffusion-cooled superconducting hot electron mixer, *IEEE Trans. Appl. Supercond.*, **5**: 2236–2240, 1995.

ILICHO ANGELOV  
 GERRI KOJUREVIC  
 HEDDERT ZHANG  
 Chalmers University of Technology

## FUSION REACTORS

### NUCLEAR FUSION

The fission giants which now provide about 20% of the world's electricity accomplish this feat through neutron-driven chain reactions in which heavy atomic nuclei, such as  $U^{235}$ , split apart into lighter nuclei, releasing large amounts of energy (on the order of 180 million eV) in the process. Beginning about 1951, when one of the first fusion programs was begun at Princeton University by Lyman Spitzer, Jr., many of the world's developed nations, as well as some developing ones, have pursued research to eventually produce a fundamentally different type of nuclear reactor. Since 1958, when research on peaceful uses of nuclear energy was declassified, these many countries have shared their research and openly collaborated, even during decades when some of the principal contributing nations were political adversaries. This new type of reactor is a fusion reactor, and the very great progress which has been made in the decades since 1951 stands as a landmark in what may be the most universal and long-lived collaborative effort in human history.

Fusion reactors, like fission reactors, will use exothermic nuclear reactions to release energy. The fusion reactions themselves, however, are very different in character than are those in fission reactors, and they require entirely different conditions in order to proceed at an acceptably high rate for power plant applications. Due to these differences, fusion reactor designs will look very different from those of fission reactors. Moreover, many of the problems that have had to be surmounted in the pursuit of fusion power are of a fundamentally different nature than those encountered in making fission reactors practical.

Nuclear fusion is a reaction in which two atomic nuclei merge to form a heavier element. If the reaction is an exothermic one, then the fusion process will result in the release of energy. This energy is carried as kinetic energy by the reaction products, consisting of the product nucleus (which is positively charged) and another particle, such as a neutron (which is electrically neutral) or a proton (which is positively charged). Because all nuclei carry positive electric charge, they repel each other, with the result that fusion reactions can proceed at significant rates only at very high temperatures which give the nuclei sufficient energy to overcome their repulsion and approach each other close enough to merge. Alternatively, this same end may be achievable by using somewhat lower temperatures, but at high densities achieved by compressing the fusion fuel with high symmetrically applied pressures.

The electrostatic repulsive force between two nuclei is proportional to the product of the positive charges (and thus to the atomic numbers) of the two reactant nuclei. This gives rise to a potential barrier, referred to as a Coulomb barrier, which the approaching nuclei must overcome in order to merge. Due to quantum mechanical tunneling, some pairs of nuclei can fuse even when their combined kinetic energy is less than that required to exceed the Coulomb barrier. However, the likelihood of this tunneling occurring declines very rapidly as the kinetic energy of the reactants falls further below the Coulomb barrier height. For most nuclear reactants, the reaction rate drops to inconsequential levels if the approaching nuclei do not have kinetic energies of at least 70% to 80% of the barrier height. For heavier nuclei, the height of the Coulomb barrier becomes several to many millions of electron volts (MeV), where an electron volt is a unit of energy corresponding to the energy an electron gains when it falls through a potential difference of 1 V. A temperature of 1000 eV (1 keV) corresponds, in more conventional units, to about 10,600,000 C.

Thus, in order to fuse most of the elements in the periodic table, one would need to cause them to approach each other at energies of several to many MeV. This can be accomplished with particle accelerators for nuclear physics research purposes, but is not practical for producing net power with significant quantities of thermal reactants. As a consequence, only the lightest elements in the periodic table have sufficiently large probabilities (also called reaction cross sections) of undergoing nuclear fusion for them to be considered as fuels for fusion reactors.

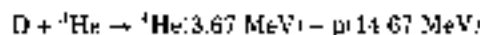
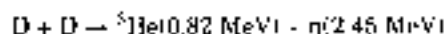
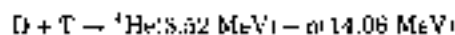
Fusion reactions drive the core of the sun at a temperature thought to be roughly 10 million to 11 million C. Fusion also powers all the other stars, and thus supplies most of the universe's observable energy. Our sun, in common with other main sequence stars, obtains its energy through a number of nuclear fusion reactions, beginning with the fusion of 2 protons into deuterium. The deuteron is accompanied by a positron and a neutrino, and together they carry 1.44 MeV of energy. The reaction probabilities (or cross sections) for this and the succeeding

solar nuclear reactions are much too small to be of any use to a commercial fusion reactor on the earth's surface. The sun is able to produce enormous amounts of energy with these reactions only because it is very large compared to the earth. The sun has a diameter of about 897,000 mi and, because of its size and its gravity, has excellent energy confinement. The time for energy produced in the sun's fusion driven core to reach the sun's surface is estimated to be of the order of thousands of years or more.

Because a practical fusion reactor for electricity production needs to be much smaller than the sun, it will have to rely upon different nuclear fusion reactions with larger cross sections, and it will also need to operate at temperatures that are at least 10 to 20 times those of the sun's core. Temperatures in this range were achieved within a mirror confinement device in 1979 and within the Tokamak Fusion Test Reactor (TFTR) at Princeton University in the mid-1980s, and have now also been obtained on the Joint European Torus (JET) tokamak in England, and the JT-60U tokamak of Japan. Ion temperatures as high as 300 million to 500 million C (roughly 30 to 50 times the temperature of the sun's core) were routinely produced in Princeton University's TFTR over a period of years.

## FUSION REACTOR FUELS

There are a number of nuclear fusion reactions which have cross sections sufficiently large to be potential candidates as fuels for commercial fusion reactors. These are listed below, where D represents a deuteron, the nucleus of  $^2\text{H}$ , T stands for a triton, the nucleus of  $^3\text{H}$ , and the energies in parentheses are the amounts of kinetic energy each of the reaction products carries away from a fusion event. An n stands for a neutron, and a p for a proton, the two types of baryons which occur in atomic nuclei.



Of these reactions, the D + T reaction has by far the largest cross section at energies of tens of keV, which is the region which should be obtainable in the first generation of commercial fusion reactors. For a fusion reactor with a thermal reacting population, the parameter which describes the fusion reaction rate is  $\langle\sigma v\rangle$ , where  $\sigma$  is the nuclear reaction cross section, and  $v$  is the relative speed of the reacting nuclei. The angle brackets mean that the product of  $\sigma v$  is averaged over the Maxwellian velocity distributions of the reactants. This fusion reactivity parameter is also much larger for the D + T reaction than for the other possible fuel mixtures. Due to the importance of the reactions arising from the high energy tail of the Maxwellian velocity distributions, the optimum range for ignition of a D + T fuel mix lies across a range of about 20 keV to 40 keV, well within the operating range of some of the largest of existing fusion devices. These experimental devices have not actually reached ignition because, although their temperatures were adequate, the product of the density and the energy confinement time was not yet sufficiently large. The energy confinement time is a measure

of how long is required for energy to leak from the plasma. In the 1950s a typical experimental device had an energy confinement time of a few milliseconds for the overall plasma; by the mid-1990s this reached values of as much as 1.4 s.

In addition to the magnitude of the reactivity as a function of temperature, there are also other factors which bear upon the ease with which different fusion fuels can be used in a commercially viable reactor. There are two types of radiation loss which take energy from the confined fuel and transport it directly to the outside, and both of these tend to be more serious for higher temperatures. The first of these is bremsstrahlung, which scales in intensity proportional to the square of the density times the square root of the temperature. For D + T fueled reactors, this would be a relatively tolerable portion of the overall power balance. However, for the other fuels it would play a larger role. For approaches to fusion reactors which employ strong magnetic fields, the electron synchrotron radiation becomes a major factor at the higher temperatures (80 keV to 120 keV) required by reactions involving D + D or D + <sup>3</sup>He. Thus, the appeal of a D + T fuel mix in at least the early electricity-producing fusion reactors arises from two major factors: the fusion reactivity is much higher at temperatures which can be reached by some of today's fusion devices, and the radiation loss rates from bremsstrahlung and synchrotron radiation are lower for likely reactor conditions. Synchrotron radiation would be less important for alternate plasma confinement schemes which do not use strong magnetic fields, but these approaches have not been developed as far as the strong field path to fusion.

### Deuterium Fuel

Deuterium is a stable, naturally occurring isotope of hydrogen. On the earth, one out of every 6700 atoms of hydrogen is deuterium. Thus, enormous supplies of deuterium are available from the earth's water. If later generations of reactors operate with the D + D reaction, then there is enough fuel to supply the world's energy needs for hundreds of millions to billions of years, depending upon assumptions about the future growth of energy usage. The deuterium can be concentrated and extracted from any water by utilizing such enrichment techniques as diffusion through a series of filters or by electrolysis.

The earth's water contains about  $4 \times 10^{22}$  kg of deuterium. If this were used to fuel D + T fusion reactors with an overall operating efficiency of 33%, then this would allow the production of about  $10^{24}$  GJ of electricity, which is close to  $3 \times 10^{11}$  the present annual electricity production of the entire world. As we will see later in this section, the real limit on the amount of energy potentially available from D + T fusion reactors is the supply of feedstock to produce the tritium used in the reaction.

If later reactors used D - D as their primary reaction, then, because this reaction is less exothermic than the D - T reaction, and because two deuterium atoms, instead of one, would be required for each reaction, then the estimate of total electricity production if all the deuterium in the waters of the world were used would drop to about  $10^{21}$  GJ, which is still a very large number, and is equivalent to about  $3 \times 10^{10}$  times the world's current annual production of electricity. The cost of deuterium is of the rough order of one dollar per gram (1), with a gram of deuterium being sufficient to produce 300 GJ of electricity if fused with tritium. The cost of the deuterium fuel for a D + T reactor is thus about \$0.003 per gigajoule of electricity. If the deuterium were instead used in a D - D reactor, the cost would

rise to about \$0.02 per gigajoule of electricity, which is still small compared to the cost of bulk electricity, which runs in the vicinity of \$20 per gigajoule. The fuel costs would be a negligible portion of the cost of electricity from a fusion plant, which would be dominated by capital costs and maintenance. It is likely that the price of electricity from a fusion reactor would be similar to or perhaps somewhat more expensive than electricity from a fission reactor, at least in the near term, while there is still adequate fuel for the simple once-through fuel cycles used in most fission power plants.

### Comparison to Fossil Fuel Energy Densities

Producing a gigawatt (10<sup>9</sup> W) of electricity for a year in a fusion reactor would require roughly 1000 kg of deuterium. Producing the same amount of electrical energy from a power plant burning coal would require about  $2 \times 10^9$  kg of carbon. The actual weight of coal required would of course be greater than this, since coal contains other elements besides carbon. The fact that the mass of fuel which has to be carried from the fuel concentration source to a fusion power plant is more than a factor of a million smaller than for competing fossil fuel plants is a significant advantage. It means that moving the fuel for an entire fusion economy would impose no requirements upon the transportation infrastructure, since the masses being moved would be thousands of kilograms instead of billions of kilograms.

The fact that fusion reactors will be able to produce a gigawatt-year of electricity while using fuel which weighs more than a million times less than that required for a fossil fuel burning plant reflects the large difference between nuclear binding energies and electron binding energies. When a reaction occurs between two parent particles, the extra kinetic energy carried by the daughter particles comes from the change in the overall binding energy. This kinetic energy is distributed among the daughters in accordance with the laws of energy and momentum conservation, with the result that the lighter daughter carries most of the kinetic energy.

Burning a fossil fuel is a chemical reaction, which involves exchanges and rearrangements of the electrons outside the nuclei. The outer electrons of atoms typically have binding energies of several to a few tens of electron volts. Thus, the changes in the net binding energy that occur in chemical reactions, such as burning fossil fuels, are typically only a few electron volts. Nuclear binding energies, on the other hand, are typically many millions of electron volts. This has the consequence that the changes in binding energy involved in nuclear reactions are also much greater than for chemical reactions. For instance, in the D - T reaction, 17.58 MeV of energy is released as kinetic energy in inverse proportion to the masses of the daughters. Thus, because a neutron weighs about one fourth as much as does an <sup>4</sup>He nucleus, the neutron carries four times as much of the kinetic energy.

The disparity in binding energies between nuclear and chemical reactions arises in turn from the fact that the strong nuclear force is much more powerful across nuclear dimensions than is the coulomb force (which binds the negatively charged electrons to the positively charged nucleus) across dimensions typical of atoms. Consequently, the huge difference in the magnitudes of potential energy available per unit of mass available from nuclear, as opposed to chemical reactions, arises from a major difference in the strengths of two of the fundamental forces of nature. Thus, there is no possibility that any chemical

reaction involving changes in electron configurations could ever begin to approach the energy releases available from nuclear reactions.

### Tritium Fuel

Tritium, the heaviest isotope of hydrogen, is unstable. It decays by emitting a beta particle (an electron) with an average energy of 5.7 keV. This beta particle is easily stopped by even a piece of paper, so it does not pose a hazard unless it is ingested. Even then, unless large amounts (millicuries or more) are taken into the body, it is not very likely to produce ill effects. This is the result of two factors. The first is that the ionizing radiation released into the body by each tritium decay is much less than the decay energies of fission products (which are typically at least hundreds of keV) or the energies in the decay chains of heavy elements, which can run to over 10 MeV. The second factor is that most of any tritium absorbed by the body would enter as water, and water is continuously excreted, with a biological half-life in human bodies of about 12 to 13 days. The absorption rate for tritium breathed into the lungs as gaseous molecular hydrogen is very low, roughly 20,000 to 25,000 times less than for tritium in water molecules.

Thus, while tritium should always be treated with due care, its possible health effects if mishandled do not appear to be significantly worse than those of many other chemicals routinely handled by an industrialized society. Indeed, tritium is already used in conjunction with phosphors to provide light without the need of electricity in several applications such as school exit lights, some airfield landing lights, and some modern illuminated watches. These applications incorporate significant amounts of tritium. Tritium-powered school exit lights typically use about 15 to 25 curies of tritium, and emergency runway landing lights use much more. This compares with an amount on the order of 100 to 150 curies injected into the Tokamak Fusion Test Reactor at Princeton University for a high-power fusion shot. This device has operated on a routine daily basis using deuterium and tritium for experiments over a period spanning three and a half years without any significant incidents involving tritium contamination of personnel or the environment. This demonstrates that tritium can be handled on a large fusion system without unduly impeding the manner in which it is operated, although the total quantities of tritium being handled were orders of magnitude lower than would be required in a power plant. The neutron-induced activation and damage of materials would be much more challenging in a fusion power plant.

While the beta decay of tritium does not result in either undue hazard potential or in excessive constraints upon the operation of fusion reactors, it does have an inconvenient consequence. The half-life (the time for half of the nuclei in any assemblage to undergo beta decay) is only 12.3 years. Thus, any primordial tritium that fell into the earth during its formation decayed away billions of years ago. Some tritium is continuously produced in the upper atmosphere through nuclear reactions initiated by cosmic rays. However, due to the short half-life of tritium it does not build up, so the equilibrium concentration of tritium in air is very low, and far too small to economically extract as fuel for a fusion fuel.

Consequently, tritium fuel for fusion must be manufactured through nuclear reactions. The tritium used for present-day fusion experiments is made in heavy water nuclear fission

reactors. However, the amount produced by these reactors would be insufficient to fuel a fusion economy. Equally important, the long-term goal of fusion research is to produce fusion reactors which supplant fission reactors as we know them, so fusion reactors must produce their own tritium. Fortunately, this is feasible using reactions such as:

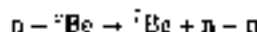


### Tritium Production

The incident neutrons that induce these reactions arise from the D - T reactions, and perhaps also from other reactions initiated by the primary neutrons. The reaction involving  ${}^6\text{Li}$  has the advantage that it is exothermic, adding another 4.5 MeV of kinetic energy to the 17.58 MeV of kinetic energy released by the D - T reaction that produced the neutron. Moreover, because it is exothermic, there is no threshold energy for initiation of the reaction, meaning that even low-energy neutrons which have undergone many elastic and inelastic collisions can still produce tritium in this way.

The reaction involving  ${}^7\text{Li}$  is less advantageous in one sense, in that it is endothermic, consuming 2.5 MeV of kinetic energy in order to occur. This also means that only energetic neutrons (above about 1 MeV (after allowing for center of mass effects) among the reactants) can initiate this reaction. On the other hand, the  $n + {}^6\text{Li}$  has the advantage that it does produce an additional neutron, which may initiate an  $n - {}^6\text{Li}$  reaction to produce still another triton.

Thus, for a fusion reactor using deuterium and tritium fuel, the raw material for the production of tritium is lithium. The natural abundances of these lithium isotopes on earth are 7.4% for  ${}^6\text{Li}$  and 92.6% for  ${}^7\text{Li}$ . In order to produce sufficient tritium to at least continuously replenish its fuel supply, a deuterium tritium fusion reactor would be surrounded by a lithium blanket which would produce tritium and capture the kinetic energy of the fusion neutrons and neutron-induced reaction products. As these particles slow down by collisions with the lithium blanket, their kinetic energy will be converted to heat. This heat will, in turn, be used to produce steam to drive electricity producing turbines. This tritium-breeding blanket may also include other materials such as beryllium. Terrestrial beryllium is 100%  ${}^9\text{Be}$ , which can act as a neutron multiplier primarily through the reaction:



This reaction is modestly endothermic, requiring an input of 1.67 MeV in order to occur, meaning that the threshold energy in the laboratory frame for incident neutrons is around 2 MeV.

If, for instance,  ${}^7\text{Li}$  is used to make tritium (which is a fairly good approximation to using natural lithium), then with a system efficiency of 33% for net electricity production, 1 kg of  ${}^7\text{Li}$  would be sufficient to produce  $7 \times 10^4$  GJ of electricity from a D + T fusion power plant. At the present cost of roughly \$20 per kg (1) for natural lithium, the lithium in a tritium breeding blanket would contribute only about \$0.001 per GJ of electricity. This is negligible compared to the bulk price of electricity of about \$20 per GJ. Thus, the price of breeding lithium could rise many fold before it had a noticeable impact upon electricity



costs. In fact, if it appeared advisable to use an isotopically enriched blanket, it would be possible to do so with only a modest effect upon the price of electricity. The energetics are such that the electricity produced with a  $\text{Li}^6$  blanket might be slightly cheaper than with a natural Li blanket, because the extra 4.8 MeV from the  $n + \text{Li}^6$  reaction is much larger than the amount of energy required to enrich the lithium.

The United States has large reserves of easily extractable lithium, mostly in dry salt lakes and saline lakes, which could be produced at prices not greatly different from the \$20 per kg of today. This reserve is estimated to be about  $5 \times 10^9$  kg (1). With an electricity production efficiency of 33%, this would yield about  $3 \times 10^{15}$  GJ of electricity, an amount that is of the order of 300–1000 times the primary energy consumption of the world in the early 1980s. Inasmuch as the United States comprises only 6% of the earth's land area, it is reasonable to assume that the total world reserves of cheaply extractable lithium might be a few times greater than these United States reserves. Thus, the reserves of easily obtainable lithium are sufficient to run an economy powered by D + T fusion reactors for a period of several to many centuries. There is a much greater amount (of order  $10^{17}$ ) of lithium dissolved in the world's oceans than in land deposits (1). This could be extracted at somewhat greater cost than from the saline lakes and dry salt beds. However, this cost premium might be relatively modest. Dikes could be built across tidal flats to isolate seawater which could then be treated the same as saline lakes and salt flats after it had undergone sufficient solar evaporation.

#### Reduced Activation Fuels

The high-energy neutrons produced by fusion reactors using deuterium or deuterium + tritium as fuels undergo nuclear reactions with the materials forming the structure of the reactor. Some of these reactions result in the production of radioactive nuclei with a variety of half-lives. This activation of structural components, and particularly of those components close to the fusion core, makes maintenance more difficult, and also will require at least short term storage of removed components, or possibly longer term, depending (as will be discussed in a later section) upon the materials used to build the reactor. In addition, the energetic neutrons can introduce lattice defects as they scatter, which in turn reduce the lifetimes of components, requiring more frequent replacement.

In order to reduce these deleterious effects of the neutrons upon fusion reactor structures, studies have been carried out to evaluate the feasibility of alternate fuels which would produce fewer neutrons per unit of fusion energy released. These reactions include the  $\text{D} + {}^3\text{He} \rightarrow {}^4\text{He} + \text{p}$  reaction and the  $\text{p} + {}^{11}\text{B} \rightarrow 3\text{He}$ . As discussed earlier, the first of these reactions requires higher temperatures than are presently obtainable, and is more subject to bremsstrahlung radiation losses, and is much more vulnerable to synchrotron radiation losses in systems with magnetic fields than is a D + T fuel mixture. Due to lower fusion cross sections and the higher atomic number of  ${}^{11}\text{B}$ , these problems are more difficult for a reactor using the  $\text{p} + {}^{11}\text{B}$  reaction.

A reactor using the  $\text{D} + {}^3\text{He} \rightarrow {}^4\text{He} + \text{p}$  reaction would have the advantage over a D + T reactor that both components of the fuel would be stable. It would still produce some 3.61 MeV neutrons from D + D reactions, and it would produce some 14.1 MeV neutrons from the burnup of tritium produced in one of

the two branches of the D + D reaction. However, almost all of the tritium produced would be consumed in D + T fusion reactions, and thus would not leave the reactor. If the temperature and fueling profiles of such a reactor could be optimized appropriately, the neutron production rate could be reduced by a factor of 100 compared to a D + T reactor (2).

Unlike all the other potential fusion fuels, which occur in abundant commercially attractive concentrations on or near the earth's surface,  ${}^3\text{He}$  is extremely rare in the earth's crust. The small supplies that are available arise from the beta decay of tritium. This tritium, in turn, has to be bred through nuclear reactions of one sort or another. At the present time, all substantial production of tritium is carried out through neutron capture reactions on deuterium in heavy water moderated fission reactors. In the future, tritium might be bred through other nuclear reactions either by fusion reactor or high-energy particle accelerators. However, if one must first produce tritium in order to obtain the  ${}^3\text{He}$  fuel for a  ${}^3\text{He} + \text{D}$  fusion reactor, then much of the rationale for using  ${}^3\text{He}$  disappears.

What is needed in order that  ${}^3\text{He} + \text{D}$  fusion reactors have the possibility of commercial viability is a naturally occurring source of  ${}^3\text{He}$  in concentrations and quantities sufficient for economic exploitation. As it happens, while no  ${}^3\text{He}$  deposits occur on earth, they do occur in abundance on the surface of the moon.

#### Lunar ${}^3\text{He}$

Because the moon, unlike the earth, lacks both an atmosphere and a magnetic field, its surface is raked by the solar wind, a flux of energetic particles driven outward from the sun. Among other constituents, this wind carries  ${}^3\text{He}$ , which embeds itself in the lunar surface when it strikes. The moon, unlike the earth, does not have enough internal heat to drive plate tectonic motions of its surface now, nor for a very long time into its past. For similar reasons, much of its surface has not experienced fresh outpourings of lava or ash since the interior cooled. The early moon probably possessed an atmosphere, but this long ago escaped into space because of the moon's weak gravity. Consequently, there has been no weather and no water to erode and rearrange the lunar surface in billions of years. When our solar system was young, the moon, in common with the earth, was subjected to an intense bombardment by large and small fragments of matter left over from the formation of the planets. The ejecta arising from this bombardment deposited new layers of material over much of the moon's surface, burying pre-existing surface layers. However, most of the solar system debris was swept out of the planetary orbits long ago through the direct interception of fragments by the planets and moons, and through deflections of the orbits of fragments arising from close encounters with the gravitational fields of the much larger planets and moons. The evidence of this early bombardment has eroded away on earth due to the effects of weather, water, plate tectonic movements of the land, and volcanoes. On the moon, however, which lost these processes in the distant past, many large craters still stand from the later periods of the bombardment.

As a result of these differences between the histories of the lunar and terrestrial surfaces, much of the lunar surface has remained relatively undisturbed for perhaps billions of years. During all this time the solar wind has continued to deposit  ${}^3\text{He}$  into the lunar surface, so that it now exists there at a

concentration of about 10 parts per billion. A study of the feasibility and economics of collecting and concentrating the  ${}^3\text{He}$  from the dust which covers much of the lunar surface found that, even after including the high transportation costs of carrying equipment to the moon and sending the  ${}^3\text{He}$  back, this would be economically feasible if a commercially viable  ${}^3\text{He} + \text{D}$  reactor could be developed (3). As discussed earlier, however, the parameters required for a  ${}^3\text{He} + \text{D}$  reactor to be viable are significantly more daunting than for a  $\text{D} + \text{T}$  reactor.

Thus, although there are a number of light element isotopes which might someday be suitable as fuels for advanced fusion reactors, the remainder of this article will concentrate on concepts for fusion reactors fueled by deuterium concentrated from water, and by tritium which would be produced by fusion-produced neutrons from lithium in blankets surrounding fusion reactors.

## POSSIBLE TYPES OF CONFINEMENT FOR FUSION REACTORS

Because temperatures much higher than those to which we are accustomed are required for any fusion reactor with a useful reaction rate, it is important that there be some method for keeping the reacting fusion fuel out of direct contact with material objects which would quickly cool the fuel to temperatures below which fusion reactions were negligible. The principal mechanism responsible for this rapid cooling would not be conduction into the intruding material, but rather radiation losses from the fuel due to enhanced bremsstrahlung and line radiation from atomic transitions due to heavier impurities entering the fuel. Thus, a reactor requires some sort of restraining force which balances the outward pressure of the fusing fuel, which, since it is absorbing part of the energy it releases, is also producing pressure. There are a number of restraining forces one might imagine, not all of which are practical in a power plant.

### Gravitational Fields

Gravitational fields produce suitably high restraining forces only for very large assemblages of mass, because the gravitational force is much weaker than the other known fundamental forces of nature. Thus, while all the successful fusion reactors we can see, namely stars, use gravitational fields for confinement, this is impractical for a commercial reactor by an enormous factor.

### Dc Electric Fields

If the reacting fusion fuel possessed a strong net electrical charge, then it might be possible to confine it with dc electric fields. However, in the types of fuel assemblages so far used in fusion research, any net imbalance in charge which developed in the fuel was much too weak to permit confinement solely through the forces that could be applied through dc electric fields. Significant electric fields do develop under some conditions inside the fuel using the magnetic field confinement approach discussed later, and they may play a role in altering the quality of the confinement when they do occur.

### Radio Frequency Electromagnetic Fields

At the high temperatures required for a practical fusion reactor, matter exists not in any of the three states with which we have

everyday experience, solid, liquid, or gas, but rather in a fourth state known as a plasma. In this state, the negatively charged electrons have been stripped away from positively charged nuclei which they otherwise normally encircle to form electrically neutral atoms. The plasma consists of unbound electrons and these nuclei, called ions. This has the result that both the ions and the electrons are highly mobile, and can rapidly rearrange themselves to counter external electric fields applied to the plasma. Because electrons are much lighter than ions, they have far higher velocities if they are at a temperature roughly similar to that of the ions. Thus, the higher electron mobility normally accounts for most of the charge rearrangement which takes place in a plasma to shield out externally applied electric fields. Plasma will be discussed in more detail later in this article.

An approach which has been considered for confining a fusing plasma is to counter the outward pressure of the plasma with the inward radiation pressure of a radiofrequency electromagnetic field. However, for plasmas that would produce fusion power densities appropriate to this type of fusion reactor, the overall outward plasma pressure would be one atmosphere or more. In order to produce a countering inward radiation pressure of about an atmosphere, the electric field strength in the applied wave would need to be about 1 million V/cm. This is difficult to achieve without inducing electrical breakdowns, and eddy current losses due to image charges in nearby walls might be large.

### Inertial Confinement

Inertial confinement is the method used in fusion bombs, more commonly called hydrogen bombs. It works well under those conditions when the radiation from the explosion of a fission bomb, usually called an atomic bomb, is used to produce soft X-rays, which isentropically compress a fuel mixture of deuterium and tritium. The goal of inertial confinement research for fusion reactor applications is to reproduce this effect on a much smaller amount of fuel, and with a far less energetic driver than an atomic bomb. This approach is referred to as inertial confinement because it is simply the inertia of the assemblage which confines the fuel while it is fusing. The fuel is confined for a time approximately given by the time required for the hot ions, with a mean velocity of about  $10^7$  cm/s, to traverse the radial dimension of the fuel, which is much less than 1 cm. Thus, the confinement time is in the sub-nanosecond range. The fusion output is proportional to the product of the density, the ion temperature, and the confinement time. This is more conveniently referred to as  $n\tau T$ , where  $n$  is density,  $\tau$  is confinement time, and  $T$  is the ion temperature. For realistic ion temperatures of a few tens of keV, the very short confinement time of the inertial approach requires compression of the fuel to very high densities, 40 to 400 times the normal density of the solid deuterium and tritium fuel (4), in order to produce fusion energy releases relevant to a reactor.

### Confinement by Magnetic Fields

An alternate strategy for maximizing the  $n\tau T$  product of density, confinement, and ion temperature is to use a much lower density, but a correspondingly longer confinement time. The technique which most naturally fits this approach is to place the plasma in a magnetic field. Since the plasma is composed of electrically charged particles, they are constrained to move

in helical paths along the lines of magnetic force, with the negatively charged electrons spiraling in one direction, and the positively charged ions in the other direction. In the simplest instance of a uniform magnetic field, and in the absence of collisions, a charged particle remains tied to its line of force, although it is free to move along it. Consequently, the overall effect of a uniform magnetic field is to restrict the outward motion of particles across magnetic lines of force, while leaving motion parallel to the magnetic field unimpeded. For a plasma with a pressure perpendicular to the magnetic field lines of 100 kPa (1 atm) (an overall pressure that is of the general order required for a fusion reactor), a countering perpendicular magnetic force of 100 kPa (1 atm) can be produced with a field strength of about 5000 Oersteds, which is quite readily achievable.

### Principal Confinement Approaches of Fusion Research

Research toward the goal of a fusion reactor began in a significant way in 1951 at Princeton University, and has since spread to many different laboratories in a large number of nations. The overwhelming majority of this research has followed either the low-density, high-confinement time approach using magnetic fields, or the high density, low-confinement time path of inertial confinement. Accordingly, the remainder of this article will cover only magnetic confinement and inertial confinement, with an emphasis on magnetic confinement because this has profited from the most research, and is presently closer to practicality.

## MAGNETIC CONFINEMENT FUSION

### Physics of Plasmas

**Debye Length.** Because a plasma is made up of positively charged ions and negatively charged unbound electrons, these two populations of particles interact strongly with each other through the coulomb force. Any substantial displacement of one species relative to the other leads to a large electrostatic potential, which would require some countering force to maintain. In the absence of any externally applied countering force, there is always an intrinsic force available from random thermal fluctuations in a plasma. The electron thermal energy density per degree of freedom is  $\frac{1}{2} n_e k_e T_e$ ; in a plasma with an electron density of  $n_e$  and an electron temperature of  $T_e$ . This energy is available to drive charge separations between the positive ions and the negative electrons. The restoring force is provided, in turn, by the electrostatic energy density associated with the electric field established by the charge separation. For a charge separation length of  $d$ , the electrostatic energy density  $E$  scales approximately as:

$$\frac{1}{2} \epsilon_0 E^2 \approx \frac{1}{2} \epsilon_0 n_e e d / \epsilon_0 d^2$$

where  $\epsilon_0$  is the permittivity of free space, and  $e$  is the fundamental charge of an electron. If this electrostatic energy density is compared to the thermal energy density, then it is apparent that substantial charge separations within the plasma can only take place over distances not significantly greater than  $d \approx \lambda_D$ , where the value of  $\lambda_D$  is given by:

$$\lambda_D = (k_e T_e / n_e e^2)^{1/2}$$

This is called the Debye length, because it was first calculated theoretically for electrolytes by Debye and Huckel in 1923. This equation is the one most commonly used for  $\lambda_D$ , although a more precise definition would include a term  $(1 + ZT_e/T_i)$  in the denominator to account for the ion effects. Here  $T_i$  is the ion temperature and  $Z$  is the ionization state of the plasma ions. It is of most importance for plasmas with high charge state ions, which is not the case for the fuel components of a deuterium and tritium fusion reactor, although there is usually some admixture of light impurities such as carbon and perhaps small amounts of oxygen in the plasma of today's experimental fusion devices. This ion term can also be of importance if  $T_i$  is much greater than  $T_e$ . However, in plasmas of interest to fusion work,  $T_i$  rarely exceeds  $T_e$ , and is often lower than it.

Since both the ions and electrons in a plasma can move freely, they will tend to move so as to neutralize the electric fields arising from charge imbalances, with electrons positioning themselves to shield the electrostatic field from positive ions, and the ions moving to cancel the field from the negative electrons. The electrons, being much lighter and therefore much more mobile than the ions for all plasmas of interest to fusion, account for most of this charge shielding.

The Debye length is one of the fundamental parameters of any plasma. Charge imbalances can occur over distances of the order of the Debye length, but over distances much greater than the Debye length the plasma will not sustain a net electric field unless there is some additional countering force to support it. Similarly, over distances much greater than a Debye length, a plasma will shield out externally applied electrostatic fields. Thus, the electric potential is normally the same throughout a plasma, unless some other force due, for instance, to rotation, alters the balance.

The two criteria for defining an assemblage of ions and electrons as a plasma both involve the Debye length. The first criterion is that the spatial extent of the plasma should be much greater than a Debye length. The second criterion is that there should be many charged particles within a Debye sphere (with a radius  $\lambda_D$ ) so that the statistical treatment underlying the definition of the Debye length is valid. The Debye length is usually small for the plasmas used for laboratory or industrial applications. For a 3 eV arc discharge with an electron density of  $10^{22} \text{ m}^{-3}$ , the Debye length is about  $3 \times 10^{-6} \text{ m}$ , with about  $10^7$  charged particles within a Debye sphere, sufficient to validate the underlying statistical treatment (5).

A plasma more characteristic of those produced in fusion research devices might have a density of  $3 \times 10^{21} \text{ m}^{-3}$ ,  $T_e = 10 \text{ keV}$ ,  $T_i \approx 0.5 T_e$ , and  $Z \approx 1.2$ , with the fact that  $Z$  is not unity arising from light contaminants in the deuterium and tritium fuel. For these conditions, the Debye length is about  $\lambda_D \approx 8 \times 10^{-5} \text{ m}$ , where almost all of the increase relative to the arc discharge is due to the much higher temperature of the fusion plasma. Since the characteristic dimensions of fusion plasmas are usually of the order of a meter or meters in magnetic confinement devices, it is always the case that they are much greater than a Debye length.

The plasmas found in electric arcs, lightning, or fluorescent lights are usually only weakly ionized, with perhaps one to a few percent of the gas molecules ionized. On the other hand, plasmas with the much higher temperatures needed for magnetic confinement fusion research are highly ionized, with thermal neutrals only penetrating into the outer few centimeters of the plasma.

**Cyclotron Motion in Magnetized Plasmas.** If a magnetic field is applied or arises within a plasma, then the motion of the constituent electrons and ions is significantly altered, becoming more ordered along a preferred axis. The equation of motion of a particle with charge  $Z$ , vector velocity  $\mathbf{v}$  and mass  $m$  in a magnetic field of vector strength  $\mathbf{B}$  is given by:

$$m(d\mathbf{v}/dt) = eZ(\mathbf{v} \times \mathbf{B})$$

If we choose the  $z$  axis to be along the magnetic field, then, in cartesian coordinates, the components of the particle motion are as follows:

$$d\dot{v}_x/dt = -\omega_c v_y$$

$$d\dot{v}_y/dt = \omega_c v_x$$

$$d\dot{v}_z/dt = 0$$

with  $\omega_c = |eZB|/m$ , and with  $B$  the scalar magnitude of the magnetic field. The fact that the particle velocity is constant along the magnetic field arises from the fact that, as shown in the vector cross product above, magnetic fields have no effect on a velocity component parallel to the field. The  $x$  and  $y$  components are both perpendicular to the magnetic field. The time derivatives of a particle in these dimensions correspond to circular motion with a frequency of  $\omega_c$ , which is called the cyclotron frequency. The combination of this circular motion perpendicular to the magnetic field and the uniform velocity parallel to the field (for uniform magnetic fields) give rise to a net helical particle orbit spiraling along the magnetic field, as shown in Fig. 1.

In a plasma with an isotropic Maxwellian velocity distribution, the temperature perpendicular to the magnetic field is equal to the temperature parallel to it. However, because the perpendicular motion encompasses two degrees of freedom while the parallel motion has only one, for these conditions the perpendicular motion of the plasma particle distribution carries twice the energy content of the parallel motion. It is not always the case that the plasma temperature is isotropic with respect to the magnetic field direction. Some heating mechanisms transfer energy preferentially into either the parallel or the perpendicular motion of the particles with respect to the magnetic field direction. Similarly, some particle loss processes involving collisions preferentially deplete either the perpendicular or the parallel energy, depending upon whether the process

is stronger for large perpendicular energies or large parallel energies.

**Larmor Radius in a Magnetized Plasma.** The radius of the perpendicular motion around the field line is called the Larmor radius, denoted as  $\rho$ . It is given by the ratio of the perpendicular velocity to the cyclotron frequency, and scales for a particle moving at the thermal velocity as:

$$\rho = [2^{1/2} v_{\perp}]/|ZeB|$$

where  $v_{\perp}$  is the thermal velocity, and the absolute value of the charge is taken to ensure a positive length. The factor of square root of two arises from the two degrees of freedom in the perpendicular motion. For any plasma, whether Maxwellian or not, there will clearly be a range of values of the Larmor radius corresponding to the perpendicular velocity distribution. Inserting the values of electron and baryon masses leads to the following more specific formulas for the cyclotron frequencies and Larmor radii of thermal particles in plasmas: for an electron:

$$\omega_c = 1.76 \times 10^{11} B \text{ s}^{-1}$$

$$\rho_e = 1.07 \times 10^{-4} T_e^{1/2} / B \text{ m}$$

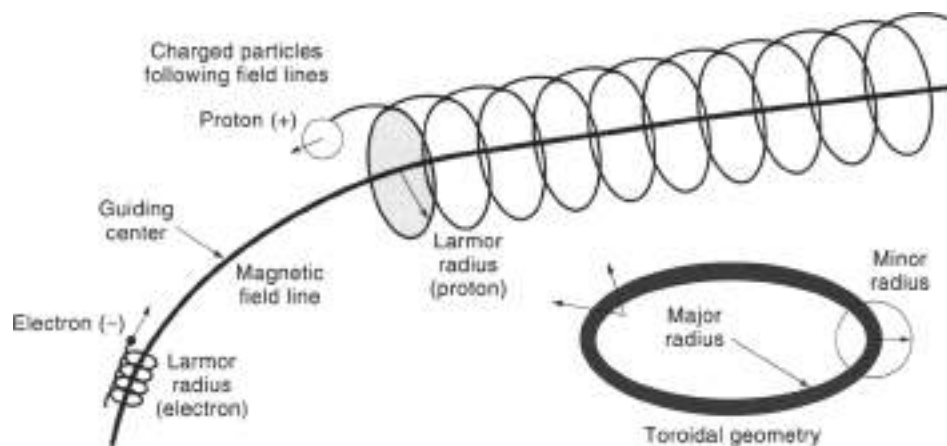
for an ion of charge  $Z$  and a mass of  $A$  atomic mass units:

$$\omega_c = 0.56 \times 10^6 Z/A B \text{ s}^{-1}$$

$$\rho_i = 4.57 \times 10^{-4} (A^{1/2}/Z) T_i^{1/2} / B \text{ m}$$

In these equations the electron and ion temperatures are expressed in keV. For comparison, room temperature is about 0.025 eV. The magnetic field is given in units of tesla, where a tesla is equal to 10,000 gauss, or roughly 25,000 times the average strength of the earth's surface magnetic field.

In a 5 T magnetic field, which corresponds to a typical field strength on the plasma axis of the Tokamak Fusion Test Reactor at Princeton University, a deuteron would have a cyclotron frequency of  $\omega_c = 2.4 \times 10^6$  per second. A plasma electron in the same field would have a cyclotron frequency of  $\omega_c = 8.79 \times 10^{11}$  per second. The much higher frequency of the electron cyclotron motion is a consequence of the much smaller mass of the electron relative to a deuteron. For the same reason, the Larmor radius of an ion is much larger than that of an electron. For the case of a 5 T magnetic field confining a plasma



**Figure 1.** Charged particles spiraling along a magnetic field line. For electrons and protons with the same energy, the Larmor radius for the electrons would be 1/43 of the proton's. The average motion can be described by the guiding center, so long as its parameters change significantly across a Larmor orbit. Drawn by T Stevenson.

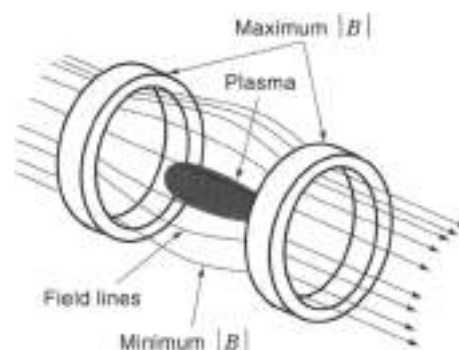
with central temperatures of 40 keV for the ions and 12 keV for the electrons, which is reasonably representative of higher performance plasma parameters on Princeton's Tokamak Fusion Test Reactor, the Larmor radius of an electron with the thermal energy would be  $r_L = 0.0073$  cm, and for an ion with the thermal energy it would be 0.58 cm. The full range of Larmor radii would include values ranging from a bit smaller than this up to a few centimeters (on the tail of the Maxwellian distribution) that would be a factor of 2 or more larger. These Larmor radii are small compared to the dimensions of the fusion devices in which they presently occur. The position of a charged particle averaged over its cyclotron motion is called the guiding center. In the absence of perturbations such as collisions, the guiding center of an electron or ion moves along a line of magnetic force (with the electron guiding centers moving in one direction, and the positive ion guiding centers in the other). So long as the gradual scale lengths for plasma properties such as density, temperature, and magnetic field strength are much larger (usually a factor of a few is sufficient) than it is a good approximation to model most types of particle behavior by following the guiding centers rather than the detailed gyromotion, which is more complicated computationally, and requires more computer time.

**Plasma Diamagnetism.** The cyclotron orbits of particles in a plasma each enclose small volumes of magnetic flux. As a consequence of Lenz's law, the directions of rotation of both the electrons and the ions are such that the tiny solenoidal currents they represent produce magnetic fields in the opposite direction from the field they are enclosing. Thus, the gyromotion at the cyclotron frequency reduces the total strength of the field inside a particle orbit. This is referred to as plasma diamagnetism. Its importance increases as the plasma pressure increases or the externally applied magnetic field decreases. For high ratios of the plasma pressure to the applied magnetic field strength, this diamagnetism can hollow out the magnetic field and produce a region in the central plasma with a magnetic well, that is, a region which is everywhere surrounded by increasing magnetic field strength. This approach has been suggested as a technique for confining plasma, but its practicality is not yet clear.

**Magnetic Moment.** A particle gyrating around a magnetic field line constitutes an electric current  $I = iZq\omega_c/2\pi$ , which encompasses an area of  $A = \pi r_L^2$ , where  $Z$  is the ionization state,  $q$  is the fundamental charge,  $\omega_c$  is the cyclotron frequency, and  $r_L$  is the Larmor radius. The product of the current and the enclosed area is called the magnetic moment of the particle orbit, and is generally denoted as  $\mu$ . Thus,

$$\mu = IA = \omega_c r_L^2 / 2B = W_{\perp} / B$$

Here  $r_L$  is the velocity of the gyrating particle in the plane perpendicular to the magnetic field line it is following, and  $W_{\perp}$  is the kinetic energy associated with this perpendicular velocity. It can be shown (5) that, in the absence of collisions or electric fields, the magnetic moment  $\mu$  of a particle orbit is an invariant. This has far-reaching consequences for plasma confinement schemes utilizing magnetic fields. The invariance of a particle orbit's magnetic moment means that as the particle gyrates along a line of magnetic force into a region of stronger magnetic field, the perpendicular velocity increases so that the



**Figure 2.** The simplest form of mirror confinement. The trapped portion of the particle velocity distribution reflects from the stronger magnetic fields at each end of the cell, forming a confined plasma in the region between the mirrors.

energy of rotation perpendicular to the field line increases by the same factor as the magnetic field strength. Inasmuch as the particle's total kinetic energy is also constant in the absence of collisions or electric fields, the increase of the perpendicular energy as a particle gyrates into an increasing magnetic field implies that the kinetic energy of the particle parallel to the field line decreases, and thus that the velocity along the field line also decreases. Thus, for a particle orbit which has a finite energy of perpendicular rotation along any part of its path, there can exist some magnetic field strength at which its velocity parallel to the field line goes to zero. When this happens, the particle is reflected back into the region of weaker magnetic field. This is referred to as *magnetic mirroring*, and plays a significant role in every form of magnetic confinement. If a region of weaker magnetic field is bounded by a stronger magnetic field at each end, then particles can be reflected back and forth between the regions of stronger magnetic field, producing a trap, as shown in Fig. 2. This works for both positively charged particles (most ions) and negatively charged particles (for example, electrons). For a particle with parallel velocity  $v_{\parallel}$  and total kinetic energy  $W$ , the parallel velocity will vary as

$$\omega v_{\parallel}^2 / 2 = W_{\parallel} - \mu B$$

Clearly there is one class of particle orbits for which magnetic mirroring would have no effect at any magnetic field strength, no matter how strong. Particles which are everywhere moving entirely parallel to the local magnetic field have no gyro-orbit, and thus  $\mu = 0$ . As a consequence, their parallel velocity is unaffected by changes in the magnetic field strength, and these particles are not reflected.

**Magnetic Mirror Confinement.** More generally, for finite ratios of magnetic field strengths in a plasma confinement device, a considerably broader range of particle orbits is not mirror confined. If we define the minimum magnetic field strength along a line of force to be  $B_m$  and the maximum to be  $B_M$ , then the constancy of  $\mu$  and  $W_{\perp}$  lead to the condition that all particle orbits with  $\mu > W_{\perp} / B_M$  are trapped by the magnetic mirror field. If this were not so, then particles could reach the point of maximum field strength with their perpendicular kinetic energy greater than their total kinetic energy, which is

not possible. Applying this principle to determine the boundary between mirror trapped and untrapped particles, one finds the conditions for marginally trapped particles, where  $W_{\perp}(\min B)$  is defined as the energy of perpendicular rotation when the particle is in the region of minimum magnetic field strength, and  $W_{\parallel}(\min B)$  is defined as the parallel energy of the particle at the minimum magnetic field:

$$W_{\perp}(\min B) = \mu R_{\perp} = W_{\parallel} B_{\min} / B_M$$

$$W_{\perp}(\min B) / W_{\parallel} = (1 - B_{\min} / B_M)^{-1}$$

These conditions can also be written in terms of the ratios of the perpendicular and parallel velocities evaluated at the minimum magnetic field with respect to the total velocity,  $v$ , of the particle, giving:

$$v_{\perp}(\min B) / v = (B_{\min} / B_M)^{-1/2}$$

$$v_{\parallel}(\min B) / v = (1 - B_{\min} / B_M)^{1/2}$$

These equations define the boundary between mirror trapped and untrapped particles in the space of parallel versus perpendicular velocities. Thus, charged particles with a velocity ratio  $v_{\perp}(\min B) / v$  which is sufficiently low will be trapped. On the other hand, charged particles will promptly escape along the mirror field if they satisfy the criterion:

$$v_{\parallel}(\min B) / v > (1 - B_{\min} / B_M)^{1/2}$$

This can be rewritten in terms of the ratio of the parallel and perpendicular velocities evaluated at the minimum magnetic field strength:

$$v_{\parallel}(\min B) / v_{\perp}(\min B) > (B_M / B_{\min})^{-1/2}$$

This condition defines two regions in velocity space, in which particles with higher ratios of  $v_{\perp}(\min B) / v_{\parallel}(\min B)$  are trapped, and ones with lower values of this ratio are lost. Since the perpendicular velocity corresponds to rotation in two dimensions, the boundary between trapped orbits and loss orbits in a three-dimensional velocity space forms a cone. Thus, particles with orbits that have ratios of parallel velocity to perpendicular velocity which are too great for trapping are said to fall into the *loss cone*, depicted in Fig. 3.

A mirror trapped plasma can never be isotropic in velocity space. This is because the transit time along a magnetic field line is very fast, so particles with velocity ratios that fall within the loss cone will rapidly escape through the mirror throat (region of maximum  $B$ ). Thus, the average perpendicular energy

in a mirror-confined plasma is larger than the parallel energy by more than the factor of 2 one would expect just from the relative number of degrees of freedom involved.

Note that neither the charge nor the mass of the particles enters directly into the criteria for trapping. In an entirely collisionless plasma, the particles within the loss cone would be entirely lost within a transit time (the time required for the slowest particles to travel the length of the region enclosed by the magnetic mirrors). After this very rapid loss, the remainder of the plasma would remain confined for as long as the magnetic field configuration was maintained.

In all systems of interest to fusion research, this static condition never arises because anisotropy collisions between mirror-confined particles alter their ratio of perpendicular to parallel velocity, creating a dynamic loss due to particles which scatter from trapped velocity ratios to untrapped ratios. Once in the loss cone, these newly scattered particles will rapidly escape the confinement device unless their ratio of perpendicular to parallel velocity is scattered back into the trapped region by a subsequent collision or collisions. Accordingly, although the loss cone is the same for different types of particles in a plasma, the particle species which collides most frequently, and thus will most rapidly scatter into the loss cone, is the one which is preferentially lost.

For a plasma in which the temperatures of the electrons and the ions are of the same order, the electrons are the more collisional species, so they initially escape through the mirror throats more rapidly than the ions. This immediately gives rise to a net negative charge leaving the mirror throats, and a net positive charge remaining in the central cell of the mirror confinement device. This charge imbalance produces an electric potential which is positive in the central cell, and thus an electric field which retards the loss of electrons along magnetic field lines through the mirror throats. This electric field causes lower energy electrons to be confined which would otherwise be lost, and, correspondingly, it slightly increases the parallel velocities of positive ions as they approach the mirror throats. Thus, the electric field strength will rapidly build up only to the strength at which it impedes the loss of lower energy electrons sufficiently so that the net electron loss just balances the loss of positive ions. In a mirror-confined plasma it will be the more energetic electrons which are able to escape over the electrostatic potential hill formed at the mirror throats, so the energy balance of mirrors is dominated by these electron thermal losses. The equilibrium rate at which electrons and positive ions escape from the mirror-confined plasma is governed by the rate at which the less-collisional ions scatter into the loss cone, so ion collisions set the rate of particle loss.

During the early decades of fusion research, devices incorporating magnetic mirrors as their primary confinement mechanism were extensively investigated. In their simplest form these machines consisted of a central cell permeated by a uniform solenoidal magnetic field, bounded by a stronger magnetic mirror field at each end. These mirror machines were examples of "open" confinement devices because their magnetic field lines left the region of plasma confinement through the throats of the mirrors. This geometry had a practical appeal because it is mechanically simpler to build a linear device than one with some more complex shape. Moreover, mirrors are devices which do not require a time varying field, and thus are intrinsically capable of continuous operation if the engineering systems are designed accordingly. In a practical mirror fusion reactor the

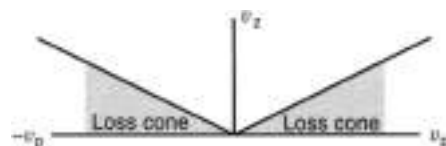


Figure 3. Schematic sample of loss cones in the perpendicular versus parallel velocity space of a magnetic mirror-confined plasma. Particles which undergo collisions such that their  $v_{\perp}/v_{\parallel}$  drops into the loss cone quickly escape along the field lines at the magnetic mirror throat.

magnetic field axis would be superconducting in order to avoid ohmic dissipation.

Simple mirror confinement machines encountered difficulties due to macroscopic instabilities which rippled the plasma surface and greatly augmented the loss of plasma particles and energy. These problems were addressed by confinement schemes which added additional current-carrying conductors to alter the magnetic field in the central cell so that it increased toward the edge of the plasma. The intrinsically anisotropic velocity distribution residing within a mirror drives another smaller scale length microinstability, called the loss-cone instability, which increases particle and energy loss. Finally, of course, there is the overriding consideration that even if all losses across the magnetic field were eliminated, the loss of plasma particles and energy along the magnetic field lines through the mirror throats would remain.

Even in the absence of macroscopic or microscopic instabilities, the theoretical limit on the energy gain of a simple mirror with a single confinement cell is relatively modest. We can define  $Q$  as the ratio of the nuclear power (thermal) produced by the plasma to the amount of power supplied via external means to the plasma in order to heat it to the high temperatures required for fusion reactions. Assuming a conservative efficiency for converting thermal energy to electricity of about 33%, and keeping in mind that a practical reactor must produce considerably more electricity than it uses, an energy multiplication factor of  $Q = 15$  to 20 or more is often considered to be sufficient for an economic fusion reactor using a fusion-heated steam cycle to drive the electrically-producing turbines.

Due to the end losses along the magnetic field lines, it would be difficult even in principle to achieve an economically acceptable  $Q$  with a simple mirror machine. For a simple mirror, free of any instabilities, and using as fuel deuterium and tritium, the energy multiplication would be at best  $Q = 2 \log(B_{\text{in}}/B_{\text{m}})$ . The logarithm in this expression can never be greater than an order of unity (that is, with a value of one to at the very most a few). This is because the mirror ratio, the ratio of the maximum to minimum magnetic field in the device, cannot be made arbitrarily large. Such practical considerations as mechanical forces and the maximum current densities superconductors can carry limit the maximum field strength that can be produced in a mirror of useful size, while the minimum field strength cannot be reduced too much without allowing instabilities and diffusion across the magnetic field in the central cell to increase.

In order to circumvent these limitations, researchers contrived more sophisticated mirror confinement schemes such as tandem mirrors. In this approach, the central confinement cell is enclosed by an additional mirror cell at each end which accepts particles escaping through the velocity loss cone of the mirror field of the central cell. The plasma in these end cells, and particularly the electron component, is heated to very high temperatures to increase the magnitude of the electrostatic field plugging the mirror throat from the central cell. Increasing this negative potential in the throat means that more energetic electrons are confined in the central cell, thus reducing the parallel heat loss, and permitting the theoretically obtainable  $Q$  to be greater than for a simple mirror.

Maintaining the proper velocity distributions of particles in tandem mirrors without driving instabilities proved quite daunting so that, in the mid-1960s, the experiments with the mirror approach to fusion were largely abandoned. Since that

time almost all magnetic confinement fusion has gone into the other general class of devices called *closed field systems*, in which the magnetic field lines pervading the central confinement region do not leave the plasma, but instead circle around to eventually reconnect with themselves.

**Closed Magnetic Field Line Confinement Devices.** Experimental confinement machines with closed magnetic fields have in most cases used a topology which is basically toroidal, or toroids elongated with straight sections. Such devices were pioneered in western countries by Professor Lyman Spitzer at Princeton University in 1951, and at about the same time by researchers at the I. V. Kurchatov Institute in Moscow. Although similar in many respects, the approaches which these laboratories pursued in the 1950s and early 1960s differed in some important ways.

**Rotational Transform.** All closed field confinement systems are defined by the fact that their magnetic field lines wrap around to reconnect with themselves after some angular translation along the major circumference of the device. The simplest topology for a closed system, and the one which has been most often used for large fusion experiments, is a torus. In this approach, current-carrying coils are arranged around a vacuum vessel so that they produce a toroidal, or donut-shaped, magnetic field along which the confined plasma ions and electrons can gyrate. This geometry is the topological equivalent of deforming a linear magnetic solenoid into a torus by bending it into a circle such that the ends touch. In a toroidal geometry, in contradistinction to the case for a linear system, each of these toroidal field coils has an inner leg and an outer leg with respect to the vacuum vessel and the plasma it contains. It is clear that the inner legs of the toroidal field coils carry the same current as the outer legs, but the outer legs on a torus are much farther apart from each other than are the inner legs, which are often sized so that they abut or nearly abut each other. This gives rise to an asymmetry between the average current density flowing along the inner circumference of the coil system where the coils nestle near each other compared to the outer coil circumference where the coils are widely separated. This asymmetry in turn produces a gradient in the strength of the magnetic field produced by the coils, the field is stronger on the inside of the torus, and weaker on the outside.

A toroidal geometry is most naturally described by two dimensions, the major and minor radii of the plasma toroid, where the major radius,  $R_0$ , is the distance from the central axis of the coil arrangement to the axis of the confined plasma. The minor radius,  $r_0$ , is the distance between the plasma axis and the outside of the plasma cross section. For circular coil systems and plasmas where the ratio of the perpendicular plasma pressure to the inward magnetic pressure is fairly low, the effective major and minor radii are quite similar to what one would get from simply looking at the plasma as a donut. For elongated coils and plasmas, and for high-pressure plasmas which rearrange flux surfaces, these effective radii are somewhat modified.

If the only magnetic field threading a plasma torus is the toroidal field, then the plasma will be only momentarily confined, no matter how strong the toroidal field is made. This arises as a result of the fact that there is a gradient in the toroidal magnetic field strength running from a high level on the inside to a lower level on the outside. This magnetic

gradient causes the negatively charged electrons to drift transverse to the gradient in one direction, and the positively charged ions to do the same thing in the other direction. These oppositely directed vertical drifts produce an imbalance in the net neutrality of the plasma above the midplane, compared to the plasma under the midplane. One side will have excess positive charge, while the other will have excess negative charge. These charge accumulations will establish a vertical electric field which is perpendicular to the toroidal magnetic field. In turn, the combination of this electric field and the magnetic field drive a type of particle drift with a drift velocity of  $v = (E \times B)/B^2$ . For toroids as described here, the direction of this drift is radially outward along the direction of the major radius. This mechanism drives both the negative electrons and the positive ions out together, so no electric field is established which would limit this drift. As the plasma drifts out to larger major radii it quickly runs into the material boundary of the containment device, and is extinguished as impurity influx leads to large line transition radiation losses. Reference 5 contains excellent descriptions of the two types of particle drifts mentioned here, as well as other sorts of drifts which occur in magnetized plasmas.

The fact that a simple toroidal magnetic field cannot confine a plasma is a consequence of the field gradient that necessarily arises from the geometry of any toroid. Thus, there is nothing which can be done to prevent the oppositely directed drifts of the electrons and ions perpendicular to the magnetic gradient.

What can be done instead is to short out the charge imbalance which would otherwise develop by adding a helical twist to the magnetic field lines so that particles gyrating along these lines of magnetic force will spend half of their time above the tokamak midplane and half below the midplane. Since the vertical drift of each species is independent of whether the particle is above or below the midplane, this means that the unidirectional drift of each particle is outward half of the time, and inward the other half, so there is no net change of position.

The helical twist in the confining magnetic field lines is described by the rotational transform, which is the amount a field line moves in poloidal angle (the angle around a cross section of the plasma donut) as it traces itself around the plasma in toroidal angle. The inverse of this quantity, called the safety factor, and denoted as  $q$ , is commonly used in describing toroidal magnetic confinement systems:

The local value of  $q$  will in general be different for each magnetic flux surface within a toroidal plasma, but it will be uniform on any given flux surface. For flux surfaces with a minor radius significantly smaller than the major radius, which is the case for all of the large devices so far built, the value of the dimensionless quantity  $q$  can be well approximated as:

$$q = (r R_0 B_z) / (R_0^2 B_\theta)$$

Here  $r$  is the minor radius of the flux surface being described,  $B_z$  is the toroidal component of the magnetic field,  $R_0$  is the plasma major radius, and  $B_\theta$  is the poloidal component of the magnetic field (the component which imparts the helical twist to field lines on the flux surface). If a field line returns to its starting position after exactly one circuit of the torus, then  $q = 1$  on that flux surface. If, for example, 2.7 transits around the torus are required for a field line to return to its initial poloidal location, then  $q = 2.7$ . For values of  $q > 1$  in a circular cross section toroid with the major radius significantly larger than the minor radius, a particle gyrating along a field line experiences

a net magnetic well when averaged along its orbit, which tends to abet confinement.

Within a confined plasma of a device of the tokamak type to be discussed shortly, there are arbitrarily many local values of  $q$  associated with flux surfaces. However, the class of surfaces with rational values of  $q$  often play a special role either in the large-scale internal stability or the gross stability of the plasma. Rational magnetic surfaces are ones for which the safety factor can be represented as  $q = m/n$ , where  $m$  and  $n$  are integers, and a field line comes back to its original position after  $m$  toroidal and  $n$  poloidal rotations around the torus. Rational magnetic surfaces have proven particularly susceptible to magnetohydrodynamic (MHD) instabilities, which are large-scale fluid-like perturbations of the plasma and the field lines within it. If they occur on rational flux surfaces well inside the plasma, they can increase the radial loss of energy and particles; if they occur near the outside of the plasma, they can cause the plasma to disrupt.

The magnetic surfaces with low values of rational  $q$  tend to be the most pathological, since they are most susceptible to current-driven instabilities. Early researchers found that it was usually easier to maintain gross plasma stability if the surfaces with low rational values of  $q$  were buried well inside the plasma, which meant that the  $q$  value of the outside was large (as much as 8 to 10 or so). Thus, the dimensionless quantity  $q$  came to be called the safety factor, since it was easier to maintain the overall stability of most sorts of toroidal plasmas if the edge  $q$  was larger.

**Ways to Introduce Rotational Transform.** The pioneers of fusion research in 1951 realized from the beginning that a rotational transform would be needed in the field lines of a toroidal or spherotoroidal confinement device. Where they differed in the mechanisms they used to produce the rotational transform. The group at Princeton University imposed the rotational transform externally by applying magnetic fields from helical coils spiraling around the vacuum vessel inside the much more powerful toroidal field coils. Alternatively, in some early experiments researchers obtained the rotational transform without the helical coils, and instead with an elongated racetrack type of design distorted so that the two straight sections crossed over each other. These closed field devices in which the rotational transform arose from the external configuration or special coils were named *stellarators* because of the astrophysics background of their inventor, Lyman Spitzer, Jr., of the Princeton University astrophysics department.

Stellarators had two particularly appealing characteristics. One was that they did not require any net current in the plasma to maintain the rotational transform, making them less susceptible to damaging disruptions and obviating the need to find a way to drive the current. Since the required fields were static, a stellarator was intrinsically capable of steady-state operation if the coils were designed for this. The other major desirable feature was that, with the magnetic fields all imposed by external coils, position control was relatively straightforward, so the plasma could be kept out of contact with material surfaces. This meant that keeping the plasma fairly free of impurities, which would dilute the fuel and increase energy losses through radiation, should be feasible.

A disadvantage of stellarators was that the helical coils producing the rotational transform crossed the toroidal field lines produced by the encircling toroidal array of coils. This produced

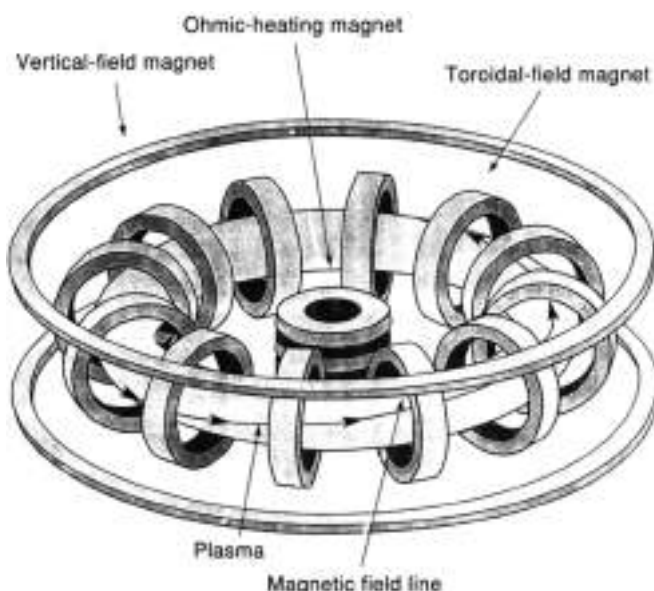


large mechanical stresses which were difficult to accommodate in the 1950s. Partly as a result of this engineering difficulty, no stellarators with large minor radii were built during that period. These early small-cross-section stellarators (typically with minor radii of order 10 cm or even less) achieved disappointing results, which may have been in some part because they were so small and relatively cool (compared to a stellar core) that low-energy neutral particles could penetrate throughout the plasma, undergoing charge exchange with the hotter confined particles which, once neutralized, would escape the confinement. In the late 1960s the stellarator approach was largely abandoned in response to more favorable results obtained at the I. V. Kurchatov Institute in Moscow using the tokamak approach, to be discussed next. In later years, as the understanding of both the physics and the engineering of fusion devices improved, some researchers have returned to the stellarator approach and variations thereon. Better performance was obtained with larger plasma cross-section stellarators built in Germany by the Max Planck Institute, and a much larger stellarator variant is being built outside Nagoya, Japan, by the National Institute of Fusion Studies. These later stellarators have mostly employed magnetic field coils with complex shapes which fulfill the role of both the toroidal field coil and the helical coils of the early stellarators.

**Tokamaks.** Soviet scientists chose another way to introduce the required rotational transform which employed the toroidal field coils of the stellarator concept, but no helical coils to introduce the poloidal magnetic field component. They instead drove an electric current which flowed toroidally through the plasma. This current produced an encircling poloidal field which vector-added to the toroidal field from the coils to produce a spiraling field with different values of the rotational transform, and thus of  $q$ , on each flux surface. The value of the outer surface  $q$ , and consequently of the inner surface  $q$ 's could be easily altered by driving different amounts of plasma current for a given value of the toroidal field. This device was named the *tokamak*, which means something like "magnetic bottle" in Russian. Figure 4 shows the basic schematic of a tokamak.

**Inductively Driven Plasma Current in Tokamaks.** In all the early tokamaks, and in many later ones, all or nearly all of the plasma current was driven inductively. This was accomplished by adding a transformer solenoid to the tokamak, with the plasma acting as a one-turn secondary winding. As with any transformer, current would continue to be driven in the secondary, that is, the plasma, only so long as the flux produced by the primary winding was changing. Since there is always a practical limitation on the volt-seconds available to drive such a flux swing, and since there also limitations upon the current densities which conductors can carry, inductively driven tokamaks are intrinsically pulsed devices.

In the present generation of tokamaks, these pulse lengths range from a few seconds to a minute. Inductively driven fusion power reactors of the future might have much longer pulses, perhaps of many hours, but they would still have to pause at some regular interval to reset the current in the primary transformer winding. The resulting period during which the fusion plasma was turned off would be short compared to the time when it was making power. Depending upon the thermal inertia of the heat blanket surrounding the tokamak, where heat exchange and tritium breeding take place, the transformer



**Figure 4.** Basic components of a tokamak magnetic confinement configuration. The coils surrounding the plasma produce a toroidal magnetic field, while the current in the plasma creates a weaker poloidal magnetic field. The combination of the two makes field lines which spiral as they are followed around the torus.

recharge interval might produce either relatively little or significant fluctuation in the electricity output. Perhaps more significantly, pulsed thermal and magnetic field loads are expected to increase the stress on components near the plasma, thereby reducing their lifetimes. As will be discussed later, much progress has been made in finding noninductive ways to drive the plasma current, allowing steady-state operation of future power reactors.

**Plasma Limiters in Tokamaks.** Another possible disadvantage of tokamaks in their simplest form is that, since the plasma is a current carrying loop, it tries to expand in major radius. This tendency can be countered, however, by adding coils which produce a vertical magnetic field. A more serious concern of early researchers was that simple circular-cross-section tokamak plasmas required contact with a material limiter in order to stabilize their position. This could act as a source of impurities, and would require large amounts of cooling in a power reactor. Most of the more recent tokamak experiments such as DIII-D at General Atomics in San Diego have replaced the limiter with divertors, to be discussed later, which appear to be much more suitable for reactors.

## MAJOR COMPONENTS OF A TOKAMAK FUSION REACTOR

### Plasma Confinement System

The quality of energy confinement that can be achieved within a tokamak design is the single most important factor in determining the most feasible design for a fusion power plant. If the rate at which energy leaks out of the plasma is too high, then prohibitively large amounts of power will be required to heat and maintain the plasma at the temperatures required to

produce useful amounts of fusion energy. The confinement time is characterized by a quantity,  $\tau$ , which gives the time required for the energy content of the plasma to decline by an  $e$ -folding. At the dawn of fusion research, energy confinement times were typically much less than a millisecond. Today they are commonly hundreds of milliseconds, and in some high-confinement plasmas the confinement time is more than a second.

The basic components of the confinement system are the various current-carrying coils which produce those magnetic fields that are externally applied. In addition to the toroidal field coils and vertical field coils, there are coils wrapped around the plasma to alter its cross-sectional shape, and in some cases additional coils might be added to control gross instabilities. More recently, the systems which deposit energy, momentum, and additional current into the plasma have also become important in improving confinement.

The coils used to produce magnetic fields for nuclear fusion experiments have usually been made of copper. While, in a few cases, they were cooled to reduce their electrical resistance at the start of an experimental pulse, these were, nonetheless, ordinary resistive coils. Actual commercial fusion reactors are likely to rely heavily upon superconducting coils, both to improve the overall electrical efficiency and to permit long pulses at high magnetic field strengths.

As fusion experiments have grown more ambitious over the decades, and as efforts to develop the engineering basis for fusion have progressed, superconducting coils have gradually become more common in operating and planned experimental devices. In the early 1970s a magnetically levitated superconducting coil was used to confine plasmas at the Princeton University Plasma Physics Laboratory. A tokamak at the Kurchatov Institute in Moscow has used superconducting coils since about 1975, and the Large Helical Device, a stellarator at the National Institute for Fusion Studies in Japan, began operations with superconducting coils in the late 1990s. Superconducting tokamaks capable of plasma pulses of 100 seconds or more are presently under construction in South Korea and India. An upgrade to the large JT-60U tokamak in Japan which would use superconducting coils to allow plasma pulses of 100 seconds or more is being proposed, and the next very large international fusion experiment is also expected to incorporate superconducting coils.

**Collisions among Confined Particles.** Getting to the high levels of energy confinement that can be obtained in today's tokamaks has been a long journey. The tokamak confinement configuration is characterized by a magnetic field topology which guides the confined ions and electrons along helical field lines on toroidal flux surfaces. A single ion or electron could travel along such field lines forever if there were no imperfections. In practical applications, there are always some magnetic field imperfections, and, much more importantly, there are many ions and electrons instead of one. The ions and electrons each collide among themselves and with each other. Sometimes when the ions collide they fuse, releasing the energy which is the purpose of the tokamak, but much more often they scatter from each other. The scattering within and among the different confined species gives rise to transport of energy and particles across the confining magnetic field lines, and thus eventually out of the plasma. If this scattering by Coulomb collisions between pairs of particles were the only mechanism by which energy leaked from the plasma, then achieving the quality of confinement

necessary for a fusion power plant would have been a daunting, but nonetheless conceptually straightforward, undertaking.

**Anomalous Diffusion.** The reality has been that the early plasma confinement experiments encountered energy leakage rates much greater than would be expected if the energy loss were primarily due to diffusion arising from two-particle collisions. This type of two-body collisional diffusion was somewhat understood, and could be calculated with neoclassical theory, which took account of effects arising from toroidicity. The excess energy leakage was called anomalous diffusion, and was presumed to be driven primarily by collective instabilities in which great numbers of ions and electrons moved together to form waves and large-scale plasma deformations that transported energy much more rapidly out of the plasma than could two-particle collisions. The early years of fusion research were plagued by anomalous energy losses so large that the confined plasmas could not be successfully heated to and maintained at the very high temperatures necessary for practical fusion power.

**Enhanced Confinement.** During the decades stretching from the 1970s through the 1990s, progress occurred as successive generations of tokamaks led to a better understanding of how the anomalous energy leakage could be reduced by varying such factors as the dimensions of the plasma toroid and the magnitude of the current flowing within it. At the same time, modes of tokamak operation were discovered with improved confinement: that is to say, significantly reduced anomalous energy leakage. These enhanced confinement regimes were characterized by different profile shapes for the density and temperature than were found in the usual tokamak plasmas, and also by reductions in the rate at which neutral gas was recycling from the walls into the outer part of the plasma. Some of these enhanced modes also involved changes in the current density profile, and in the relative speed at which different portions of the plasma rotated.

The improved modes were usually called by phenomenological names because they were mostly discovered experimentally rather than first being predicted theoretically: fusion research makes progress by building new tokamaks. Two of the most famous confinement modes were the H-Mode, which was discovered in Germany with the H standing for high confinement to distinguish it from low confinement, and the Supershot Mode, discovered at Princeton in 1986 in the Tokamak Fusion Test Reactor. A supershot plasma was distinguished by having a region in its core with very good energy confinement, and by temperatures much higher than had previously been achieved. Supershot plasmas were also the first ones in a tokamak to exhibit the phenomenon of bootstrap current. The bootstrap current arises from a dynamic effect within the plasma as the charged particles press across magnetic field lines. Finding the bootstrap current enhanced the economic viability of future fusion power plants because its existence means that most of the electric current needed to sustain the poloidal component of the magnetic field can be supplied at little or no extra cost by the plasma itself. Similar sorts of plasmas were subsequently produced in Japan's JT-60U tokamak and in Britain at JET, the principal tokamak of the European Union. It was the Supershot mode which was selected as the best route to follow in the Tokamak Fusion Test Reactor to produce fusion thermal power levels of up to 10.7 MW.

**Enhanced Reversed Shear Mode.** In the mid-1990s, an even better operating mode was found at Princeton's TFTR as well as at the European Union's JET, and subsequently at General Atomics in San Diego, and Japan's JT-60U. This discovery grew out of a study of reversed shear plasmas which had been pursued on several tokamaks. Shear is the rate of change of the rotational transform of the magnetic field lines in passing from one nested flux surface to the next one. It is governed by the shape of the current within the plasma, which produces the poloidal field component that determines the transform. Most tokamak discharges have current profiles that yield shear which is always in the same direction in passing through successive flux surfaces. Techniques were found in recent years which modified the plasma current profile in such a way (by putting more current in the outer portions of the plasma) that the magnetic shear was reversed in the plasma interior. In the central plasma the rotational transform changed oppositely from what it did in the outer plasma in going through successive flux surfaces.

In many cases this reversed shear magnetic field configuration did not produce dramatic improvements in confinement. However, in some cases, the confinement in the central region of the plasma improved markedly in terms of ion energy leakage, and particle confinement of both the ions and electrons. Named the *enhanced reversed shear mode* by researchers at Princeton University's TFTR, who were the first to observe the abrupt transition to this mode with high levels of energetic particle heating power, it was quickly produced on other tokamaks. In retrospect, it was found to be very similar to a type of confinement mode which had been studied earlier in the European Union's JET tokamak with a different heating technique.

Enhanced reversed shear plasmas have nearly perfect particle confinement in their cores, and the rate of energy loss through ion-ion collisions is for the first time down to the theoretical minimum due to two-body collisions. In fact the conductivity is so low that it required a revision of the neoclassical theory, which had previously not taken account of orbit effects in regions with very steep gradients in plasma parameters.

At present this enhanced reversed shear regime can only be achieved transiently for periods of roughly 0.2 s to 1 s because the required plasma current profile is produced by rapidly increasing the total plasma current in an already hot plasma. Because the plasma is hot, its electrical conductivity is high, so the extra current takes hundreds of milliseconds or more to diffuse into the central plasma. During this period, the current density in the outer plasma is elevated. It is expected that in the future it will be possible to maintain the required reversed shear profile with active current profile control using injected beams of energetic particles or injected waves. This confinement mode may lead to fusion power plants which are physically smaller and cheaper than had been previously envisioned, but much work remains to be done in understanding how to control the mode, and in how to fuel, and remove helium ash from it.

### Plasma Heating Methods

Alongside these improvements in energy confinement, equally significant progress occurred in technologies for heating the confined plasma. There are four general ways of heating a magnetically confined plasma. One is with the fusion reactions themselves. When a deuterium nucleus and a tritium nucleus

fuse, 50% of the released energy is immediately carried out of the plasma by a neutron which, being electrically neutral, is not confined by the magnetic field that holds the plasma. In an actual fusion power plant, the neutron will be captured in a special blanket where it produces heat for electricity production and new tritium for fuel. The remaining 20% of the energy is borne by the nucleus of a helium atom, called for historical reasons an alpha particle. Since this nucleus does not have any negative electrons bound to it, it carries a net positive electrical charge, and is confined by the magnetic fields in the plasma.

Consequently, the alpha particles produced by fusion reactions remain in the plasma, giving up energy to the plasma particles through collisions until the alphas are thermalized, that is, until they cool down to the temperature of the plasma. The energy transferred from the alpha particles, which are initially much hotter than the plasma, heats it. In an ignited fusion reactor, such as will one day be used to produce electricity, this will be the principal source of plasma heating, in fact the definition of ignition is that the alphas supply sufficient energy to maintain the plasma temperature. However, some other technique will be needed to heat the plasma of even a future electricity-producing reactor to the ignition conditions under which alpha particle heating can take over, and, in any event, the experimental tokamaks in operation today are of too modest a size and capability to reach ignition conditions.

The other three types of plasma heating technologies are called ohmic heating, wave heating, and neutral beam heating. Of these, ohmic heating is the most readily implemented, since it arises automatically from the electric current that flows through the plasma to maintain the poloidal component of the magnetic field which confines the plasma. In much the same way that an electric current flowing through a copper wire heats it as the electrons carrying the current scatter as they move along it, the current moving through the tokamak plasma also heats it.

There is, however, an important difference in how an ordinary electrical conductor and a plasma behave when heated. In an ordinary conductor, such as a copper wire, the electrical resistance rises as the wire's temperature increases, which is to say that the scattering of the electrons increases. Thus, in an ordinary conductor, if the current passing through it is kept constant by raising the driving voltage as the conductor's temperature increases, then the current will become progressively more effective in heating it until the wire eventually melts. In the sorts of plasmas of relevance to fusion, the behavior is just the opposite. The frequency of scattering decreases as the temperature increases, so the electrical resistance drops. This has the consequence that, as the plasma gets hotter, more and more current must be run through it to achieve smaller and smaller additional increases in the temperature. It would be extremely difficult to heat a plasma to thermonuclear ignition temperatures in this way, and even if one could handle the large current required, the poloidal field it would produce might be greater than the optimum for maintaining plasma stability and confinement. In light of these limitations, ohmic heating serves only as the initial heating mechanism in tokamaks, typically raising the temperature to 10 million to 20 million °C, modest by fusion standards, during the startup phase of a tokamak pulse.

One or both of the other two heating techniques—waves or neutral beams—must be used to further raise the temperature to the point where significant numbers of fusion reactions can

occur. Wave heating works in much the same way that a microwave oven does, except that instead of heating food by causing molecules to vibrate, the waves increase the energy of the ions or electrons gyrating along the magnetic field within the tokamak plasma. Various forms of wave heating have been used in many tokamak experiments, and it is expected that wave heating will be important for future reactors. However, so far the highest temperature and fusion power results in tokamaks have been brought about with the other technology: neutral beams.

**Neutral Beam Injection.** Most neutral beams in use today are born in ion sources where the negatively charged electrons are stripped from the positively charged atomic nuclei to produce ions. These ions are then accelerated by passing them between grids with different electric voltages applied to them. This forms a beam of energetic ions moving toward the tokamak in much the same way that the electron source in the back of a television picture tube forms a beam of energetic electrons moving toward the phosphor screen to form the image. The ions in a beam at Princeton University's TFTR tokamak, however, have more energy than the electrons in the beam of a picture tube—120,000 V instead of a few tens of thousands of volts. In addition, the electric current in the ion beam is thousands of times greater than the current in a picture tube.

The electrically charged ions that emerge from the accelerator grids would not be able to enter the tokamak plasma in their charged state. This is because the very magnetic field which confines the plasma would bend the ion beam out of its path and prevent it from entering the plasma. To circumvent this difficulty, the ion beam is passed through a neutralizer cell filled with low pressure gas, where a portion of the ions each pick up an electron from the gas molecules. After picking up an electron, a beam ion becomes electrically neutral, and thus once again an ordinary atom. Unlike an ordinary atom, which at room temperature has an energy of about one fortieth of a volt, these atoms in what is now a neutral beam have energies of 120,000 V. The remaining ions in the beams are bent out of it with a magnet, and then the purely neutral beam is able to pass unimpeded across the tokamak's magnetic fields to enter the plasma. Once inside the plasma, the neutral atoms of the beam are again ionized through collisions with the plasma particles which detach the electrons from the beam atoms. The resulting 120,000 V ions, being electrically charged, are confined by the magnetic field of the tokamak. They circulate along the magnetic field, colliding with the plasma particles. Since the beam ions are much more energetic than the plasma electrons and ions, they transfer energy to them, and thereby heat the plasma. Eventually the beam ions slow down and become part of the bulk plasma. At this point they are said to have been *thermalized* because their energy is similar to that of the bulk, or thermal ions.

In 1973, when neutral beams were selected as the principal heating technique for the Tokamak Fusion Test Reactor, this technology had reached the point of development where it could inject a few tens of thousands of watts of power into a tokamak plasma. In the experiments which took place at Princeton University in the 1980s, the neutral beams injected a maximum of forty million watts of power into the Tokamak Fusion Test Reactor, an increase of about a thousand fold in the power capability of the technology from the time when the decision was made to use it.

The fact that the beam ions do become part of the confined thermal plasma means that it is important not only that they carry an energy to heat the plasma, but also that they be the right hydrogen isotopes to fuel it as well. Thus, the TFTR plasma heating systems inject high-energy neutral beams of both tritium and deuterium in order to maintain the correct fuel mix in the reacting core of the plasma.

As the plasmas in successive generations of tokamaks become larger and denser, there is a corresponding increase in the beam energy required to ensure that most of the energy and fuel are deposited in the central plasma. However, the efficiency with which positive ions can be converted back to neutral atoms is a strong function of the ion velocity. For velocities corresponding to a beam energy above 120 keV for deuterium or 180 keV for tritium, the neutralization efficiency is steeply declining into unacceptably low values.

In response to this limitation, a new technology is being developed based upon ion sources which produce negative ions, or ions which have one more electron than the neutral atom would have. Production of negative ions of deuterium or tritium is much more difficult than is producing their positive ions, but the neutralization efficiency of high-energy negative ions is nearly independent of energy across a range of several million electron volts, with neutralization efficiencies of 58% to 69% being quite feasible. The first generation of a negative ion beam system began operating on the JT-60U tokamak at Naka, Japan, in 1996. When it reaches full power it will inject 10 megawatts of 500 keV neutrals. Another large negative ion system will go into operation in the late 1990s on the Large Helical Device, a type of stellarator, outside Nagoya.

**Current Drive.** All of the tokamaks which have been operated through the mid-1990s normally drove most or all of their plasma current inductively. However, as discussed earlier, this technique necessarily requires that the tokamak plasmas operate in pulses. In order to run future tokamaks steady state or in very long pulses, other techniques must be used to drive the current. Over the years, a number of techniques have been demonstrated to drive substantial amounts of plasma current, or even all of it in some special cases. The methods have used either the high-energy neutral beams which also heat the plasma, or waves which either transfer momentum preferentially to one of the confined species in one direction along the magnetic field lines or, alternatively, waves which preferentially heat the electrons or ions in the direction perpendicular to the magnetic field, thus changing their collisionality. The different current drive methods lead to drive current in different regions of the plasma, meaning that appropriate combinations of them can modify the shape of the current profile to achieve better stability and less energy leakage.

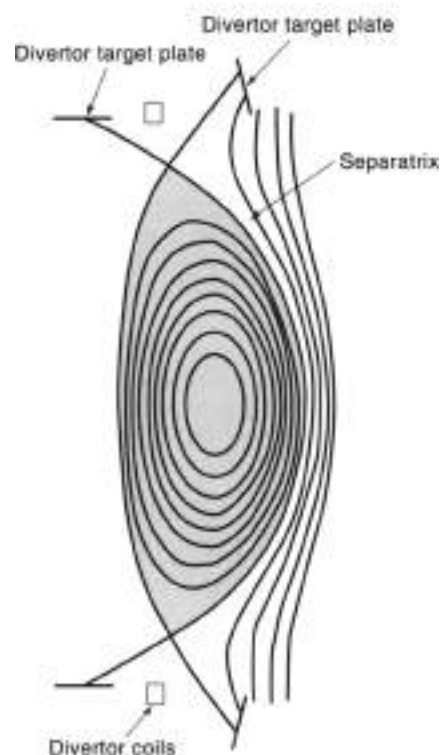
**Bootstrap Current.** All of the continuous direct current drive techniques described above have in common that their efficiency for driving current is significantly less than the efficiency of neutral beams or some sorts of waves in heating the plasma. Thus, it would be more appealing if much of the plasma current could be driven by the thermal energy of the plasma, rather than being driven by a less efficient externally applied drive. A thermal energy drive would be fed primarily by the efficient neutral beam or wave systems heating a sub-ignited plasma, or by the energy of the fusion-produced alpha particles in an ignited plasma.

A current drive mechanism arising from the bulk thermal plasma was theoretically predicted in the 1970s, and was first identified in the early 1960s on the octupole (a type of magnetic confinement device) at the University of Wisconsin. It was named the *bootstrap current* because it was generated by the plasma itself through a sort of dynamo effect as particles press against magnetic field lines. The strength of this effect depended upon the plasma pressure gradient. Thus, it was not until 1986 that bootstrap current was found experimentally in a tokamak when high confinement plasmas with steep gradients were produced on the TFTR tokamak at Princeton. The existence of the bootstrap current was subsequently verified on the other major tokamaks of the world, and it is now expected to supply much of the current for future tokamaks. One or more of the other direct-current drive techniques may also be needed to provide lesser amounts of localized current drive to produce the optimum profile for confinement and stability.

Because the strength of the total magnetic field in a tokamak increases sharply as one progresses towards its major radius, particles with sufficiently small ratios of parallel velocity to perpendicular velocity with respect to the field will become trapped in local magnetic mirrors, much as happens to the main plasma in mirror confinement machines. These particles are referred to as *trapped particles* in distinction from the *passing particles* which circulate freely along field lines around the tokamak. Their orbits, when projected onto a plasma cross section, resemble bananas, with the tips corresponding to mirror reflection from higher magnetic field strengths. In the Soviet Union, these were sometimes referred to as *sickle orbits*. In lower temperature collisional plasmas, the trapped particles often scatter out of trapped velocity space before they complete a mirror bounce, and thus have relatively little effect on the plasma behavior. At the high temperatures typical of today's tokamaks, the particle collisionality is much reduced, and trapped particles may complete many bounces. It is these trapped particles which are particularly important in the production of bootstrap current. Trapped particles in a tokamak are an effect which arises from the toroidicity, since this is what creates the higher magnetic field at smaller major radius. Thus, the fraction of the total particles which are trapped is proportional to  $r/R$ , ratio of the minor radius to the major radius, which is the inverse of the aspect ratio. Thus, in order to maximize bootstrap current, which is in some sense free, reactor designs are sometimes driven toward lower aspect ratios, in the general vicinity of about 3 or less. However, this is somewhat offset by the fact that the fraction of the current that can be driven by bootstrap effects is in part determined by  $(R/r)^{3/2}$ . In addition, there are also other constraints, such as the scaling of confinement in some plasma regimes, which may drive the optimum aspect ratio to larger values (6,7).

Comprehensive design studies of a number of variants of tokamak reactors for power plants have been carried out at the University of California at Los Angeles, resulting in a number of designs dubbed with the name *Arise* (8). These for the most part anticipate providing much of the plasma current with the bootstrap effect, and the balance with one or more of the direct-current drive techniques.

**Plasma Exhaust System.** A long pulse or steady-state tokamak reactor requires something better than a simple limiter to handle the power coming out the edge of the plasma, and to dispose of the helium ash which results from fusion reactions.



**Figure 5.** Example of tokamak plasma with top and bottom diverters. The divertor coils produce open flux surfaces after the separatrix field line, allowing the outer plasma to leave the main chamber to strike target plates. This reduces the influx of impurities to the main plasma, allows the power density striking the target plates to be reduced through the expansion of flux tubes, and may allow further power density reductions through radiation in the divertor region.

This is done by altering the closed field line topology of the normal tokamak configuration. By placing an additional shaping coil in the vicinity of the plasma top, bottom, or in both locations, it is possible to redirect the outer field lines so that a thin layer of the plasma on the edge of the plasma cross section is diverted out of the main plasma chamber into a divertor chamber where the plasma, and the helium ash it entrains, are converted to neutrals through a combination of encountering a region of higher neutral particle density and of directly striking plates of carbon or some other material. The divertor region is baffled from the main chamber to reduce reflux, and much of the neutralized plasma outflow can be pumped away as gas. Figure 5 displays one possible divertor configuration.

Ensuring that the divertor plates have an economically attractive lifetime is a significant engineering problem which is not yet fully solved for fusion power reactors: the power density in the scrapeoff plasma flowing toward the divertor will be high, causing thermal problems, and the energetic particle flux can cause high sputtering rates on the divertor plates. Some of the approaches being pursued to ameliorate these problems include reducing the incident power density by reducing the angle of incidence, and by expanding the flux envelope of the diverted plasma, and reducing the sputtering rate by introducing strongly radiating heavier impurities such as noble gases into the divertor chamber plasma to lower the temperature of the ions so they will produce less sputtering. This deliberately enhanced radiation can also be used to disperse energy at reduced intensity over a larger expanse of material.

**Tritium Breeding and Thermal Conversion Blanket.** About 80% of the energy produced in a deuterium-tritium fusion reactor immediately leaves the plasma in the form of 14.1 MeV neutrons, since plasmas with dimensions and densities appropriate to reactors are almost perfectly transparent to high-energy neutrons. The remaining 20% of the energy carried by the fusion alpha particles heats the plasma, and eventually either leaks from the plasma as electromagnetic radiation that will be absorbed and converted to heat by the first wall, or it flows in the scrapeoff plasma to the divertor, where it heats surfaces. Thus, all of the alpha particle energy is ultimately collected on surfaces. These surfaces will be cooled with some fluid or gas, of which lithium or helium are examples of possible choices. This extracted heat will then be used to produce electricity, probably by generating steam to drive a conventional turbine.

The neutrons, which carry the other 80% of the fusion power, will be slowed down and captured in a blanket surrounding the tokamak. In the process, their kinetic energy will be converted to heat, which will be further augmented by exothermic nuclear reactions with  $^6\text{Li}$ . This heat can then be removed from the blanket to produce electricity, either through steam turbines, or possibly by more efficient techniques taking advantage of the fact that the blanket may contain a liquid metal.

This blanket must also breed enough tritium to at least replace the amount consumed in nuclear reactions by the plasma, and it is preferable that it produce somewhat more in order to start up additional fusion power plants. The reactions of neutrons with lithium isotopes which produce tritium were discussed in an earlier section. Proposed blanket designs have of necessity included either natural lithium or  $^7\text{Li}$ , but they have differed in what neutron multipliers were added to the mix. The most commonly assumed multiplier is beryllium, which undergoes an  $(n, 2n)$  reaction, but heavy multipliers such as lead have also been considered. The chief difficulty with some heavy multipliers is that they tend to produce radioactive daughter nuclei with undesirable half-lives. Beryllium does not pose this difficulty.

**Maintenance and Materials.** Developing techniques to maintain and repair fusion power plants is an area which still requires substantial engineering development. It may be somewhat simplified by the time commercial plants begin operating some decades from now by advances in robotics. Present tokamaks are commonly made of stainless steel alloys or inconel, primarily because they are relatively inexpensive, and can be machined and formed by drawing upon a large body of fabrication experience. These materials, however, would stay significantly radioactive for many years after removal from a fusion plant, meaning they would require storage. They would not be volatile or very reactive, so there would be little threat of contaminating water or air.

However, one of the advantages of fusion relative to fission is that one has some degree of freedom to choose materials which reduce activation. This is not possible with fission, since the nuclear waste products are produced by the fissioning of the fuel itself. Accordingly, fusion power plants probably will use different materials which will have lower initial levels of radioactivity per gigawatt year of exposure, and shorter half-lives which decrease the storage time required for removed components. Two of the low activation materials which have been considered by studies of the Aries group (5) include a vanadium alloy with small amounts of chromium and titanium, and silicon carbide. In their reactor models, they found that the

total radioactivity per watt of reactor thermal power remaining in a fusion power plant one year after shutdown would be a factor of  $10^2$  lower than for a comparably rated fission breeder reactor, if the fusion plant was built of HT-9, a ferritic steel. If the plant were instead constructed of the vanadium alloy, the one year radioactivity was depressed to a level about  $10^5$  less than for a fission breeder, and if silicon carbide was used for the fusion reactor, then the one-year radioactivity of the fission reactor declined further to a value about  $10^8$  less than the fission breeder. If  $\text{D}^3\text{He}$  reactors ever proved feasible through lunar mining, then the silicon carbide design would give a one-year radioactivity that would be more than  $10^{10}$  less than for a fission breeder.

These and other low-activation materials require engineering validation through tests exposing them to large neutron fluences. In addition, some materials, especially silicon carbide, require progress in fabrication technology. The structures normally made of silicon carbide are significantly smaller than would be required for a reactor.

## INERTIAL CONFINEMENT

### Parameter Regime

A practical fusion power plant fusing deuterium-tritium fuel requires temperatures of over 10 keV, with a fuel density,  $\rho$ , and an energy confinement time,  $\tau$ , such that their product lies in the range of  $\rho\tau = 10^{14}$  to  $10^{25}$  s/cm $^3$ . A practical fusion reactor using the magnetic confinement approach discussed earlier will likely operate in a regime with a density of about  $10^{14}$  fuel nuclei/cm $^3$ , and a confinement time of about a second or so. Inertial confinement operates at the other extreme of the parameter range. Because the small nuclear explosions in an inertial confinement reactor would occur with high-velocity fuel particles traversing very small distances, the confinement time would be of the order of  $10^{-12}$  s, requiring that the target fuel be compressed to about  $10^{25}$  fuel particles/cm $^3$ .

This compression must be carried out by a strong implosion. A typical inertial confinement target capsule is a small sphere formed of an ablator material which is lined with solid deuterium-tritium fuel. The central spherical cavity thus formed, which composes most of the target capsule's volume, is filled with deuterium-tritium gas. The implosion occurs when the ablation shell is illuminated with a brief burst of extremely high-energy density, on the order of  $10^{16}$  W/cm $^2$ . This causes the outer portions of the ablation shell to leave it with high momentum which, on the average, will be directed radially outward from the shell. Conservation of momentum requires that an equal inward-directed momentum be imparted to the remaining target, driving the collapse. For a reactor to have energy gains of economic interest, it is also necessary that this implosion proceed isentropically, with as little preheating ahead of the compression wave as possible. This, in turn, requires that the compression be uniform to a level of about 1% over the entire sphere. Figure 6 shows one possible design for a target capsule.

### Direct and Indirect Drive

The most obvious way to compress a target sphere is to shine the driver energy directly onto the ablation sphere. The driver might be an array of lasers, light ion accelerators, or heavy ion accelerators. One of the principal challenges in implementing

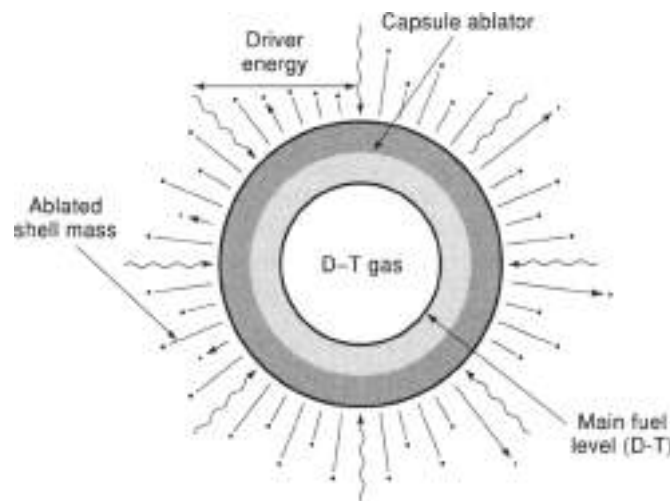


Figure 6. Target capsule for inertial confinement fusion are compressed by ablation from the outer shell. After ignition in the center, the fusion burn region will need to spread outward to produce significant energy gain.

this approach is that obtaining the required high degree of uniformity in the driver power at the sphere has proven very difficult. In addition, with some types of drivers there may be pre-heating of the fuel due to penetration by driver ions, and laser-driven instabilities may degrade the symmetry of the compression. Ways of ameliorating these problems are being studied.

The other implosion method is indirect drive. In this approach the fuel capsule is not struck by the primary driver beams (4). Instead, the capsule is placed inside a Hohlraum which is of considerably larger dimensions than the capsule, as depicted in Fig. 7. This Hohlraum is made of a high atomic number material, and has openings through which the driver beams can enter so as to directly strike the inner wall of the Hohlraum, but not the fuel capsule. A portion of the driver energy striking the Hohlraum is converted to soft X rays, and they in turn drive the implosion of the fuel capsule. An advantage of this method compared to direct drive is that quite uniform irradiation of the capsule with the soft X rays can be achieved even with anisotropic primary driver beams, enabling fewer beams to be used. In addition, there is a large body of

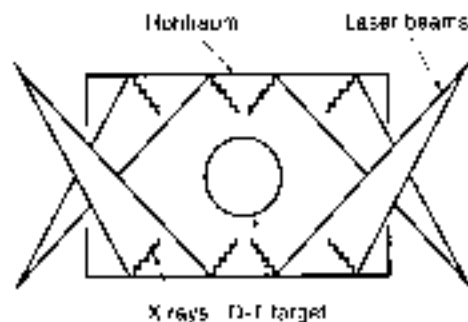


Figure 7. Target placement for indirect drive inertial confinement fusion. High-power laser beams strike the inner surfaces of the Hohlraum, producing X rays which compress the target capsule. A variant of this would use heavy ion beams striking absorbers at the ends of the Hohlraum. Indirect drive requires much less perfect symmetry in the driver than does direct drive.

experience available with this sort of indirect drive from the nuclear weapons program. The majority of research being conducted on inertial confinement in the US today follows the indirect drive approach.

### Inertial Confinement Drivers

**Laser Drivers.** Over recent decades, a number of classes of drivers have been considered, and in some cases tested, for inertial confinement applications. Experiments have been driven by neodymium-glass lasers, and krypton fluoride gas lasers have been studied as possible primary drives. Most of the laser drivers suffer from low efficiencies for conversion of electricity into laser light. In the case of KrF this efficiency is about 6% to 8% (9). This would require that the target capsule produce 140 to 160 times more energy than was in the driver in order to make a power plant feasible. Such high gains are vastly beyond current achievements, but may, in principle, be achievable. Laser drives have much higher efficiencies of as much as 60%. If the prices of these can be reduced by a large factor, and if techniques can be found to interface them with the thermonuclear environment of a reactor cell, then these may become attractive.

**Ion Drivers.** The other classes of drivers are light ions and heavy ions. In the light ion approach, extremely high currents of lithium ions are produced by an array of diode sources and accelerated to energies of hundreds of keV to a few MeV. Lithium diode drivers have been studied at the Sandia National Laboratory. The other method would use lower currents of ions from the high mass end of the periodic table such as xenon or cesium, which would be accelerated to energies of 2 GeV to 10 GeV. About  $10^7$  A of beam at these energies would be required to achieve the necessary illumination intensity. Although the required pulse duration is very brief, such currents are well beyond the capabilities of existing high-energy accelerators. High-current accelerator concepts such as inductive linacs are being studied.

Both of the ion drive techniques have the major advantage, relative to most of the laser approaches, that their efficiency for conversion of electricity to ion energy can be quite high, 25% to 30%. With the heavy ion approach a target energy gain of a factor of 40 might be sufficient for a feasible plant, which is roughly a factor of 4 below that required for a KrF laser driver. On the other hand, for either technology the final path of approach of the ion beams toward the Hohlraum will be after the final focusing elements of the ion optical system. Maintaining the necessary illumination intensity and uniformity will be difficult as these high space charge density beams converge toward even higher concentrations of space charge. Space charge effects, which arise because of the coulomb repulsion between like charges, are strongly defocusing.

### Inertial Confinement Plant Systems

An inertial confinement fusion power plant would consist of a target chamber, the driver system of ion accelerators or lasers, and a target fabrication system introducing several capsules (and Hohlraums if needed) per second into the chamber. A lithium layer would breed tritium and would convert the fusion energy to heat, which could then be converted to steam to drive turbines. Concepts for this lithium layer have included liquid metal flowing over the inner surface of the target chamber, as well as jets of liquid metal around the target.

## PROGRESS IN FUSION ENERGY

Over the past quarter century, large advances have been made in understanding the physics of plasmas suitable for a magnetically confined reactor, and the technologies for heating the plasmas have similarly progressed. In this period the power released through fusion reactions has increased a factor of  $10^4$ , with over 10 MW produced by the TFTR tokamak funded by the US Department of Energy at Princeton University. Similarly the energy leakage of the ions has dropped from being many times higher than that theoretically predicted to values which in some cases are at or near the theoretically best confinement that could ever occur. Temperatures in tokamaks have climbed from a few hundred electron volts to as much as 40 keV to 45 keV in TFTR and in JT-60U in Naka, Japan. Energy confinement times have grown from a few milliseconds to a second or more on Japan's JT-60U and Europe's JET tokamaks. The power of neutral beam systems heating the plasma has climbed from the tens of kilowatts prevalent a quarter century ago to 40 MW on TFTR and JT-60U.

Inertial confinement fusion has made advances in understanding the physics of matter at the very high pressures and densities required for this approach, and understanding has been gained about the ways in which energy and momentum are transferred to the target capsule. A major new laser-driven test bed, named the National Ignition Facility, is to be constructed to further these studies, and investigations continue on ways to make more efficient drivers.

Producing an economically attractive fusion power plant will require more work for either the magnetic or inertial confinement approaches. Based upon the present state of these fields, it appears that the magnetic confinement approach enjoys a clearer path forward to a reactor, but this could change in the future depending upon progress in inertial confinement.

## BIBLIOGRAPHY

1. J. Weiser, *Tokamaks*. Oxford: Clarendon Press, 1987.
2. J. G. Cardy, R. J. Goldston, and R. R. Parker, Progress toward a Tokamak fusion reactor. *Phys Today*, 22-30, 1992.
3. L. J. Wittenberg, J. F. Santarius, and G. L. Kulcinski, Laser source of  $^3\text{He}$  for commercial fusion power, *Fusion Technol.* 10, 367, 1986.
4. J. Lindl, Development of the indirect drive approach to inertial confinement fusion and the target physics basis for ignition and gain. *Phys Plasmas*, 2, 111: 3033-3024, 1995.
5. H. J. Goldston and P. H. Rutherford, *Introduction to Plasma Physics*. Bristol and Philadelphia: Institute of Physics, 1995.
6. L. R. Grisham et al., The scaling of confinement with major radius in TFTR. *Phys Rev Lett.* 67: 66-69, 1991.
7. L. R. Grisham et al., Scaling of ohmic energy confinement with major radius in the Tokamak fusion test reactor. *Phys Plasmas*, 1, 3396-4001, 1994.
8. R. W. Conn et al., Economic, safety, and environmental aspects of fusion reactors, *Nucl. Fusion*, 30: 1919-1934, 1990.
9. W. J. Hogan, R. Bongester, and G. L. Kulcinski, Energy from inertial fusion. *Phys Today* September, pp. 42-50, 1992.

LARRY R. GRISHAM  
Princeton University

## GYROTRON

The gyrotron is a vacuum tube capable of delivering high levels of radio-frequency (RF) power at frequencies from several gigahertz to more than 300 GHz, which covers most of the microwave and millimeter wave band. Because of the structure of the RF fields in the cavity, the magnitude of the electric fields and the RF losses in the cavity walls are much lower in gyrotrons than in most microwave and millimeter wave vacuum tubes. As the RF increases, it is not necessary to reduce the size of the cavity and output waveguide. For these reasons, the gyrotron is the principal RF device for delivering high levels of RF power at frequencies between 30 GHz and 200 GHz.

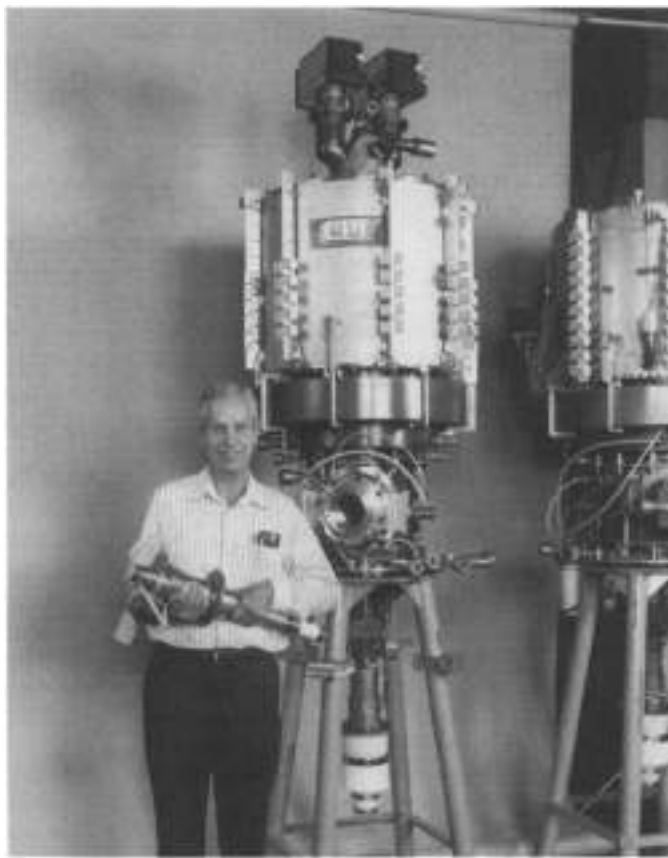
Gyrotrons typically require a very high magnetic field to provide the cyclotron electron motion for power extraction, and most gyrotrons above 30 GHz require a superconducting magnet. This significantly increases system cost and complexity, limiting applications for these devices, however, recent advances in magnet technology are reducing this problem. Gyrotrons also have a narrow RF bandwidth, typically less than 1%, although gyrotron amplifiers are capable of wider bandwidths. Gyro-TWTs with 20% bandwidth have been demonstrated, and octave bandwidths are predicted with more complex circuits. Gyrotron oscillators produce more average or CW power levels since gyro-amplifiers are much more difficult to develop at higher frequencies due to the overmoded nature of the circuits. Still, gyro-amplifiers have significantly more capability than conventional tubes at millimeter-wave frequencies, and development of high-power gyroamplifiers is in progress at several locations around the world for use as high-power RF sources for accelerators and high-resolution radar.

Gyrotrons are used for electron cyclotron resonance heating (ECRH), electron cyclotron current drive (ECCD), and diagnostic measurements in fusion plasma devices. They are also being used for industrial heating applications such as ceramic sintering. Figure 1 is a photograph of Dr. Howard Jory, one of the pioneers in gyrotron development, holding a 28 GHz, 10 kW continuous wave (CW) harmonic gyrotron used for industrial heating. Behind him is a 110 GHz gyrotron rated at approximately 450 kW CW and 1 MW for pulses less than 1 s, that is used for electron cyclotron resonance heating. Communications and Power Industries, Inc. in Palo Alto, California manufactures both devices.

Research on gyro-type devices began in the 1950s when the astrophysicist R. Q. Twiss described an amplifying mechanism for monochromatic radiation of angular frequency  $\omega$  from stimulated emission of an ensemble of electrons (1). Twiss' formula predicted amplification for Cerenkov radiation and for cyclotron radiation. Working independently, Schuender described the stimulated emission of radiation from electrons in a magnetic field in 1959 using a quantum mechanical model (2). Also in 1959, Gaponov described this mechanism using a waveguide model (3).

The first experimental results describing a fast wave cyclotron interaction were reported by R. H. Pantell in 1959 (4). His device produced radiation between 2.5 and 4.0 GHz from a 1 kV, 3  $\mu\text{A}$  electron beam. In the former Soviet Union, Gaponov and others were performing experiments in 1959 using a helical beam in a longitudinal magnetic field, but the efficiency was low. In 1963, M. Filulin presented the first theory involving helical beams at a conference in the Ukraine, which led to an





**Figure 1.** Dr. Howard Jury holds a 28 GHz, 10 kW CW harmonic gyrottron used for industrial heating. Behind him is a 110 GHz, 450 kW CW, 1 MW pulsed, gyrottron, used for electron cyclotron heating of fusion plasmas.

experiment in 1964 that produced 200 W. In 1966, the power level reached 1 kW. A number of additional experiments were reported during the early 1960s (5-8), but the experiment that confirmed the cyclotron maser interaction was performed by J. L. Hirschfield and J. M. Wachtel in 1964 (9).

During the mid- to late 1960s, major advances were made by scientists in the Soviet Union, though much of this work went unreported in the Western World due to the political climate at the time (10). Soviet scientists embarked on a major program to develop megawatt gyrottrons and made major advances in electron mode selection, open resonators, and waveguides (11), ray tracing to optimize the interaction between the electron beam and the RF wave in the circuit, and large cross-section cavities utilizing whispering gallery modes (12). Whispering gallery modes are transverse electric modes where the number of azimuthal field variations significantly exceeds the number of radial field variations. Use of these modes was further facilitated by the development by S. N. Vlasov et al. of a quasi-optical device for converting the whispering gallery waveguide mode into a Gaussian mode that can be propagated in a narrow wave beam without waveguides (13). An equally important Soviet development was the magnetron injection gun, which generates the required electron beam for efficient gyrottron operation.

For a more complete historical description of gyrottron development, the reader is referred to three publications that cover the subject in more detail (14-16). The reader is also referred

to a listing of journals in the bibliography that contain most of the published work in this area.

## BASIC THEORY OF OPERATION

Gyrottrons exploit the negative mass instability to obtain azimuthal bunching of a cyclotroning electron beam. A transverse electric field in the cavity modifies the energy of the electrons such that higher-energy electrons gyrate more slowly around the magnetic flux lines than do lower-energy electrons. A schematic cross section of a gyrottron beam is shown in Fig. 2. The beam is hollow with individual electrons rotating at the cyclotron frequency around magnetic field lines with orbit diameters equal to twice the Larmor radius. The Larmor radius is given by

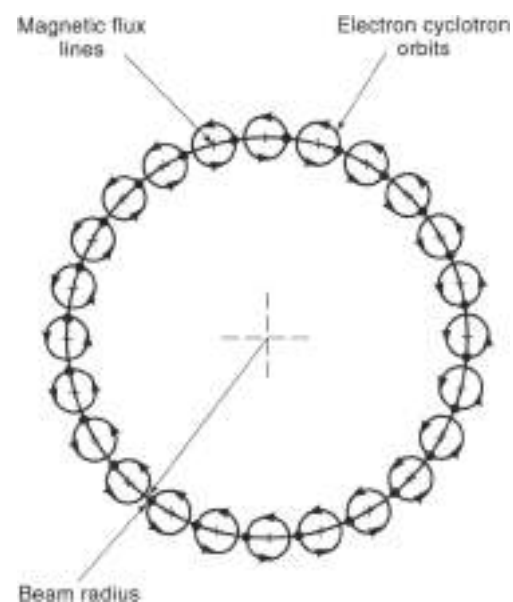
$$R_L = \frac{\gamma m v_\perp}{e B_0}$$

where  $\gamma = (1 - v^2/c^2 - v_\perp^2/c^2)^{-1/2}$  is the relativistic factor,  $m$  is the rest mass of the electron,  $v_\perp$  is the velocity of the electron in the plane perpendicular to the direction of the magnetic field,  $v_\parallel$  is the electron velocity in the direction parallel to the magnetic field,  $e$  is the charge of the electron, and  $B_0$  is the applied magnetic field strength.

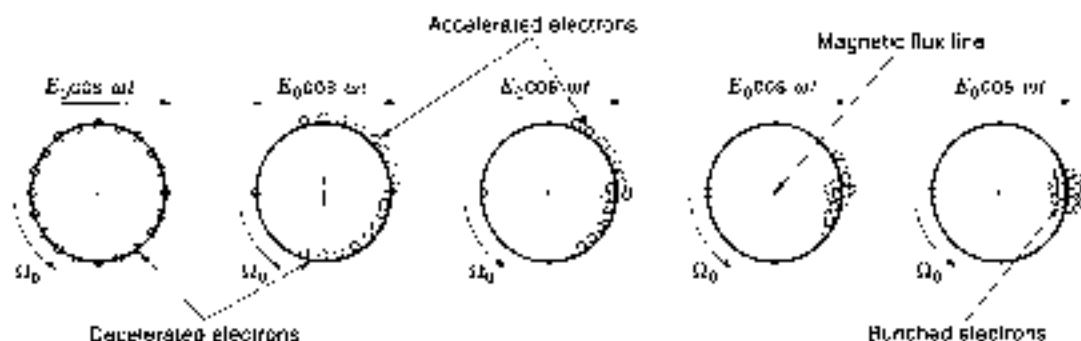
Figure 3 more closely examines the effect of the electric field on the individual electrons in a cyclotron orbit when the frequency of the electric field is equal to the cyclotron frequency. The cyclotron frequency is given by

$$\Omega = \frac{e H_0}{\gamma m c}$$

Most gyrottrons employ an interaction circuit that can be approximated by a right-cylindrical cavity. In the gyrottron interaction, only electron beam energy that is transverse to the axis



**Figure 2.** The cylindrical electron beam of the gyrottron consists of electrons orbiting around magnetic flux lines.



**Figure 3.** The cavity electric field interacts with the electrons orbiting around the magnetic flux lines. Each diagram is an integral number of cyclotron periods later in time as the electrons traverse the cavity. The force on the electrons is given by  $\vec{F} = -e\vec{E}$ .

of the cavity is extracted from the beam, so that the interaction is with the transverse (TE) modes of the cavity. For this geometry, the electric field components have the form

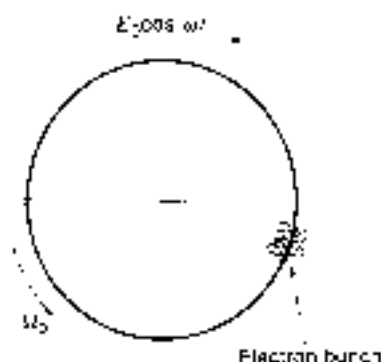
$$E_z = E_0 J_m'(k_r r) \sin k_z z \cos \omega t$$

$$E_r = \sum_{m=1}^{\infty} J_m(k_r r) \sin k_z z \cos \omega t$$

where  $E_r$  and  $E_z$  are the azimuthal and radial components of the electric field, respectively;  $J_m$  and  $J_m'$  are the Bessel function and Bessel function derivative, respectively;  $m$  is the azimuthal mode number of the electric field;  $k_r = X'_{nm}/r_0$ , where  $X'_{nm}$  is the  $n$ th root of the corresponding Bessel function derivative, and  $r_0$  is the circuit radius; and  $k_z = p\pi/L$ , where  $p$  is the axial mode number of the circuit field (typically = 1), and  $L$  is the cavity length.

In Fig. 3, each successive image is an integral number of RF periods later in time as the electrons traverse the cavity. Electrons accelerated by the electric field gain energy and, as a result of the relativistic mass increase, their angular velocity decreases as  $\gamma$  increases. Conversely, electrons decelerated by the electric field lose energy and gain angular velocity. This causes azimuthal bunching of the electrons, sometimes referred to as the *cyclotron resonance maser* (CRM) instability.

If the frequency of the electric field,  $\omega$ , exceeds the cyclotron frequency,  $\Omega$ , the bunch will eventually fall back in phase and more electrons will undergo deceleration than acceleration, as shown in Fig. 4. This will result in transfer of energy from



**Figure 4.** When  $\omega > \Omega_0$ , more electrons will be decelerated than accelerated, resulting in transfer of energy to the electric field.

the electrons to the electric field. Optimized performance of gyrotrons requires careful design of the circuit geometry and control of the magnetic field. The circuit geometry affects the frequency and choice of the mode for the electric field, as well as the strength and profile. The magnetic field affects frequency of the interaction, and the magnetic compression ratio affects the size and location of the beam as well as the transverse energy. In many cases the magnetic field is tapered; that is, it is modified in strength through the cavity, to optimize the efficiency of power extraction.

Harmonic gyrotrons operate at a multiple of the cyclotron frequency. As the harmonic number increases, beam placement becomes more critical, and the theoretical efficiency decreases. The advantage of harmonic gyrotrons is that the magnitude of the magnetic field is  $1/n$  times that required for nonharmonic operation. This can eliminate the need for a superconducting magnet.

For a more complete description of gyrotron theory, both linear and nonlinear, the reader is referred to several excellent publications on the subject (16–18).

## BASIC COMPONENTS

Figure 5 shows a schematic layout of a typical gyrotron oscillator. Basic components consist of an electron gun, input beam tunnel, circuit, output taper, collector, window, and magnets. There are numerous variations of each of these components, depending on the operating characteristics of the gyrotron. The discussion that follows describes the most common types of components, including their purpose and performance characteristics.

### Electron Gun

The function of the electron gun is to produce the electron beam required for interaction with the desired operating mode in the circuit. The gun must produce an electron beam that will be located at the proper radius for efficient interaction with the cavity electric fields and must contain most of its energy in cyclotron motion. Most gyrotrons employ a magnetron injection gun (mig) that emits electrons from a cathode placed in a region of crossed electric and magnetic fields. This geometrical configuration causes the individual electrons to spiral around the magnetic field lines as they are accelerated by the electric

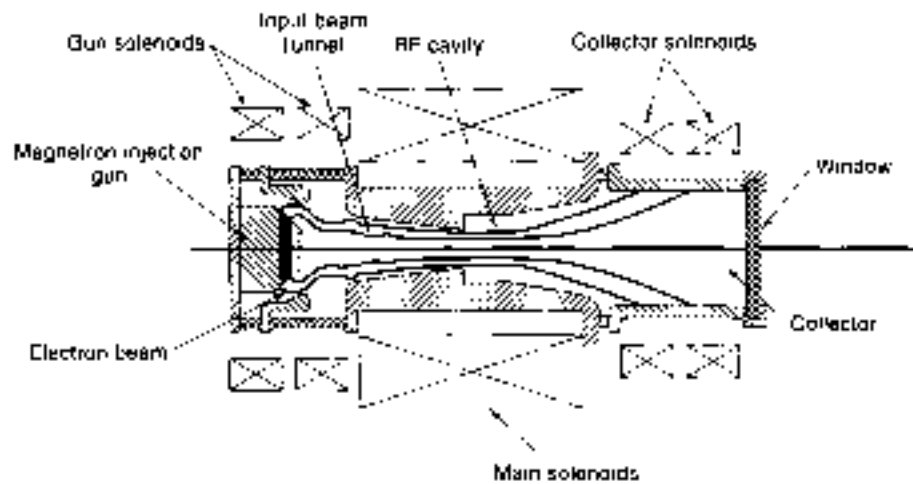


Figure 6. The basic components of a typical gyrotron are the electron gun, beam tunnel, cavity, output beam tunnel, collector, window, and magnet solenoids.

field toward the circuit. This development was pioneered in the Soviet Union in the 1960s and made rapid development of gyrotrons possible (19).

Gyrotron cathodes typically operate temperature-limited, which means that the amount of current emitted from the cathode is primarily determined by the temperature of the emitting surface. The reader is referred to the excellent treatment of cathode emission mechanisms presented by A. S. Gilmour, Jr. (20). Typical cathode temperatures are between 950 °C and 1000 °C.

Gyrotron guns come in two basic types, mod anode and diode. A mod anode gun contains an electrode near the emitting surface to modulate the electron beam, that is, it uses an applied voltage to modify the characteristics of the beam. This electrode is maintained at a voltage between the cathode and ground and can have a significant effect on the beam characteristics and, hence, the performance of the gyrotron. By varying the voltage on this electrode, the user can control the output power by varying the amount of transverse energy that is present in the electron beam. Most commercially available gyrotron oscillators using mod anode guns operate at cathode voltages between -60 kV and -80 kV with mod anode voltages ranging from 15 kV up to about 30 kV above cathode potential.

A diode gun possesses no intermediate electrode, so the cathode voltage or the magnetic field controls gyrotron operation. Because the cathode power supply must also supply the current in the beam, it cannot be modified as easily as a mod anode supply, so diode guns typically operate where RF power is turned on or off by pulsing of the cathode power supply. For applications where it is not necessary to modify the operating characteristics of the tube, this represents a lower-cost gyrotron and power supply configuration.

#### Magnetic Solenoids

The circuit requires an electron beam where the electrons are orbiting the magnetic field lines at a frequency near that of the RF frequency of the cavity. For gyrotrons operating at frequencies above 30 GHz, this typically requires a superconducting solenoid to generate the required magnetic field in the circuit. An exception can occur if the circuit is operating at a harmonic of the cyclotron frequency. The requirement of a superconducting solenoid limits applications for gyrotrons to those where the additional weight and complexity do not present a severe

burden; however, advances in superconducting magnet design and technology are rapidly reducing this problem. Gyrotrons are now being considered for mobile or airborne applications that were previously considered impractical.

In addition to the solenoids required to produce the cavity magnetic field, other solenoids can be present around the electron gun and collector regions. Solenoids around the electron gun modify the amount of transverse energy in the beam or its size in the circuit. These coils are typically operated at room temperature because the fields in the electron gun are typically a few hundred gauss. The ratio of strength of the magnetic field in the circuit to that in the electron gun is referred to as the *beam compression*. This value plays an important role in the design of the electron gun and the combination of electric and magnetic fields required for efficient circuit interaction. Beam compressions of 10 to 20 are typical.

Solenoids around the collector are used to distribute the power deposited by the electron beam exiting from the circuit. The fringing magnetic fields from superconducting coils impose a magnetic field in the collector that is higher than in other linear beam tubes. This field prevents the electrons from spreading due to space charge and results in a relatively thin beam of electrons impacting a small region of the collector. This can result in excessive localized heating of the surface and possible destruction of the gyrotron. The collector coils are used to "buck" the field from the main solenoid and spread the electron impact area over a larger region of the collector. In some more advanced applications, an oscillating power supply drives one or more solenoids to sweep the beam back and forth along the collector surface.

#### Beam Tunnel

The beam tunnel must perform two important functions. First, it must prevent RF fields from the circuit from traveling back toward the electron gun. RF fields in the electron gun can severely affect the electron beam and cause excessive current on the body of the tube or the mod anode. It can also result in heating of the cathode emitter, which will affect the beam current. Either condition will adversely affect tube operation.

The second function of the beam tunnel is to prevent parasitic oscillations between gun and circuit. Parasitic oscillations result in excessive body or mod anode current and can prevent tube operation. In some cases, dielectric material

with a high RF loss is used to load out any electromagnetic fields present.

### Circuit

The RF circuit typically consists of a right circular cylinder whose radius  $a_c$  is chosen such that

$$a_c = \frac{cX'_{n0}}{2\pi f}$$

where  $X'_{n0}$  is the  $n$ th root of the Bessel function derivative for the  $TE_{n0}$  cavity mode,  $c$  is the speed of light, and  $f$  is the operating frequency. The gyrotrom operates near the cut-off frequency of the circuit, and the equation for the circuit radius is the equation for the cut-off frequency for a waveguide of radius  $a_c$ . The cavity mode depends primarily on the desired RF frequency and output power level. The application for the tube can also play an important role. The first commercial gyrotroms used  $TE_{01}$ ,  $TE_{10}$ , or  $TE_{20}$  modes because they were reasonably close to the fundamental mode, did not couple to nonsymmetric modes that could travel down the beam tunnel toward the electron gun, and had very low RF loss in the circuit walls. More than 100 gyrotroms at frequencies between 28 GHz and 100 GHz and power levels up to 340 kW CW were produced by Varian Associates, Inc. (now Communications and Power Industries, Inc.) from 1978 through 1990. These tubes were used for electron cyclotron resonance heating (ECRH) of plasmas in magnetic confinement fusion research. Gyrotroms of this type have also been developed in Germany, China, Japan, France, and Russia (21).

As the demand for higher-power tubes increased, gyrotroms that used whispering gallery modes were developed. Whispering gallery modes allow the electron beam to be larger and minimize mode competition from other circuit modes. These modes also have reasonable RF power dissipation in the circuit. This work was pioneered in the former Soviet Union beginning in 1968.

Designers of high-power gyrotroms must balance a large number of factors to achieve a circuit configuration that will provide the power and frequency required with reasonable efficiency while avoiding parasitic mode competition, instabilities, and excessive RF power densities on the circuit walls. For additional information on the design of high-power gyrotrom circuits, the reader is referred to Ref. 22.

### Output Taper

Most all gyrotrom circuits are open-ended toward the collector, and the RF power diffracts into the output taper. The output taper is usually a shallow, tapered section of waveguide that terminates the circuit interaction and transmits the RF power toward the collector. The waveguide radial dimensions must be increased in such a way that the purity of the output mode is not compromised. For gyrotroms where the RF power is extracted along the axis of the tube, the output taper continues to increase the radial dimension until the desired collector radius is achieved. The requirement to maintain mode purity often conflicts with the necessity of achieving the collector radius within a reasonable distance. The size of the collector is driven by the necessity to dissipate the spent electron beam without incurring excessive power densities on the walls, which could lead to melting or loss of vacuum integrity.

For gyrotroms with average RF power levels exceeding 200 kW, particularly those employing whispering gallery circuit modes, it is more expedient to extract the RF power radially to allow more freedom in collector design. For these tubes, the output taper transmits the circuit power to an RF launcher for eventual extraction from the vacuum envelope. This removes further requirements on the output taper to maintain mode purity and allows the designer more freedom in transitioning to the collector region.

### Collector

Typical linear beam devices rely on termination of the magnetic field and space charge depression to spread the spent electron beam in the collector. Most of these devices employ an iron polepiece to terminate the magnetic field at the entrance to the collector. While this works well for solid electron beams, it does not apply to gyrotroms. Because of the high magnetic fields required in gyrotroms, this is often not practical, particularly when superconducting solenoids are used. Forces imposed by nearby iron on the superconducting coils would dramatically increase the complexity and cost of the magnet. As a result, sufficient magnetic fields exist in the collector to limit spreading of the electron beam from space charge forces. In addition, the beam used in gyrotroms is typically a thin cylindrical beam with high current density. This also exacerbates the problem of beam dissipation in the collector. Power densities up to 500 W/cm<sup>2</sup> are typical; 1000 W/cm<sup>2</sup> is considered the upper limit.

For gyrotroms at lower average power levels, these complications may not be a critical issue. For these devices, the output taper conducts the RF power to the collector with high mode purity, and the collector also serves as the output waveguide. In some cases, a downtaper is employed at the end of the collector to reduce the tube diameter before the output window. For these devices, the designer must also be concerned about modes that can be trapped in the collector between the uptaper from the circuit and the downtaper to the window.

For gyrotroms exceeding 500 kW of average power, radial RF power extraction allows the designer more freedom in collector design. The size of the collector can be based on the dissipator requirements of the spent beam alone. As a result, collectors for high-power gyrotroms often constitute a major portion of the device. In Fig. 1, the collector of the 500 kW CW gyrotrom behind Dr. Jory begins above the output window. The designer must balance low power densities with the increasing size, weight, and cost of large collectors.

### RF Launcher System

The RF launcher system converts the circular cavity/waveguide mode to a quasi-optical Gaussian mode. This development was pioneered in the Soviet Union and allows transmission of the beam with very low loss using a series of metallic mirrors. Modern computer codes and integration with computer numerically controlled (CNC) mills and lathes allow precise control of the power in the RF beam. Modern designs convert the power from the circuit into the desired Gaussian mode with greater than 95% efficiency. The conversion of the waveguide mode to a quasi-optical Gaussian-type mode also results in a more convenient RF mode for most user applications.

Figure 6 shows a typical RF launcher configuration consisting of a waveguide launcher and two mirrors internal to the gyrotrom. The exit angle of the RF power from the launcher is

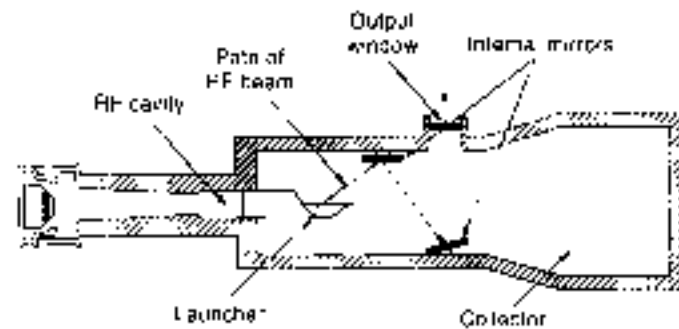


Figure 6. The Gaussian mode launcher converts the waveguide mode to a quasi optical beam, and the internal mirrors shape the RF beam profile.

at the waveguide bounce angle, and the mirrors concentrate the power into a small circular beam and tailor the distribution of power within the beam as required by the output window.

#### Output Window

A dielectric window is used to extract the RF power from the vacuum envelope of the gyrotron. Common window materials are alumina and sapphire. However, beryllia oxide, boron nitride, silicon nitride, and other materials have been used in special devices. At higher average power levels, the thermal characteristics of standard materials is such that the heat deposited in the ceramic cannot be adequately removed by cooling around the edges as in most linear beam devices. Figure 7 shows a schematic diagram of a double disk window that uses a dielectric fluid to convectively face-cool the disks.

The power-handling capability of the window is the major factor limiting output power for high average power gyrotrons. For tubes utilizing Gaussian mode output, internal mirrors are carefully designed to shape the beam profile for optimum window performance. In addition, new materials, such as chemically vapor deposited (CVD) diamond, are being studied to provide extremely high thermal conductivity and low dielectric loss.

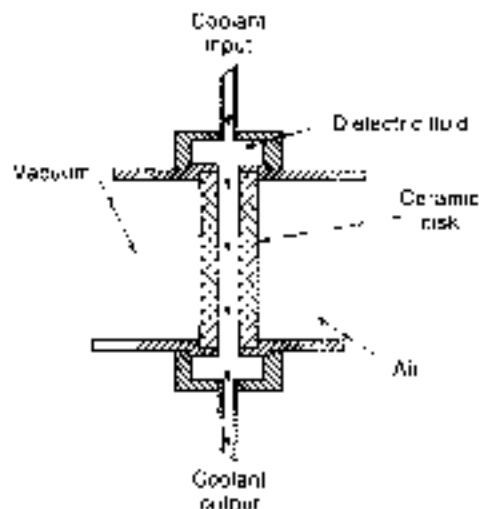


Figure 7. In the double-disk window, a dielectric fluid flows between the ceramic disks for enhanced cooling.

## AMPLIFIERS

While most gyro devices are oscillators, research is in progress to develop gyroamplifiers as drivers for linear accelerators (23) and for high-frequency, high-resolution radar (24). While oscillator design is relatively straightforward, amplifier design presents several challenges. Since the circuits are typically overmoded, it becomes more difficult to prevent RF leakage through the drift regions between cavities. Also, the high cavity  $Q$  factor characteristic of gyrotron output cavities limits the bandwidth that can be achieved. While several amplifier configurations have been studied, the gyroklystron and the gyro-TWT (traveling wave tube) have achieved the most success.

The gyroklystron consists of two or more cavities separated by drift spaces cut off for the operating mode. An azimuthal drift is imposed on the electrons by a sinusoidal energy modulation in the first cavity. These drifts are enhanced by intermediate buncher cavities, and the output cavity extracts the power from the azimuthally bunched electron beam. Most gyroklystrons employ cavity modes that are above the fundamental  $TE_{11}$  circular waveguide mode. Extreme care must be used to ensure that any RF power converted from the cavity mode to the  $TE_{11}$  mode is not allowed to propagate through the drift spaces and cause spurious oscillations. This also limits the circuit modes that can be used to those close to the fundamental (typically  $TE_{01}$  or  $TE_{21}$ ) circuit modes. Bandwidths for gyroklystrons are typically less than 1%. Operating efficiencies greater than 30% are common, and gains range from 20 to 40 dB.

In the gyro-TWT, an azimuthally bunched helical beam interacts with a traveling wave. If the periodicity and velocity of the electron bunches are such that synchronism occurs between the electrons and the traveling wave, cumulative bunching and energy extraction can occur. These amplifiers have considerably larger bandwidths than gyroklystrons; however, they operate at lower efficiency ( $< 20\%$ ) and generally produce less power.

Research has also been performed on gyro backward wave oscillators, gyroexcitrons, and gyropeniotrons. The reader is referred to the additional reading list for information on these devices.

## FUTURE DEVELOPMENT

Most gyrotron research is focused on increasing the output power of the device, primarily for ECRH and ECCO of fusion plasmas. It is anticipated that several hundred megawatts of RF power will be required for heating these plasmas. It is desirable to maximize the output power of the individual gyrotrons to reduce the number of gyrotrons required. The current goal is to develop gyrotrons producing 1 MW of CW power at frequencies between 100 GHz and 170 GHz (27-29). Much of this research is focused on individual components, such as the RF window (30) and depressed collectors for energy recovery (31-33). Other researchers are investigating alternative cavity configurations (34, 35).

Amplifier research is focused on developing gyroamplifiers for high-resolution radar and for high-power linear accelerators. Current research goals for radar applications include bandwidth enhancement, increased efficiency, increased output power, and gains exceeding 30 dB. It is anticipated that peak powers approaching 100 kW and average powers exceeding 10 kW will be required. Accelerator applications do not

require large bandwidths, but peak powers exceeding 100 MW at frequencies between 11 GHz and 20 GHz will be needed. High efficiency and reasonable gain will also be major goals. Advances in gyrotron components, such as windows, electron guns, launcher systems, and depressed collectors, will also be applicable to amplifier development.

Research is also in progress to develop gyrotrons for industrial heating applications, particularly for sintering of fine grain ceramics. Studies indicate that microwave and millimeter sintering can produce extremely high heating rates or selective heating in multiphase systems leading to novel ceramic materials with compositions and microstructures not possible using standard techniques (36). Several experiments are in progress to develop sources for these applications (37,38). It is anticipated that other applications will materialize when efficient, cost-effective sources are available.

## BIBLIOGRAPHY

1. R. D. Twiss, Radiation transfer and the possibility of negative absorption in radio astronomy, *Austr. J. Phys.*, **11**: 564-573, 1958.
2. J. Schneider, Stimulated emission of radiation by relativistic electrons in a magnetic field, *Phys. Rev. Lett.*, **2**: 504-505, 1959.
3. A. V. Gapunov, Interaction between electron fluxes and electromagnetic waves in waveguides, *Izv. Vyssh. Uchebn. Zaved. Radiofizika*, **2**: 460-462, 1959.
4. H. H. Pantell, Electron beam interaction with sea waves, *Proc. Symp. Millimeter Waves, Microw. Res. Inst. Symp. Ser.*, **0**: 1959.
5. K. K. Chow and R. H. Pantell, The cyclotron resonance backward wave oscillator, *Proc. IRE*, **48**(11): 1960.
6. I. B. Batt, Tunable source of millimeter and submillimeter wave radiation, *Proc. IEEE*, **52**: 830-831, 1964.
7. I. B. Batt, A powerful source of millimeter wavelength electromagnetic radiation, *Phys. Lett.*, **14**(4): 293-294, 1965.
8. J. Feinstein, Research on electronic interactions with the fields of mirror resonators. In *Proc. Int. Congr. Microwave Tubes 5th*, 1964.
9. J. L. Hirschfield and J. M. Wachtel, Electron cyclotron maser, *Phys. Rev. Lett.*, **12**(19): 533-536, 1964.
10. M. I. Petelin, Relativism of microwave electronics. In *Proc. Twenty-Second Int. Conf. Infrared Millimeter Waves*, Wintergreen, VA, July 1997.
11. L. A. Vaynshteyn, Open resonators and open waveguides, *Sovetskoye Radio*, 1966.
12. Y. V. Bykov et al., presented at the VII Inter-rodigiats Conf on SHP Electron., Tomsk, 1972.
13. S. N. Vlasov, I. I. Zagryadskaya, and M. I. Petelin, Transformation of a whispering gallery mode, propagating in a circular waveguide, into a beam of waves, *Radio Eng. Electron. Phys.*, **20**: 16-17, 1975.
14. V. A. Flyagin et al., *IEEE Trans. Microw. Theory Tech.*, **MTT-25**: 512-521, 1977.
15. J. L. Hirschfield and V. L. Granatstein, The electron cyclotron maser—an historical survey, *IEEE Trans. Microw. Theory Tech.*, **MTT-25**: 522-527, 1977.
16. J. L. Hirschfield, Gyrotrons, In *Infrared and Millimeter Waves*, Vol. 1, New York: Academic Press, 1978, Chapter 1, pp. 1-54.
17. B. S. Synnook and H. R. Jury, Cyclotron resonance devices. In *Advances in Electronics and Electron Physics*, Vol. 55, New York: Academic Press, 1961, pp. 103-185.
18. J. Mark Brind, Gyrotron theory. In V. L. Granatstein and I. Alexeff (eds.), *High Power Microwave Sources*, Norwood, MA: Artech House, 1983, pp. 103-145, 1987.
19. V. A. Flyagin et al., The gyrotron, *IEEE Trans. Microw. Theory and Tech.*, **MTT-25**: 514-521, 1977.
20. A. S. Gilmour, *Principles of Traveling Wave Tubes*, Hudson, MA: Artech House, 1984, pp. 103-149.
21. M. Thumm, State-of-the-Art of High Power Gyro Devices and Free Electron Masers, Update 1995. Forschungszentrum Karlsruhe, ITP, Association: FZK-Eurotron Postfach 3640, D-676021 Karlsruhe, Germany.
22. K. E. Kreischer et al., The design of megawatt gyrotrons, *IEEE Trans. Plasma Sci.*, **PS-13**: 364-373, 1985.
23. V. L. Granatstein and W. Lawson, Gyro-amplifiers as candidate RF drivers for TeV linear colliders, *IEEE Trans. Plasma Sci.*, **24**: 648-665, 1996.
24. B. Danyl et al., Development of W-band gyrokystron for radar applications, *Proc. Twenty-Second Int. Conf. Infrared Millimeter Waves*, Wintergreen, VA, July 1997.
25. M. Black et al., Experimental study of a high power W-band gyrokystron amplifier, *Proc. Twenty-Second Int. Conf. Infrared Millimeter Waves*, Wintergreen, VA, July 1997.
26. J. J. Choi et al., Experiments on high power 36 GHz gyrokystron amplifiers, *Proc. Twenty-Second Int. Conf. Infrared Millimeter Waves*, Wintergreen, VA, July 1997.
27. V. E. Myasnikov et al., Long pulse operation of a 110 GHz 1 MW gyrotron, *Proc. Twenty-Second Int. Conf. Infrared Millimeter Waves*, Wintergreen, VA, July 1997.
28. K. Sakamoto et al., Development of high power 170 GHz gyrotron for ITER, *Proc. Twenty-Second Int. Conf. Infrared Millimeter Waves*, Wintergreen, VA, July 1997.
29. V. E. Zepelov et al., Development of 1 MW output power level gyrotron for ITER, *Proc. Twenty-Second Int. Conf. Infrared Millimeter Waves*, Wintergreen, VA, July 1997.
30. M. Thumm, Development of output window for high power long pulse gyrotrons, *Proc. Twenty-Second Int. Conf. Infrared Millimeter Waves*, Wintergreen, VA, July 1997.
31. A. Singh et al., Integrated design of depressed collectors for gyrotrons, *IEEE Trans. Plasma Sci.*, **25**: 480-491, 1997.
32. R. L. Ives et al., Development of a multi stage depressed collector for a 1 MW CW gyrotron, *Proc. Twenty-Second Int. Conf. Infrared Millimeter Waves*, Wintergreen, VA, July 1997.
33. V. A. Flyagin et al., Investigation of coaxial gyrotrons at IAP RAS, *Proc. Twenty-Second Int. Conf. Infrared Millimeter Waves*, Wintergreen, VA, July 1997.
34. B. Piosczyk et al., Operation of a coaxial gyrotron with a dual RF-beam output, *Proc. Twenty-Second Int. Conf. Infrared Millimeter Waves*, Wintergreen, VA, July 1997.
35. B. Piosczyk et al., A 1.5 MW, 140 GHz, TE<sub>25,16</sub> Coaxial Cavity Gyrotron, *IEEE Trans. Plasma Sci.*, **26**: 461-469, 1997.
36. J. D. Calame et al., The narrowwave sintering of ceramics: New insights, models, and applications based on realistic ceramic microstructures, *Proc. Twenty-Second Int. Conf. Infrared Millimeter Waves*, Wintergreen, VA, July 1997.
37. A. W. Fillet et al., Millimeter-wave sintering of ceramic compacts, *Proc. Twenty-Second Int. Conf. Infrared Millimeter Waves*, Wintergreen, VA, July 1997.
38. A. W. Fillet et al., Pulsed 35 GHz gyrotron with unrounded applicator for sintering ceramic compacts, *IEEE Int. Conf. Plasma Sci.*, 1993, pp. 105-106.

## Reading List

The following journals contain most of the publications related to gyrotron research and development.

*IEEE Trans. Plasma Sci.*  
*IEEE Trans. Microw. Theory Tech.*  
*Int. J. Electron.*  
*Int. J. Infrared Millimeter Waves*  
*Proc. Int. Conf. Infrared Millimeter Waves* (yearly)  
*Proc. Int. Conf. Plasma Sci.* (yearly)  
*Izv. Vysok. Chistota. Zvezd. Radiofiz.*

Additional sources of information are listed below.

- R. B. Miller. *An Introduction to the Physics of Intense Charged Particle Beams*, New York: Plenum Press, 1982, pp. 238-249.
- H. R. Jory, R. Luen, and R. S. Symons. *Final Report of Millimeter Wave Study Program*, performed for Oak Ridge National Laboratory on Order No. Y-12 11Y-199433V, Varian Associates, Inc., Palo Alto, CA, 1975.
- K. R. Chu, *Theory of Electron Cyclotron-Maser Interactions in a Cavity at the Harmonic Frequencies*, NRL Memorandum Report 3672, Washington, DC: Naval Research Laboratory, 1977.
- K. S. Symons and H. R. Jory. Small-signal theory of gyrotrons and gyroklystrons. *Proc. Symp. Eng. Prob. Fusion Res.* 7th, 1977.
- H. R. Jory et al., Gyrotrons for high-power millimeter wave generation. *Proc. Symp. Eng. Prob. Fusion Res.* 7th, 1977.
- V. L. Bratman et al., Theory of gyrotrons with a nonlax structure of the high frequency field, *Izv. Vysok. Chistota. Zvezd. Radiofiz.* 18 (4): 1973.
- A. V. Gaponov et al., Induced synchrotron radiation of electrons in cavity resonators, *Sov. Phys. JETP Lett.* 2:97, 1965.
- M. Caplan and C. Tarrington. Improved computer modeling of magnetic injection guns for gyrotrons, *Int. J. Electron.* 51: 415-426, 1981.
- J. Neilsen et al., Determination of the resonant frequency in a complex cavity using scattering matrix formulation. *IEEE Trans. Microw. Theory Tech.*, 37: 1165-1171, 1989.
- V. A. Flyagin and G. S. Kosmach. Gyrotron oscillators. *Proc. IEEE*, 76: 644-656, 1988.
- K. Felch, H. Huey, and H. Jory. Gyrotrons for ECH applications. *J. Fusion Energy*, 9 (1): 59-75, 1990.
- V. V. Rykuz et al., Experimental investigations of a gyrotron with whispering gallery modes, *Izv. Vysok. Chistota. Zvezd. Radiofiz.* 18 (10): 1544-1546, 1975.
- J. B. Mead. Millimeter wave radars for atmospheric remote sensing. *Proc. Twenty Second Int. Conf. Infrared Millimeter Waves*, Wintergreen, VA, July 1997.

#### Peniodrons

- G. Nobler. *Int. J. Electron.* 56: 617-627, 1984.
- P. V. Ielko. *IEEE Trans. Microw. Theory Tech.* 32: 917-921, 1984.
- L. Zima et al., *Int. J. Electron.* 67: 1065-1075, 1984.

#### Amplifiers

- K. R. Chu and A. T. Drobot. Theory simulation of the gyrotron traveling wave amplifier operating at cyclotron harmonics. *IEEE Trans. Microw. Theory Tech.* MTT-28, 313-317, 1980.
- K. R. Chu et al., Characteristics and optimum operating parameters of a gyrotron traveling wave amplifier. *IEEE Trans. Microw. Theory Tech.* 27: 178-187, 1979.
- G. G. Denisenko et al., Gyr-TWT with a helical operating waveguide: New possibilities to enhance efficiency and frequency bandwidth. *Proc.*

*Twenty Second Int. Conf. Infrared Millimeter Waves*, Wintergreen, VA, July 1997.

- Q. S. Wang, D. B. McDermott, and N. C. Luhmann. Operation of a stable 200 kW second-harmonic gyro-TWT amplifier. *IEEE Trans. Plasma Sci.*, 24: 700-706, 1996.
- E. V. Zasyukin et al., Study of X-band three stage gyrogyrotron amplifier. *Proc. Twenty Second Int. Conf. Infrared and Millimeter Waves*, Wintergreen, VA, July 1997.

R. LAWRENCE (MS)  
 Calabazas Creek Research Inc.

## HALL EFFECT TRANSDUCERS

Hall effect transducers belong to a family of semiconductors that produce an output voltage proportional to a magnetic field or magnetic induction. Initially, the Hall effect was a curiosity comparable to the Seebeck or Peltier effects, but soon it was realized that it could be used for the study of conduction mechanisms in semiconductors. From 1960 on, the first technical Hall effect applications were introduced—for example, the magnetic field measurement probe, magnetic multiplier, and current-to-voltage transducer. At present, a range of powerful Hall integrated circuits are available for magnetic field measurement purposes, the position control of objects, and angle position of wheels. The latest development was the discovery of the quantum Hall effect. This currently has no technical applications, but it has permitted electrical metrology to make considerable advances.

In 1879, E. H. Hall made the remarkable discovery (1) that a magnetic field perpendicular to a conducting foil deviates the charge carriers in a conductor in such a way that a voltage is generated transversally to the foil. The use and exploitation of this effect was, however, severely restricted to research purposes due to the shortage of good materials. Favorite materials during the exploration period were bismuth and even germanium, but the stability obtained with these elements was very disappointing due to their temperature characteristics. With the expansion of research on semiconductors after World War II, Hall measurements were refined with success for the determination of resistivity, carrier concentration, carrier type, and mobility in semiconductors. This research led to still better semiconductor materials, soon, the first industrial instruments for magnetic field measurements with sensitive calibrated Hall effect sensors were introduced on the market.

With the development of integrated circuits (IC) the applications domain of Hall devices was further extended because it became possible to integrate an amplifier with the Hall sensor in one unit. Such an IC can be used for several purposes: position detection, measurement of current, and electronic compass. In 1980, the German scientist Klaus von Klitzing (2) discovered a new effect that was strongly related to the classical Hall effect: the quantum Hall effect. By studying a two-dimensional electron gas at low temperatures in the presence of a very strong magnetic field (7.4 kT), it was found that the ratio of the Hall voltage to the device current depends only on the Planck constant  $h$  ( $6.626 \times 10^{-34}$  Js) and the elementary charge  $e$  of the electron ( $1.6021 \times 10^{-19}$  C). The von Klitzing effect was very important to electrical metrology because it became possible to replace the material ohm standard in the same way as the volt standard was replaced by the Josephson junction.

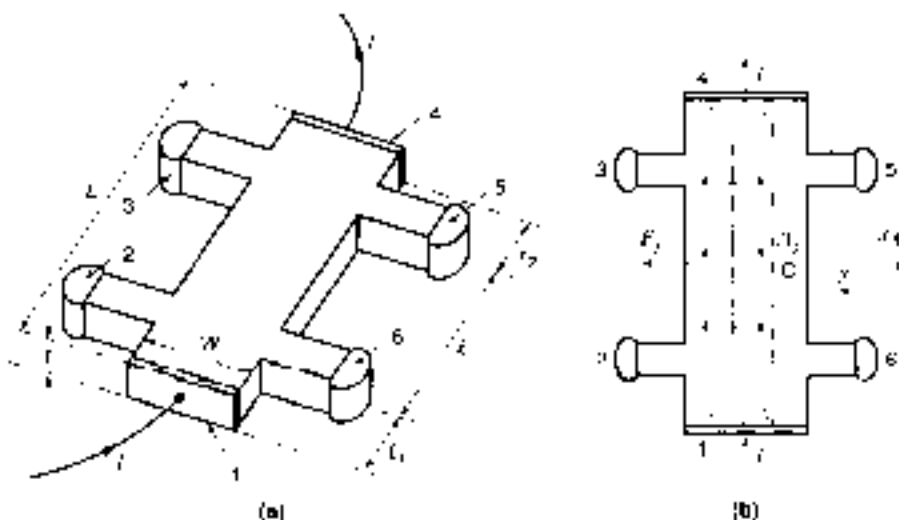


Figure 1. (a) Hall bridge-type sample with four voltage contacts (2, 3, 5, 6) and current contacts (1, 4); (b) deviation of carrier lines due to an applied  $B$ -field. The contacts 1 and 4 are supposed ideal, and  $l_1$ ,  $l_2$ , and  $l$  are much larger than the width  $W$  of the sample. The carrier density at the left and right side will differ due to the Lorentz force  $F_L$ .

## HALL EFFECT

For the derivation of the classical Hall effect (3), a rectangularly conducting sample (Fig. 1) with width  $w$ , length  $L$ , and thickness  $t$  is considered. At the four side planes, the contacts 1, 2, 5, 6 have been added for the connection of a current source  $I$  (1, 4) and for potential measurements (2, 3, 5, 6). It is supposed that the width of the contacts is small in comparison with the distance  $l$  between (2, 3) and (5, 6). Also, the distances  $l_1$  and  $l_2$  are much larger than the width  $w$ . The sample is placed in a magnetic field with induction  $B$  perpendicular to its largest plane. Figure 1(b) shows that this plane is associated with a rectangular coordinate system with the  $x$ -axis in parallel with the length axis of the sample. As long as  $B = 0$  in this system, the charge carriers in the sample move along the dotted straight lines between contacts 1 and 4.

The particular shape has the advantage that it can also be used for the determination of the classic resistivity of the sample material. The resistivity  $\rho_{xx}$  in the  $x$ -direction of the sample is determined with  $B = 0$ , by measuring the potential difference  $V_{23}$  (or  $V_{56}$ ) between the contacts (2, 3) or (5, 6). It follows that

$$\rho_{xx} = \frac{V_{23}}{I} \frac{L}{t} \quad (10)$$

If the contacts are symmetrically placed around the length axis of the sample, then we expect the voltage between contacts 2 and 6 (or 5 and 6) to be zero and in this case we also have  $V_{23} = V_{56}$ .

In the presence of a  $B$ -field, the moving charge carriers deviate from the straight lines due to the Lorentz force that the field exercises on the carriers. If the current flows in the  $x$ -direction, the carriers with an elementary charge  $q$  have a velocity  $v_x$ , and the corresponding force will be in the  $y$ -direction. The absolute value  $F_L$  of the Lorentz force is:

$$F_L = Bv_x q$$

As a result, the carrier concentration will be somewhat higher at the left-hand side than at the right-hand side of the sample [Fig. 1(b)]. It means that internally an electric field  $E_y$ , counteracting the effect of the Lorentz force is generated. This field

exercises a repelling Coulomb force  $qE_y$  on the carriers and in a steady-state situation both forces are equal and thus

$$E_y = Bv_x$$

The carrier concentration  $n$  (number of carriers per  $\text{m}^3$ ) and the velocity  $v_x$  are related to the device current  $I$ , because the amount of charge displaced in  $t$  s is equal to the product  $wtqv_x t$ . Therefore we have

$$v_x = I/wtq$$

Because the field  $E_y$  is constant for all points between contacts (2, 6) or (3, 5), it follows immediately that the voltage  $V_H$  between (3, 5) or (2, 6) is equal to  $w \cdot E_y$ ,

$$V_H = Bt/wtq = R_H BI/t, \quad \text{with } R_H = 1/wq \quad (11)$$

The constant  $R_H$  is the Hall coefficient (independent of the width  $w$ ), and it has the dimension  $\text{m}^3/\text{C}$ . For the sensitivity of the device to be large, it is thus required that the sample is thin and the carrier concentration not too large. If these conditions are fulfilled, the mobility of the carriers (proportional to the carrier velocity) has to be large in order to give a reasonable conduction in the  $x$  direction. This explains why in ordinary conductors the Hall effect is extremely small. From the foregoing derivation we can deduce some important facts.

1. Measurement of the Hall voltage permits determination of the average polarity of the charge carriers. It was indeed only supposed that the direction of the current was known, but this current could as well be established with holes or electrons depending on the type of semiconductor used. In the first case the left-hand side of the sample will have an excess positive charge and point 2 will become positive with respect to 6 (or 3 to 5). In the second case, 2 will be negative with respect to 6. For extrinsic  $p$ -type materials we find from  $R_H$  the carrier charge density (in  $\text{C}/\text{m}^3$ ),  $nq = 1/R_H$  and for the  $n$ -type this becomes  $-1/R_H$ .

2. From Eq. (11) of the Hall voltage we obtain the Hall resistance  $R_H = V_H/I$

$$R_H = R_H B/t$$



This quantity increases linearly with the applied field  $B$ . The Hall effect device is also called a Hall transducer because it converts the value of the magnetic field into a voltage.

3. It can be seen that the current lines in the sample are slightly longer when the field is applied than when it is not. It can therefore be expected that the longitudinal resistivity  $\rho_{xx}$  increases when the field increases, this is the magnetoresistance effect. In the Hall sample, however, the magnetoresistance effect is of minor importance because the ratio  $L/w$  is large.

Until now it has been supposed that the contact pairs are symmetrically placed with respect to the sample. If this is true, no Hall voltage can be generated between the contact pairs when the field is zero. In practice there always remains a slight asymmetry and, because of the finite resistivity in the  $x$ -direction, a small offset voltage  $V_{10}$  will add to the Hall voltage. This offset voltage can only be eliminated by reversing the direction of  $B$  or by rotating the Hall device over  $180^\circ$  around its  $x$ -axis and combining the two resulting Hall voltages. The two voltages corresponding to this procedure are, respectively,  $V_1 = V_{10} + V_H$  and  $V_2 = V_{10} - V_H$ . The Hall voltage thus is found from  $(V_1 - V_2)/2$ . For ac measurements, the offset can be eliminated by blocking the dc component of the output voltage of the Hall device.

## QUANTUM HALL EFFECT

In order to observe the quantization of the Hall resistance, several conditions have to be fulfilled: the charge carriers have to behave like a two-dimensional electron gas (2DEG), temperature has to be very low (e.g., below 4.2 K) and a very strong magnetic induction has to be applied (4.5 T).

A two-dimensional electron gas is the special condition that occurs in strong inversion or accumulation layers at planar semiconductor-semiconductor or semiconductor-insulator interfaces. In the high electric field at the interface, the electrons are only allowed to move freely in a very thin layer in parallel with the interface. When a magnetic field is applied perpendicular to the interface, the electrons will tend to move along circular paths due to the interaction of the field and the electron charge  $e$ . The frequency  $\omega_c$  of this rotation is called the cyclotron frequency and it can in a classic way be derived from  $r$ ,  $B$  and the cyclotron effective mass  $m_c$  by equating the centrifugal force and the centripetal force due to the  $B$ -field. From this it follows that

$$\omega_c = eB/m_c$$

Under normal conditions the electrons are scattered before a full revolution occurs. However, if the electron mobility is very high, a number of revolutions can exist, and the effect is that an angular momentum quantization occurs that influences the density states of the 2DEG. The net effect is that the resistivity  $\rho_{xx}$  shows an oscillating behavior as a function of the magnetic induction (Shubnikov-Haus effect). A necessary condition is that the product of the cyclotron frequency and the scattering time  $\tau$  of the electrons is much larger than one ( $\omega_c \tau \gg 1$ ). Because the scattering increases with temperature and  $\omega_c$  is proportional to  $B$ , the experiments have to be performed at low temperatures ( $< 40$  K) and in the presence of a high field (e.g., 15 T) in order to observe quantization. Figure 2



Figure 2. Hall resistance  $R_H$  and longitudinal magnetoresistivity  $\rho_{xx}$  on an arbitrary scale and as a function of the induction  $B$ . The dashed lines correspond to temperatures above 44 K while the solid lines are for less than 4.2 K. The Hall resistance shows quantized plateaus in this case.

shows the resulting Hall resistance  $R_H$  and the magnetoresistivity  $\rho_{xx}$  as a function of  $B$  under these conditions. As derived in the preceding, the Hall resistance increases linearly with  $B$  as long as the temperature is high, for example, above 40 K, and the magnetoresistance also increases with  $B$ . If temperature is low enough (about 1 K),  $\rho_{xx}$  becomes very small for certain ranges of  $B$  and the Hall resistance  $R_H$  becomes quantized following the relation

$$R_H(i) = h/e^2 \quad i = 1, 2, 3, \dots \quad (2)$$

In the special case  $i = 1$ , the Hall resistance becomes the von Klitzing constant  $R_K$

$$R_K = R_H(1) \quad (3)$$

For practical applications, the 2D gas can be implemented in Si-metal-oxide semiconductor field-effect transistors (MOSFET) or in a heterostructure with layers of GaAs and AlGaAs.

Because the quantized Hall effect is used in metrology, scientists have tried to detect deviations from the theoretical behavior by experiments or by extending the models for the samples in order to see whether or not  $R_H(i)$  is universally valid.

### Dependence on the Semiconductor Material

As the 2D gas can be implemented in MOSFET structures as well as in AlGaAs/GaAs structures, both materials have been carefully investigated (7,8). Measurements have confirmed that the value of  $R_K$  is with certainty independent of the material. The precision obtained with these measurements was at a relative error level  $3 \times 10^{-6}$ . A small number of researchers have found anomalies, but it has been shown that these can always be attributed to bad samples (9). It was possible to develop a test for the samples, in order to see in advance if the correct conditions for observing the quantized Hall effect are fulfilled.

### Dependence on the Sample

Several heterostructure samples of the same material but with very different carrier mobilities have been studied (5) and it

appears that the differences for  $R_H$  all fall within the uncertainty range of the measurement setup. On the other hand, the measurements also show that  $R_H$  varies as predicted with the integer  $i$ .

#### Width of the Sample

It is predicted by some theories (e.g., 10.11) that there could exist a relation between the width of the sample and  $R_H$ . However, studies performed on samples with widths of between 10  $\mu\text{m}$  and 1 mm did not reveal deviations larger than the measurement precision (12).

#### Contact Resistance

For connection of the Hall sample to the outside world, metallic contacts are required. Good contact resistances are of the order of a few milliohms, but in some samples values in the kilohm range have been observed. If the latter is the case, relative errors of the order  $1 \times 10^{-3}$  can be observed. The effect of bad contacts can not be predicted due to their microscopic nature. The experiments (13,14) show that this behavior is strongly influenced by the device current or temperature. It is estimated that the error caused by a good contact lies in the range  $1 \times 10^{-3}$ . It has to be noted that the errors encountered for samples with bad contacts are not due to the loading of the device by the voltmeter circuit. If the device is bad, there is always a measurable longitudinal voltage  $V_{xx}$  associated with it. Finally, the most important electrical attributes of a good quantized Hall resistor can be found in Ref. 15.

### APPLICATIONS OF THE CLASSICAL HALL EFFECT

The Hall effect can be used in semiconductor research as well as in technical applications.

#### Research Applications

The Hall effect is used intensively in semiconductor research for the determination of carrier concentration ( $n$ ) and mobility ( $\mu$ ) in combination with the determination of conductivity ( $\sigma$ ) or resistivity ( $\rho$ ). In the first instance, it is assumed that the layers under study are uniform. If the magnitude of the carrier charge is  $q$ , then Eq. (1) for  $R_H$  gives the concentrations for  $p$ -type and  $n$ -type semiconductors, respectively

$$\rho = 1/qR_H \text{ and } n = 1/qR_H \quad (4)$$

For the case where two carrier types are involved, the general relation between the conductivity mobilities  $\mu_p$ ,  $\mu_n$  and the conductance  $\sigma$  is given by

$$\sigma = q(\mu_n n + \mu_p p) \quad (5)$$

and, if  $p \gg n$ , this reduces to

$$\sigma = q\mu_p p = \mu_p / R_H \quad (6a)$$

For electron carriers the corresponding equation is:

$$\sigma = q\mu_n n = \mu_n / R_H \quad (6b)$$

In principle, it is possible to determine the mobility  $\mu_p$  from this equation by measuring  $\sigma$  and  $R_H$ . Nevertheless, some problems that can arise require corrections.

A first problem is the occurrence of carrier scattering mechanisms that require a correction factor to be introduced when  $\mu_p$  or  $\mu_n$  are calculated from Eq. (6). A scattering factor  $r$  has to be introduced in Eq. (6a) and its equivalent (Eq. (6b)) for the electrons as follows:

$$\rho = r/qR_H \text{ or } n = r/qR_H \quad (7)$$

Depending on the scattering mechanism, the scattering factor lies in the range 1 ... 2. Therefore, the values of the conductivity mobilities derived from  $R_H$  have to be corrected with  $r$

$$\mu_p = r\mu_p' \text{ or } \mu_n = r\mu_n' \quad (8)$$

The mobility  $\mu_H$  is derived from the Hall voltage and  $\sigma$  is the Hall mobility and  $\mu_p$  and  $\mu_n$  are the ordinary conductivity mobilities. Although the latter are more important, the Hall mobility is easier to determine and can still give comparative information about the material.

Second,  $R_H$  depends also on the  $B$ -field and, when both types of carriers are present as well as a  $B$ -field, the Hall coefficient becomes rather complex. In Ref. 16 the following equation is given:

$$R_H = r \frac{1/p - b^2 n / (q \mu_p B^2) - n / (q \mu_n B^2)}{q(\mu_p + b n \mu_n / \mu_p + \mu_n B^2 / p - n^2)} \quad (9)$$

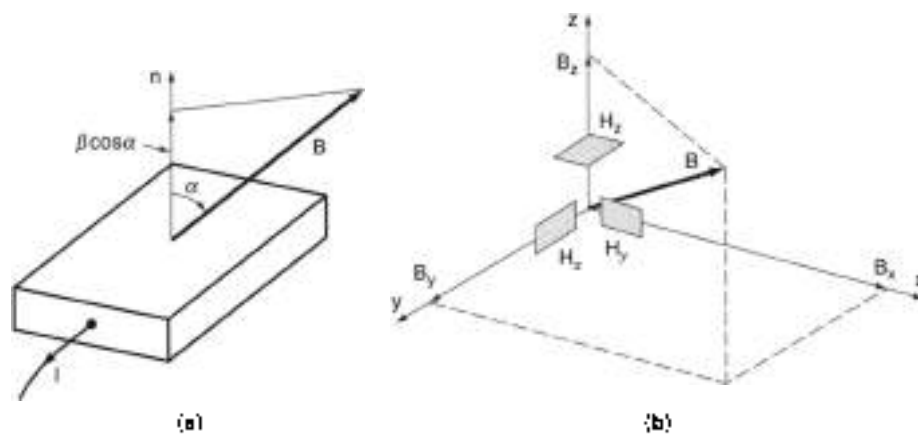
The coefficient  $b$  is equal to the ratio of the mobilities  $\mu_n$  and  $\mu_p$  or  $b = \mu_n / \mu_p$ .

For  $B \rightarrow 0$  and  $B \rightarrow \infty$ , the general expression (Eq. (9)) can be simplified to a low-field and a high-field expression, respectively. Depending on the mobilities it is possible to have semiconductors which obey either the low- or high-field expression, or the general equation (Eq. (9)). The low-field expression will be valid when  $B\mu_n \ll 1$  if  $p \ll n$  or  $B\mu_p \ll 1$  if  $p \ll n$ . Likewise, the high-field limit is valid when  $B\mu_n \gg 1$  if  $p \gg n$  or  $B\mu_p \gg 1$  if  $p \gg n$ .

For some semiconductors the conduction mechanism changes with temperature. For example, in HgCdTe one finds that for  $T < 200 \text{ K} \dots 300 \text{ K}$  electrons are the main carriers. At the lower temperatures a mixed carrier conduction occurs and, as a result,  $R_H$  eventually can change its sign. Therefore, Hall measurements are always performed at different temperatures and induction in order to extract the necessary information from the measured data. An extra complication is the dependence of  $r$  on  $B$ .

In order to be able to determine the spatial resistivity and mobility profiles in nonuniform layers, it is necessary to remove small layers from the sample so that Hall coefficients can be measured as a function of thickness. This can be done by etching or, preferably, by anodic oxidation. If another method used for layers on an insulator, a  $p$ - $n$  junction, a MOSFET or a mesa capacitor is formed at its upper surface. If, for example, a metal gate layer is used, the penetration depth of the space-charge region can be controlled by means of the reverse bias of the junction and accurate measurements of the mobility and carrier concentration are the result.

In Ref. 3 one can find a review of the most common sample shapes used, and the precautions to be taken for the reduction



**Figure 8.** (a) Component normal to Hall device for an arbitrary direction of the field and (b) setup with three elements for the determination of magnitude and direction of field.

of errors. Especially, if parameter estimation techniques are used the results can deliver much information about the material. Hall measurements on irregular sheets are also possible, provided some conditions are met. The theory for such samples was derived in 1958 by van der Pauw; see Ref. 17.

### Technical Applications

Because it became possible to integrate the electronics and the Hall effect transducer on the same substrate, a number of important industrial applications have been demonstrated. The Hall transducer can be used for direct measurement of magnetic induction and power, switching applications, position control, and so on. A Hall device will not be destroyed by strong magnetic fields and when semiconductors of the III-V type (GaAs, InSb) are applied, it is possible to enhance the sensitivity and the allowable temperature range. With ion-implanted GaAs a temperature range of  $-40$  to  $+250$  °C is obtained. The advantage of using silicon lies in the fact that the integration of an amplifier with the Hall element is easy. However, sensitivity is lower due to the low mobility of the carriers in silicon.

In combination with magnets fixed on moving parts, the Hall transducer effectively becomes a position sensor. By monitoring the output, one can get an idea about the velocity or the acceleration of a moving object. Such a transducer will never wear out because there is no friction. Aging will also be very low because of the large mean time between failures (MTBF) of an integrated circuit. In contrast with optical devices, the operation of this transducer will not be disturbed by moisture, dust, or dirt except when the dust contains iron particles. The Hall device is also relatively fast; the rise and fall times are of the order 0.2  $\mu$ s and magnetic ac fields with frequencies of up to 100 kHz can be detected. The substrate containing the transducer is either a ceramic material or a ferromagnetic substance. The latter is used if it is necessary to decrease as much as possible the permeability of the magnetic circuit in which the device is mounted.

**Linear Hall Effect Transducers.** The output of a linear Hall effect device is proportional to the product of the induction  $B$  and the device current  $I$ ; in this way, it performs an analog multiplication. It is also obvious that such an element can be dedicated to the measurement of magnetic fields if  $I$  has a constant value.

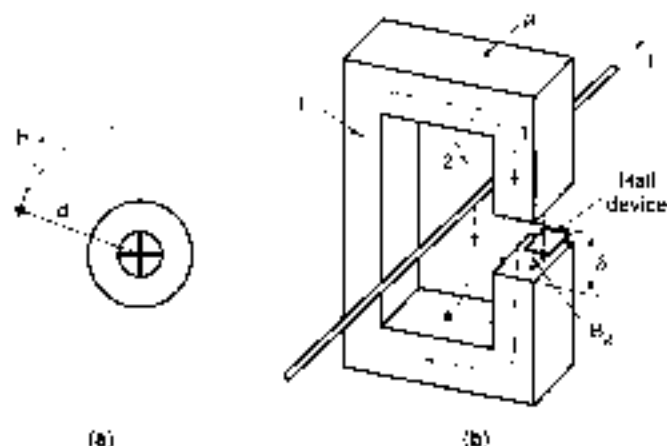
**Measurement of Magnetic Fields.** The instrument using Hall effect sensors for the determination of magnetic induction

(18–22) is called a gaussmeter although it should be termed a teslameter, as the unit of induction is weber (T) in SI units.

From the preceding it follows that the Hall voltage is due to the component of the field normal to the main sheet plane of the device. If the angle between the field  $B$  and the normal is  $\alpha$  [Fig. 3(a)], then the normal component is  $B \cos \alpha$ . For maximum sensitivity it is required that  $\alpha = 0$  and if the field has an unknown direction it is necessary to search for maximum output by rotating the device in space. If three separate Hall devices are orthogonally arranged [Fig. 3(b)] so that their main planes coincide with the planes of a rectangular axis system, the outputs will be proportional to the spatial components  $B_x$ ,  $B_y$ , and  $B_z$  of  $B$ . The magnitude and direction of the vector  $B$  in space can then be derived from the three outputs. It is, however, required that the probes be carefully calibrated. This can be done by adjusting the individual device currents, or by measuring the three sensitivity factors and storing them in a memory for further use with the software calculations. If large spatial gradients exist in the field, then the device output represents a kind of average. This error can be reduced by using very small elements; mostly it is possible to adapt the magnetic circuit to minimize this effect.

As sensitivity depends on  $I$  it seems appropriate to take  $I$  as large as possible. However, this also increases both device dissipation and current consumption. Therefore the advantage gained by increasing  $I$  is rather marginal and it is better to amplify the Hall voltage electronically if a higher sensitivity is required. For spatial magnetic field measurements it is preferable that the Hall device has a ceramic substrate as this ensures that the field is not disturbed by the device. Due to the large field of applications of gaussmeters, several types of probes with a calibrated built-in Hall element are available on the market. For measurement of fields in air gaps, for example, in electrical machines or relays, thin probes are preferred but these probes are very sensitive to mechanical stress. The Hall device is mounted at the end of a long flat isolating stylus so that less accessible points can be probed as well. The transverse probe is sensitive for fields perpendicular to the stylus plane. The axial probe is cylindrical and the Hall device is mounted perpendicular to the axis of the cylinder; it measures fields parallel to the probe axis. Full-scale sensitivity can be as large as 15 T and rms readings of ac fields are also possible.

**Measurement of Electrical Current.** Although not strictly necessary, a magnetic circuit is commonly used with the Hall device for current measurement (19,23) purposes [Fig. 4(b)]. This



**Figure 4.** (a) The radial field around a current-carrying conductor depends on the position of the Hall element. (b) magnetic circuit with a Hall device in the air gap. At low saturation all field lines in the core follow the magnetic path (1). When saturation occurs, some of the field lines (2) are lost for the air gap.

has two advantages: the sensitivity of the measurement is increased, and the position of the current carrying conductor becomes unimportant. The magnetic field at a distance \$d\$ of a straight conductor (Fig. 4(a)) carrying a current \$I\$, has a value \$H = I/2\pi d\$ and this is position sensitive. A Hall device at this position will give an output proportional to \$\mu\_0 H\$. If the current encircles a magnetic circuit, the force lines (e.g., 1 in Fig. 4(b)) will concentrate in the magnetic material (mean length \$L\$, magnetic permeability \$\mu\$) and in the air gap with width \$\delta\$. The value \$B\_g\$ of the magnetic induction in the air gap becomes, with Ampère's law and because of the constant flux in the core and air gap section,

$$B_g = \mu_0 I / (L + \mu \delta) \quad (10)$$

For \$\mu = 10^3\$, \$L = 0.1\$ m, \$\delta = 1\$ mm (\$10^{-3}\$ m); the value \$\mu\delta = 10\$ and \$L\$ can thus be neglected. It means that in the range where \$\mu\delta \gg L\$, the gap induction \$B\_g\$ can be approximated by \$\mu\_0 I/\delta\$ and the sensitivity becomes independent of \$\mu\$ and yet can be very high because \$\delta\$ is so small. If the core encircles the magnetic circuit \$N\$ times, then \$B\_g\$ also becomes \$N\$ times larger. The material properties of the core material hardly influence the precision of the measurement, but a very stable air-gap width \$\delta\$ is required. Also, when the current increases, the core induction increases and the material has a tendency to saturate. As a result, some of the magnetic field lines can leave the legs of the core and pass through the window area (e.g., line 2 in Fig. 4(b)). Because a part of the flux is lost, the air-gap induction becomes too low and a nonlinearity is introduced in the transducer characteristic. The existence of an air gap is favorable in this respect because it reduces the core induction at the cost of a lower sensitivity. The design of such a circuit is therefore a trade-off between sensitivity and full-scale behavior. The nonlinearity can remain below \$\pm 1\%\$ with ordinary core materials such as ferrite. Besides the electrical offset of the Hall device, a magnetic offset is caused by the core material. This is due to the remanent magnetism that remains after a heavy magnetization of the core. The remanent field gives an additional electrical output offset. Reduction of the remanent magnetism is done by choosing a very soft magnetic material. It is not possible to correct for this remanent field using software

because its value depends on the history of magnetization and, as precise models for the core behavior are lacking, a precise offset error correction is unpredictable. In Ref. 24, one can find how the problem of the intrinsic Hall plate offset voltage can be alleviated by combining the output of several elements. The simple circuit can still be very useful if the highest precision is not required because it requires only a Hall IC and a magnetic circuit, and power consumption is very low.

**Closed-Loop Hall Effect Current Transducer.** The linearity and offset problems can be reduced considerably by applying feedback (25) to the simple circuit (Fig. 5). In this case, the Hall voltage amplifier drives a second winding \$N\_2\$ coupled with the core and a resistor \$R\$. The voltage drop \$RI\_2\$ over this resistor is taken as the new output voltage. By choosing the correct winding, the amplifier output current \$I\_2\$ can counteract the effect of the unknown current \$I\_1\$. The number of ampere turns generating the field is now equal to the difference of the ampere turns of the current and feedback windings. Equation (10) for the magnetic circuit becomes, with \$H\_1\$ and \$H\_2\$, respectively, the magnetic field in the magnetic circuit and in the air gap

$$H_1 L + H_2 \delta = H_1 L - I_1 N_1 = N_1 I_1 - N_2 I_2 \quad (11)$$

The amplifier output voltage \$E\_0\$ is proportional to the air-gap induction and with an induction-to-voltage gain \$K\$, we obtain

$$E_0 = K \mu_0 H_2 / \delta$$

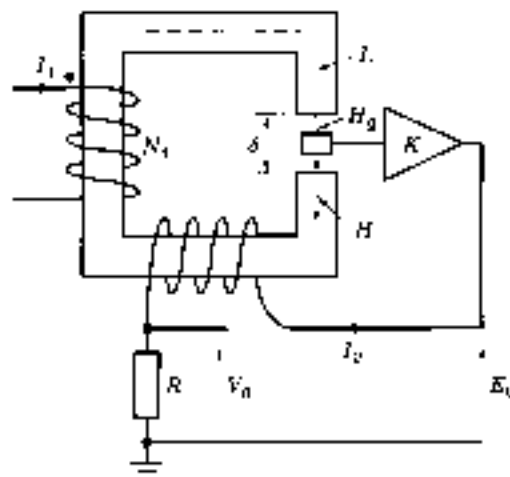
As the corresponding current \$I\_2\$ is \$E\_0/R\$, the air-gap field becomes

$$H_2 = N_1 I_1 / (L\delta - L/\mu + N_2 K \mu_0 / R)$$

If the amplifier gain \$K\$ is made very large, it is seen that \$H\_2 \to 0\$ and the field in the air gap disappears due to the feedback. It follows that \$N\_1 I\_1 = N\_2 I\_2\$ and

$$V_0 \approx (N_1/N_2) I_1 \quad (12)$$

As is the case in every feedback system, the static error can be neglected if \$K\$ is high enough. The output \$V\_0\$ is a very accurate



**Figure 5.** Closed-loop Hall effect device in magnetic circuit. If the loop gain is very large the primary and secondary ampere turns will cancel and \$V\_0\$ becomes proportional to \$I\_1\$.

image of  $I_1$  because the winding resistances, leakage inductances, and losses in the core, have no effect on accuracy. If the material is initially sufficiently demagnetized, the permeability will be very high and the magnetic offset error is reduced. In contrast with the open-loop system, the length of the gap barely influences the measurement precision, as a result, the temperature behavior is also enhanced. When the current  $I_1$  exceeds the output current capability of the amplifier, for example, under short-circuit conditions, the core is likely to be saturated and, therefore, a means for demagnetizing the core is standard in such an instrument. Demagnetization is obtained by driving temporarily the winding  $N_2$  with an exponentially damped oscillating (e.g., sinusoidal) current. If the initial value of this current drives the core into strong saturation, eventually the core will be demagnetized. For current probe applications the core is divided into two parts, which can be separated in such a way that the conductor can be enclosed by the core window without breaking the electric circuit. Especially, in this case it can be appreciated that the possible instability caused by imperfect air halves contact is eliminated by the feedback. The required current output capability of the amplifier depends entirely on the winding ratio  $N_2/N_1$  and the full-scale current  $I_1$

$$I_2 = I_1 N_1 / N_2 \quad (13)$$

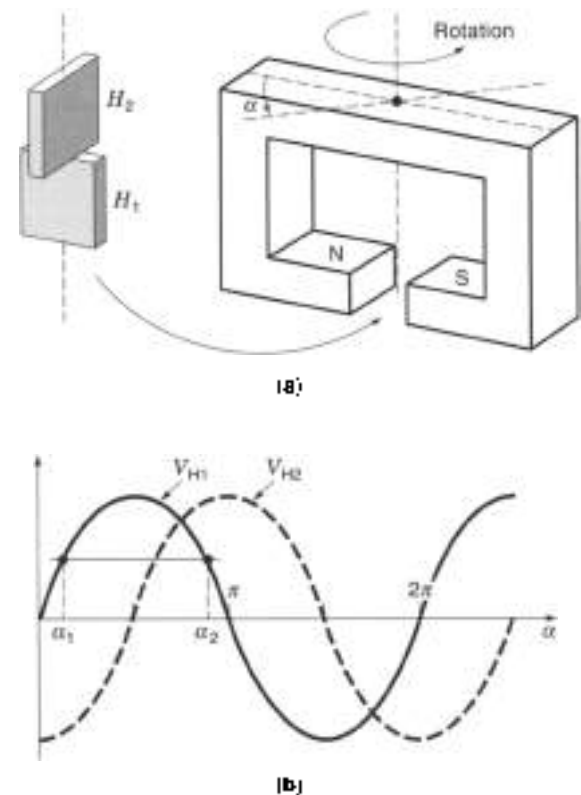
For practical reasons, the winding ratio has to be limited. This means that the amplifier in general will have to deliver a much larger current than in the open-loop system and its power consumption will be large. Another problem is that additional circuitry is required for wideband measurements because the frequency response of the Hall device is limited to about 100 kHz. A solution is to split the circuit into two parts. First, the frequency response of the Hall amplifier will deliberately be limited to a few kilohertz by first-order lowpass filtering of the output  $V_H$ . The purpose is to reduce the effect of the poles in the open loop transfer of the Hall device. Next, for the higher-frequency part of the spectrum, the two windings can be made to operate as a current transformer. The signal obtained from this current transformer is also first-order high passed, with the same 3 dB cut-off frequency and transfer gain as for the Hall device. By combining the two filter outputs a flat frequency characteristic is then obtained. Typical probes show a flat frequency response from dc to 20 or 100 MHz and have full-scale ranges from 1–10 A with a full-scale output voltage of 50–200 mV.

**Multiplier Applications of the Hall Effective Device.** Although ac power can be measured by sampling the current and the voltage of a consumer, in general, the potential drop, and, hence, the dissipation in the current shunt, are relatively large. Further, it is difficult to protect the electronics unconditionally against a short-circuit at the consumer side. On the other hand, when the load current flows through the magnetizing coil of Fig. 4, a short will not harm the Hall device or the magnetizing coil because the coil is very rugged. As suggested by Eq. (11) the Hall device output is proportional to the product of  $B$  and  $I$ . In fact, one of the first applications of the device was as an analog multiplier (19,20). Although this application has been superseded completely by the use of the Gilbert cell analog multiplier ICs, the device still has attractive features for power applications. For example, for determination of the power an ac load consumes, the product of the momentary voltage  $v(t)$  and current  $i(t)$  has to be calculated. The former current-measuring circuit can easily be adapted for this purpose. It suffices to vary the

Hall current  $I$  proportional to the voltage  $v(t)$  so that the output voltage  $v_H(t)$  of the device becomes proportional to the momentary power. By filtering this voltage, a measure for the average power is obtained. If  $v_H(t)$  is fed to a voltage-to-frequency converter followed by a counter, the counter output will represent the energy consumed by the load. For household applications the preferable counter is a mechanical one because it gives a clear indication and holds the current value, even during a power loss. The advantage of using the Hall device lies in the fact that the current scale for the instrument can be adapted through a change of the winding  $N_1$  in the same way as a classical Watt-hour meter with a rotating disk induction motor.

**Clinical Hall Effect Devices as Position Sensors.** A number of industrial applications require knowledge of the position of a mechanical part; the Hall effect ICs have been developed especially for this task (23). For example, in brushless dc motors the angular position of the rotor has to be known in order to activate the drive electronics of the field coils. More recently the Hall sensor has been applied in magnetic bearings (26) for magnetic levitation purposes.

If there is no magnetic field available, the Hall devices will be used in combination with permanent magnets (27). Suppose the rotation angle  $\alpha$  of a permanent magnet has to be determined (Fig. 6a). If a linear sensor  $H_1$  is positioned in a zone between the pole shoes, where the field of a rotating magnet is homogeneous and has a value  $B$  the output  $V_{H1}$  will be proportional to



**Figure 6.** (a) Hall device used as an angle sensor: a magnet is fixed on a turning axis and the Hall device is mounted between the pole shoes of the magnet. The output of  $V_{H1}$  is proportional to  $\sin \alpha$  because for  $\alpha = 0$  its plane coincides with the plane of the magnet. (b) A second device is necessary in order to remove all possible ambiguity from determination of  $\alpha$  from the Hall voltages.

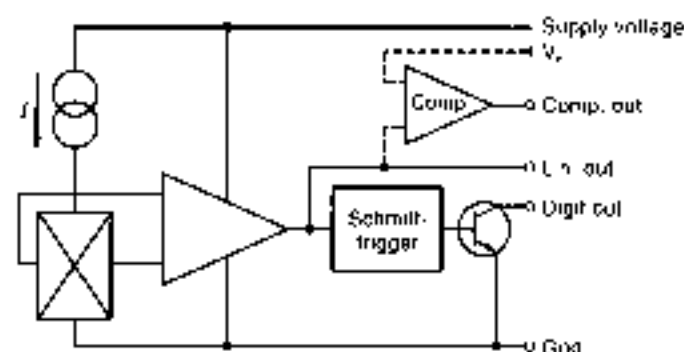


Figure 7. Internal circuit of a Hall IC with Schmitt-trigger and linear output. A comparator output is also useful for limit detection.

the cosine of the angle of rotation  $\alpha$  on the condition that the rotation axis coincides with the length-axis of the Hall sensor plane. From output  $V_{H1}$  it is, however, impossible to obtain  $\alpha$ , as there are two angle solutions for every voltage ( $\alpha_1$  and  $\alpha_2$  in Fig. 6(b)). This ambiguity can be resolved by placing a second transducer  $H_2$  (output  $V_{H2}$ ) perpendicular to  $H_1$ . From  $V_{H1}$  and  $V_{H2}$  and taking into account the signs of both outputs, the position of the magnet can be determined in the whole  $\alpha$ -range ( $0, 2\pi$ ).

A digital output can be obtained by including a Schmitt-trigger circuit (Fig. 7), driven by the analog output. Mostly, an open-collector transistor forms the output as this permits adaptation of the IC to the logic level of the system connected to it. Sometimes a small magnet is mounted close to the sensor in order to provide for a magnetic bias. The output can be made to react in two ways as shown in Fig. 8. In Fig. 8(a), the purpose is to discriminate between the presence or absence of a magnetic pole. Without field the output is at the logical level "1" (line 1). When the magnet approaches the device its analog output increases. At the trip point of the Schmitt-trigger corresponding to a field  $B_1$ , the digital output falls to the logical level "0" (line 2). In a third phase, the magnet is removed from the device but output remains low (line 3) as long as  $B$  lies above the lower trip point  $B_2$ . In Fig. 8(b), the lower trip point corresponds to a negative field, that is, the other pole of the magnet. Operated in these two ways, the Hall device acts like a switch. The hysteresis of the Schmitt-trigger reduces the effect of branching and noise. There exist three common ways for detecting the position of a moving magnet: Fig. 9(a-c). It is assumed that the magnet moves along the indicated  $x$ -axis in all three cases. In Fig. 9(a), the magnet moves in a direction perpendicular to the Hall device plane. The analog output falls continuously with increasing displacement  $x$ , and a digital action is possible if

the circuit is followed by a comparator. In Fig. 9(a), the comparator switches from low to high if the Hall voltage exceeds reference voltage  $V_2$ . The main drawback is the poor linearity of the analog output with respect to the movement of the magnet. In Fig. 9(b), the movement is along an axis in parallel with the device plane and the south pole is near the magnet. The output will be at maximum when the south pole is positioned at  $x = 0$ . The system can be used as a mechanical window detector, if the magnet is positioned within the range  $(-x_1, +x_1)$ , the comparator output is low and vice versa. The sensitivity is controlled by adapting the distance  $d$  between the magnet and the device. Linearity can be acceptable in a small part of the displacement curve around the point of the curve where the second derivative with respect to  $B$  vanishes. In the third case of Fig. 9(c), the magnet is turned over 90° and a linear device sensitive for negative and positive fields is used. Sensitivity can be controlled as in the former case and linearity can be extremely good but the sensitivity control will also influence the linearity. In the setup of Fig. 9(c), sometimes both magnet and Hall device are fixed but a high permeability screen moving between them can interrupt the field. Such a system can detect the teeth of a cogwheel or it can replace the mechanical breaker of the ignition in a car.

**Quantum Hall Effect in Electrical Metrology.** In the International System of Units (SI) the four base units are the meter (length), kilogram (mass), second (time) and the ampere (current). The ampere is the current that generates a force of 1 N per length between a pair of very thin and infinitely long conductors. This definition is artificial, and of little practical value, but it ensures that the mechanical and electrical units of force are the same. Apart from this, some of the mechanical units derived are introduced for practical purposes, for example, the Henry, coulomb. These secondary quantities can always be expressed in the base units. The magnitude of magnetic permeability in the SI system is fixed to  $4\pi \times 10^{-7}$  H/m. The second is defined by means of the period of a well-defined transition in Cesium 133. The velocity of light ( $c$ ) has a fixed value and the meter is the distance traveled in a vacuum by light during  $1/299\,792\,458$  s. This makes the use of the material standard meter obsolete. All electrical units can be defined in this system if two of them, for example, the volt and the ampere, are related to the mechanical units meter, second, kilogram. This comparison can be implemented via a kelvin current or volt balance. In practice the ampere has been replaced by the ohm as the latter is related to the ampere by Ohm's law.

Because the realization of physical standards requires complex instruments, national metrology laboratories still possess material standards such as Weston cells and 1  $\Omega$  resistors for

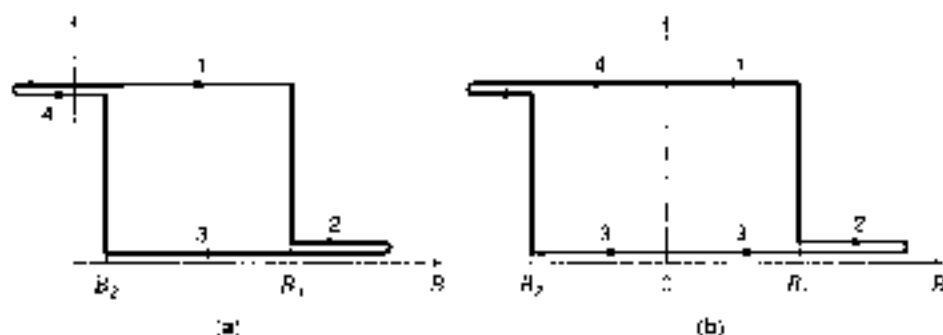
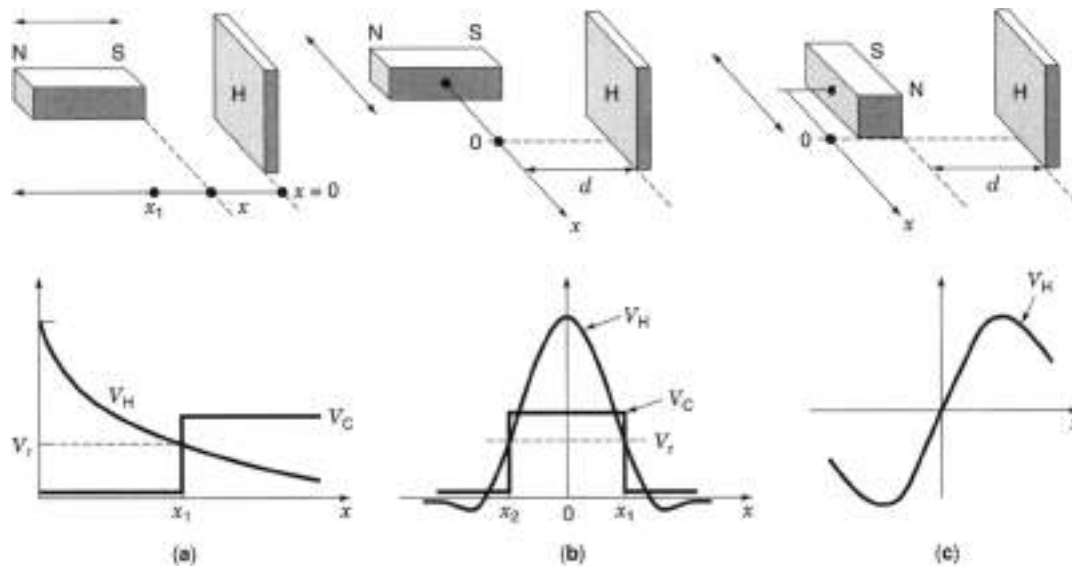


Figure 8. (a) Schmitt-trigger output for a unipolar sensitive Hall IC; (b) output for bipolar sensitive IC.



**Figure 9.** Three ways to detect a magnet movement: (a) magnet axis and movement normal to Hall plate and resulting linear ( $V_H$ ) and comparator ( $V_C$ ) output voltages; (b) movement axis parallel with Hall plate gives a window function for  $V_C$ ; its output is linear over a large displacement range.

their electrical measurements. These standards are considered to be a kind of flywheel in which an image of the physical unit is conserved for a limited time. By regularly comparing the physical standards with the material standards, the correction for each of these is determined. The natural volt ( $V_{1,AN}$ ) and ohm ( $\Omega_{1,AN}$ ) of a nation is then defined from the average values of these material standards. Before the invention of the Josephson junction and the quantified Hall effect, it was imperative to perform regularly cumbersome international comparisons in order to determine the corrections between the national standards  $V_{1,AN}$  and  $\Omega_{1,AN}$  of the different countries.

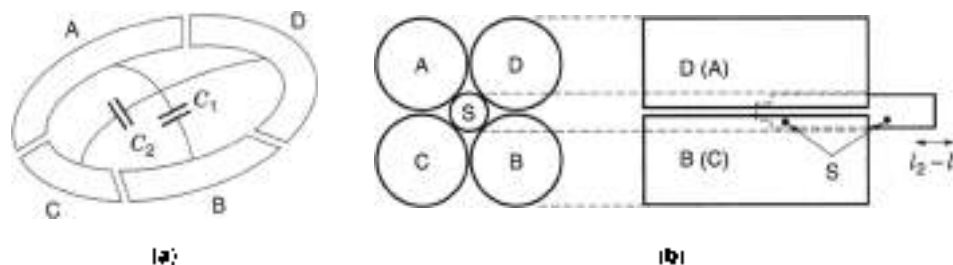
In 1936, Thomson and Lampard (28,29) derived a striking theorem concerning a special cylindrical four-electrode capacitor structure. It was proved that for the structure of Fig. 10(a), independently of the shape of the cross section of the electrodes A, B, C, and D, the capacities  $C_1$  and  $C_2$  per unit length of opposite electrodes obey the expression

$$\exp(-\pi C_1 / \epsilon_0) + \exp(-\pi C_2 / \epsilon_0) = 1 \quad (14)$$

with  $\epsilon_0$  the permittivity of the absolute vacuum. The value of  $\epsilon_0$  is known because, in vacuum, the velocity of light is  $c$

$(3 \times 10^8 \text{ m/s})^2$  and  $e$  as well as  $m_e$  are fixed. For a symmetrical capacitor Fig. 10(b),  $C_1 = C_2 = C$  and from Eq. (14) we obtain  $C = 1.957549043 \text{ pF/m}$ . In practical applications, a capacitor with finite length is used and this length is varied by inserting a guard tube S in the space between the four electrodes (30-32). By careful construction it is possible to reduce the effect of the fringing fields at both ends of the capacitor. In this way, the physical length of the capacitor is changed by displacing the guard tube from  $l_1$  to  $l_2$  and the corresponding variation of capacity is  $1.953549043(l_2 - l_1) \text{ pF}$ . The capacity variation in this system is, therefore, only determined by one dimension, the displacement  $l_2 - l_1$  of the guard tube can be measured very accurately with a laser interferometer. The very complicated setup for this system has been called the calculable capacitor because of its capacitance formula. By employing a chain of ac comparators bridges the material ohm standard can be related to the impedance of the calculable capacitor, and thus with the mechanical units.

The quantum Hall effect has opened the door to a new material ohm standard, which is solely dependent on the universal physical constants  $e$  and  $h$ . On the other hand, it is generally accepted (33) that the von Klitzing constant  $R_K$ , that is, the



**Figure 10.** (a) Calculable capacitor structure with four electrodes; (b) simplified system with four cylindrical electrodes A, B, C, and D. The effective length is varied by moving a screen S in the opening between the cylinders. Not shown are a similar but fixed screen at the left side, the screen surrounding the tubes, interferometer, and trimming screws. The smaller part of the screen is the "spike." It can be shown that this "spike" reduces the effect of non-polarities in the cross section of the capacitor (Ref. 32).

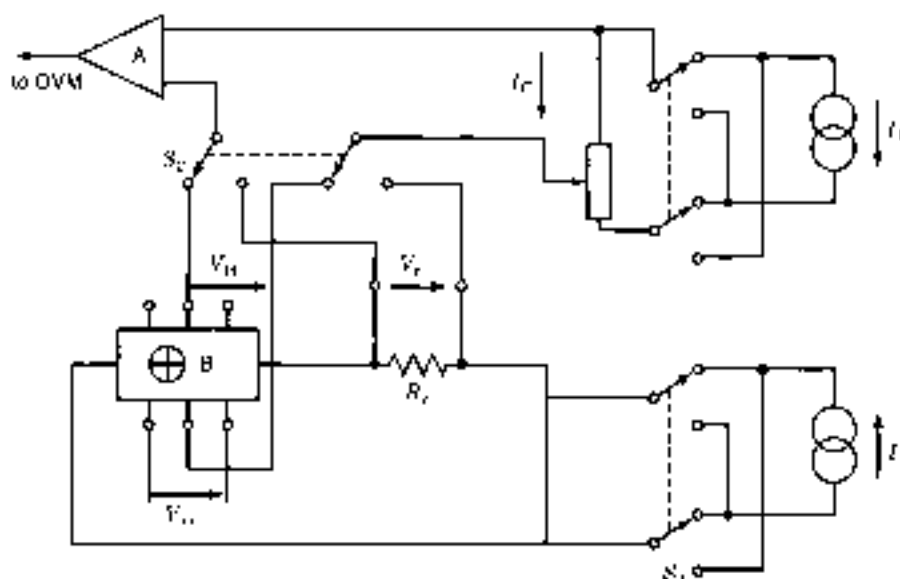


Figure 11. Potentiometric setup for comparing the quantified Hall resistance with the laboratory standard resistor.

value of  $R_H$  for  $i = 1$ , can also be calculated from  $\mu_B$ ,  $e$  and the dimensionless fine structure constant  $\alpha$ . This constant can be derived via complex quantum electrodynamic calculations from measurements of the anomalous magnetic moment of the electron. The relation between  $\mu_B$ ,  $e$ ,  $\alpha$ ,  $\hbar$  and  $c$  is given by

$$R_K = R_H i = 1 = \hbar / e^2 = \mu_B c / 2\pi e \quad (15)$$

with a value  $\alpha = 1/137.0359895$ . Since  $\mu_B$  and  $e$  have well-defined values, the accuracy of  $R_K$  and  $\alpha$  is the same in Eq. (15).

As it became clear that the quantum Hall effect could be used for definition of the unit of resistance, metrologists succeeded in refining the measurement procedures so that  $R_K$  could be expressed in  $\Omega_{\text{KJ}} \cdot h$ . It also became possible to compare the existing material ohm standards with a relative error of  $10^{-6}$ . At that moment the reproducibility of the quantum Hall measurements was better than the implementation of the material ohm in the International System of Units (SI).

In 1990, the Comité Consultatif d'Electricité (CCE) decided, based on an analysis of all the known international comparisons of the material resistance standards by the Bureau International des Poids et Mesures (BIPM), to attach a conventional value to the von Klitzing constant

$$R_{K-90} = 25812.807 \Omega$$

The relative uncertainty for this value is  $2 \times 10^{-7}$  and this definition makes it possible to replace the maintained standards with the quantum Hall effect device. At the same time, for the second quantum effect used in metrology, that is, the Josephson effect, a conventional value  $K_{J-90}$  has also been assigned to the Josephson constant. The material standards for the volt and ohm are now replaced by the Josephson junction and the quantum Hall device; the outputs of these devices are independent of the site on earth and the shape or materials used for their construction. Intercomparisons are therefore no longer essential.

**Measurement Setup.** Comparisons of the Hall resistance with standard resistors can be performed in different ways: potentiometric or with a cryogenic current comparator (CCC). Figures 11 and 12 show generic and simplified diagrams representative for various measurement setups. In the potentiometric method (Fig. 11) a reference resistor  $R_s$  of the same nominal value of  $R_H(i)$  is used. As the plateau  $i = 1$  is difficult to obtain with reasonable fields, often the integers  $i = 2$  or  $i = 4$  are preferred. Superconducting magnets that generate an induction of 15 T between its poles are commonly used. The value  $I$  of the dc current source is in the 50  $\mu\text{A}$  range and this alleviates the work of the temperature controller because of the low power dissipation in  $R_s$ . Measurement of the voltage  $V_{H,i}$  permits inspection if the correct conditions for the occurrence

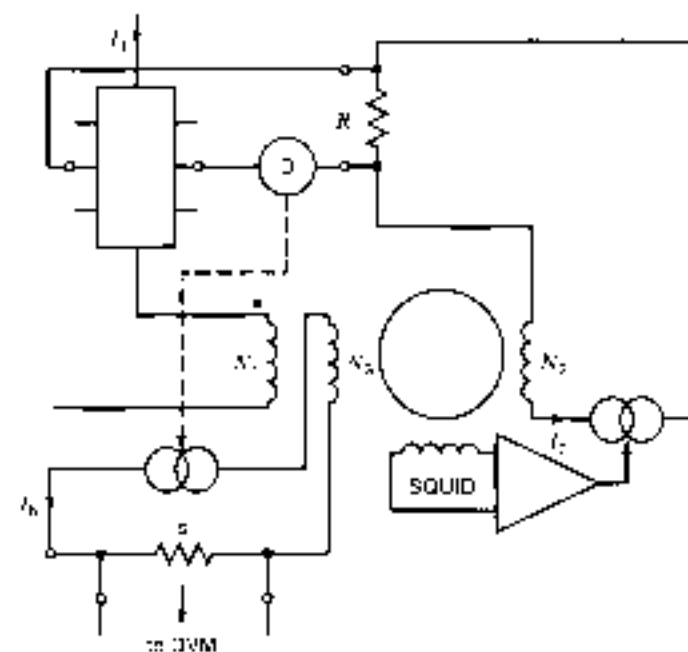


Figure 12. Cryogenic current comparator setup for comparing the quantified Hall resistance with the laboratory standard resistor.



of the quantum Hall effect are met. The Hall plateaus will appear when  $V_{xx}$  is very small. A zero detector is lacking and, instead, a high-gain amplifier followed by a digital voltmeter (DVM) is used. A compensating source  $V_c$  is derived from a potentiometer fed with a second current source  $I_1$ . By changing the position of  $S_2$ , two voltage differences are measured;  $A_1 = V_{H1} - V_c$  and  $A_2 = V_{H2} - V_c$ . From the  $\Delta$ -values measured it is possible to calculate the current  $I$  and  $V_c$  and thus  $V_H$ . To eliminate the drifts of the current sources and the thermal emfs it is necessary to perform measurements over a longer period, for example, an hour, and calculate averages. Because all current sources can change direction by turning over  $S_1$  the effect of thermal voltages can be eliminated. If the CCC is used as shown in the simplified circuit of Fig. 12, it is possible to compare  $R_H$  with smaller resistors, for example, 100  $\Omega$  because the turns ratio  $N_2/N_1$  can be varied. In Fig. 11, there are two balance conditions to be met:  $V_H = R_0 I_2$  for the zero detector balance, and  $N_1 I_1 = N_2 I_2$  for the zero flux condition in the core. The magnetic flux is measured by the very sensitive superconducting quantum interference device (SQUID), and the output of this device controls the current source  $I_2$  in such a way that the core flux becomes almost zero. If the balance of D is obtained by varying  $N_1$ . Then we find with  $V_H = R_0 I_1$ , the quantum Hall resistance  $R_H = R_0 N_1/N_2$ . In order to increase the precision to the level  $10^{-9}$  or better, a fixed ratio  $N_2/N_1$  is more appropriate. In this case a third current winding  $N_3$  is added to the comparator and a balance current  $I_3$ , derived from the detector output performs the automatic balancing. This balance current is measured by means of a shunt resistor  $s$  and a DVM. The balance current is much smaller than the main current and, therefore, the limited precision of the digital voltmeter does not impair the measurement accuracy.

If the two methods are compared (Ref. 5), it can be calculated for a specific case that the potentiometric method generates more system noise than the CCC. In a typical case the CCC method is therefore almost 45 times faster than the potentiometric method. As a result, the measurement time is reduced to a few minutes and the random uncertainty still is at the level  $2 \times 10^{-7}$ .

## BIBLIOGRAPHY

1. E. H. Hall, On a new action of the magnet on electrical currents, *Am. J. Math.*, **2**: 287-293, 1879.
2. K. von Klitzing, G. Dorda, and M. Pepper, *Physical Rev. Lett.* **45**: 494-497, 1980.
3. D. K. Schroder, *Semiconductor Device Characterization*, New York: Wiley, 1980.
4. G. Landwehr, The discovery of the quantum Hall effect, *Metrologia*, **22**: 119-127, 1985.
5. A. Hartland, The quantum Hall effect and the resistance standards, *Metrologia*, **20**: 175-190, 1982.
6. B. Jeckelmann and B. Jeanneret, Application de l'effet Hall quantique à la métrologie, Swiss Federal Office of Metrology info. (OMET info.) 4(2), 8-14, 1987.
7. B. Jeanneret et al., Comparison of the quantized Hall resistance in different GaAs/Al<sub>x</sub>Ge<sub>1-x</sub>As heterostructures, *IEEE Trans. Instrum. Meas.*, **40**: 231-235, 1991.
8. B. Jeckelmann, A. D. Inglis, and B. Jeanneret, Material, device and step independence of the quantized Hall resistance, *IEEE Trans. Instrum. Meas.*, **44**: 269-272, 1995.
9. B. Jeckelmann, A. D. Inglis, and B. Jeanneret, Are anomalous values of the quantized Hall resistance really anomalous? *Metrologia*, **33**: 489-502, 1996.
10. A. H. MacDonald and P. Streda, Quantized Hall effect and edge currents, *Phys. Rev. B*, **20**: 1616-1619, 1984.
11. B. Shapiro, Finite-size corrections to the quantum Hall effect, *J. Phys. C*, **19**: 4703-4721, 1986.
12. B. Jeanneret et al., Influence of the device-width on the accuracy of quantization in the integer quantum Hall effect, *IEEE Trans. Instrum. Meas.*, **44**: 254-257, 1995.
13. K. C. Lee, Bonding wires to quantized Hall resistors, *IEEE Trans. Instrum. Meas.*, **44**: 249-253, 1995.
14. B. Jeanneret and B. Jeckelmann, Influence of the voltage contacts on the four-terminal quantized Hall resistance in the nonohmic regime, *IEEE Trans. Instrum. Meas.*, **44**: 276-279, 1995.
15. A. D. Inglis and I. Minowa, Fabrication of precision quantized Hall devices, *IEEE Trans. Instrum. Meas.*, **46**: 205-207, 1997.
16. R. A. Smith, *Semiconductors*, Chapt. 5, Cambridge, UK: Cambridge University Press, 1959.
17. L. J. van der Pauw, A method of measuring specific resistivity and Hall effect of discs of arbitrary shape, *Phil. Res. Rep.*, **13**: 1-9, Feb. 1958.
18. C. L. Chien and C. R. Westgate, *The Hall Effect and Its Applications*, New York: Plenum, 1980.
19. E. Schwabald, Der Halleffekt und seine technische Anwendung, *Archiv für Technische Messen*, V013-2, 3, 1956.
20. E. Schwabald, *Halbleiter Bauelemente in der Messtechnik*, Berlin: VDE Verlag, 1961, pp. 201-213.
21. J. Janke, Putting the Hall effect transducer to work, *Int. Electron.*, 33-37, February 1965.
22. H. P. Baltes and R. S. Popovic, Integrated semiconductor magnetic field sensors, *Proc. IEEE*, **74**: 1107-1132, 1986.
23. E. J. Oliver, *Practical Instrumentation Techniques*, New York: Hayden Book Company, pp. 46, 106, 138, 249, 287-288, 1971.
24. A. Bidetti, G. Monreal, and R. Vio, Monolithic magnetic Hall sensor using dynamic quadrature offset cancellation, *IEEE J. Solid-State Circuits*, **32**: 839-846, 1997.
25. Z. L. Worsica, Hall effect feedback transducers and their application, *Acta Inche 1976*, Elsevier, 5th IMEKO Congress, London, 551-558, 10-14 May 1976.
26. J. Buchta, R. G. Dale, and N. R. C. Kiley, Sensor for magnetic bearings, *IEEE Trans. Magn.*, **20**: 2962-2964, 1983.
27. J. Spitzer and M. Perkins, Hall-effect works in two-state switches, *Electron. Design*, 125-127, Jul. 5, 1980.
28. D. G. Lampard and A. M. Thompson, A new theorem in electrostatics with application to entrainable standards of capacitance, *Nature*, **177**: 898, 1956.
29. D. G. Lampard, A new theorem in electrostatics, *J. IEE*, **104C**: 271-280, 1957.
30. W. K. Cluthier, A calculable standard of capacitance, *Metrologia*, **1**: 36-38, April 1965.
31. N. Fincková, An absolute determination of the ohm based on calculable standard capacitors, *ECROMES-72*, Precise Electrical Measurement Conference publication no. 152, IEE 1977: 53-57, 5-9 Sept. 1977.
32. G. H. Rayner, NPL calculable capacitor, *IEEE Trans. Instrum. Meas.*, **21**: 361-365, 1972.
33. B. W. Palley, *The fundamental physical constants and the fraction of consistency of*, Bristol: Adam Hilger, 1985.

## HIGH-ENERGY PHYSICS PARTICLE DETECTOR MAGNETS

### THE DEVELOPMENT OF DETECTOR SOLENOIDS

The discovery of type two superconductivity in 1961 (1) was celebrated by the particle physics community. Suddenly it appeared to be possible to create a large volume of magnetic field at an induction not heretofore considered to be economical using conventional magnets. In 1960, one of the largest operating particle detectors using a magnetic field was probably the 72 inch bubble chamber at Berkeley. Within days of the announcement of the discovery of niobium tin as a high field superconductor, particle physicists at Berkeley and other locations set up research groups to make superconducting magnets for particle detectors. There was excitement and constant communication between groups in both the United States and Europe. The cryogenic expertise for the development of the new superconducting magnets came from the people who had developed the cryogenic systems for hydrogen bubble chambers.

The hype and hopes of type two superconductivity soon faded with the realization that the niobium tin then available was not a usable superconductor for large magnets. Early on the experimental work shifted to Nb-Ti and Nb-Zr alloys. Before 1964, Nb-Zr was the alloy of choice because it was more stable than Nb-Ti. The alloy superconductors that performed well in short samples would reach only 30% or 50% of their critical current in a magnet (2). Wires plated with copper appeared to perform somewhat better than bare wire (3). Degradation due to flux jumps was the topic of the day. Improvements in magnet performance were not spectacular because there was no general understanding of what was happening within the superconductor. By 1964, the construction of large superconducting particle detector magnets appeared to be nearly hopeless.

#### Early Superconducting Detector Magnets

The paper on cryogenic stability of superconductors by Stekly and Zar (4) caused excitement in the particle physics community. The paper stated that if the superconductor was put in a low resistivity matrix, it didn't matter whether the superconductor flux jumped as long as the matrix remained at a temperature below the superconductor critical temperature. The discovery of cryogenic stability led to the first large detector magnets being built for particle physics. The first of these magnets was the 12 foot bubble chamber magnet at Argonne in 1969 (5,6). The 12 foot bubble chamber was followed by a 7 foot bubble chamber magnet at Brookhaven in 1970 (7), a 15 foot bubble chamber at the Fermilab 1973 (8), a 3.6 meter bubble chamber at CERN 1973 (8,9), the LASS magnet at SLAC 1974 (10), and a number of smaller devices. In 1972, M. Morpurgo tested a hollow conductor, forced-cooled, cryostable solenoid for the OMEGA experiment at CERN (11,12). Cryostability solved the scale problem for large superconducting magnets, but magnets built in this way operated at low current densities and were far from being thin from the standpoint of particle transmission through the magnet.

#### Low Mass Thin Detector Magnets

Truly thin detector solenoids required a superconductor that could operate at higher current densities without flux jumping.

Work by Bean et al. (13), Haneux (14), Chesler (15), and Smith et al. (16,17) paved the way to understanding the intrinsic stability of superconductors, which led to the development of modern, twisted multifilamentary conductors with a low matrix metal to superconductor ratio. Increasing the current density in the magnet winding was one way of making the magnet more transparent to particles.

In addition, thin superconducting solenoids had to be cooled in a different way. Helium bath cryostats contain too much material for them to be transparent to particles. In order to reduce the mass of the cryostat, it was found that thin solenoids had to be cooled indirectly by conduction to tubes that contain helium. Experimental work in the 1970s suggested that two-phase helium cooling would result in a lower operating temperature than supercritical helium cooling (18).

The first experiment calling for a thin solenoid was at the ISR at CERN (19). In 1975 a thin solenoid was proposed for the MINIMAG experiment proposed by the Lawrence Berkeley Laboratory in 1975 (20). This experiment required a one meter diameter solenoid that was 0.35 radiation lengths thick, including the cryostat. Two 1 m diameter test coils were built and tested in 1975 and 1976 (21,22). The conductor in the test coils was operated at matrix plus superconductor current densities as high as  $1250 \text{ A mm}^{-2}$ . The MINIMAG experiment was not built, but a larger detector for an experiment at PEP colliding beam ring at the SLAC was embarked upon. This detector required a clear bore diameter of 2 m with a gap of 3.3 m between the iron poles. A uniform 1.5 T induction (better than 1 part in 1000) within a 2 m diameter, 2 m long volume was required. The coil and cryostat had to be less than 0.7 radiation lengths thick so that calorimeters and muon detectors could be located outside of the magnet. Work began on a 2 m diameter test coil in late 1976. This coil was tested in 1977 and 1978 (23). The thin coil experimental work at Berkeley led directly to the CELLO detector at Cornell University (19) and the FEP-4 detector (24,25) at the PEP colliding beam facility at the SLAC.

A group at CEN Saclay outside Paris decided to build their detector magnet using a conductor that had a low copper to superconductor ratio soldered to very pure aluminum high residual resistance ratio (RRR) matrix. The advantages of the aluminum matrix were as follows: The minimum propagation zone was lengthened so that the energy needed to induce a quench in the magnet was increased by over three orders of magnitude, and the quench propagation velocity along the wire was faster than for a comparable copper matrix conductor. The 2 m diameter CELLO detector magnet was first tested in 1979 (26). A conductor made with the copper matrix superconductor mixed into pure aluminum (RRR > 1000) was developed in a number of locations at about the same time (27). This type of conductor was used on the CDF detector magnet at Fermilab (28), on the VENUS detector (29), the TOPAZ (30) detector, and the AMY detector (31) at KEK in Japan, on the ALEPH (32) and DELPHI (33) detectors at CERN, on the H-1 (34) and ZEUS (35) detectors at DESY in Germany, on the GSI solenoid (36) at Darmstadt, on the CLEP-2 detector (37) at Cornell University, and on the CLEF detector at Frascati. A thin solenoid SSC experiment test coil was tested before the SSC was canceled (38). The Japanese flew a 1 m diameter helium solenoid (39,40) that used a low matrix to superconductor ratio (RRR  $\geq 1000$ ) aluminum matrix conductor to achieve a very low radiation thickness (about 0.25 radiation lengths) for a cosmic ray experiment. Detector solenoids for the Harker (41) experiment at the

Table 1. Parameters for Various Thin Superconducting Detector Solenoids

| Magnet        | Central Induction (T) | Warm-Bore Diameter (m) | Cryostat Length (m) | Matrix Material | Conductor Location | Radiation Thickness (Had. Len) | Stored Energy (MJ) | Matrix $J$ (A-mm <sup>2</sup> ) | Type of Cooling |
|---------------|-----------------------|------------------------|---------------------|-----------------|--------------------|--------------------------------|--------------------|---------------------------------|-----------------|
| CLEO-1        | 1.5                   | 2.0                    | 3.7                 | Cu              | Outside            | 0.7                            | 10.0               | ~350                            | Forced          |
| PEP-4         | 1.5                   | 2.04                   | 3.84                | Cu              | Outside            | 0.83                           | 10.9               | 645                             | Forced          |
| CELLSI        | 1.3                   | 1.5                    | 4.02                | Al              | Outside            | 0.6                            | 5.0                | ?                               | Forced          |
| CDF           | 1.5                   | 2.55                   | 5.4                 | Al              | Inside             | 0.84                           | 30                 | 64                              | Forced          |
| TOPAZ         | 1.2                   | 2.72                   | 5.4                 | Al              | Inside             | 0.70                           | 20                 | 56                              | Forced          |
| VENUS         | 0.75                  | 3.4                    | 5.6                 | Al              | Inside             | 0.52                           | 12.0               | ?                               | Forced          |
| ALEPH         | 1.5                   | 4.96                   | 7.0                 | Al              | Inside             | 1.6                            | 136                | 30.6                            | Natural         |
| AMY           | 3.0                   | 2.39                   | 2.11                | Al              | Outside            | >2                             | 40                 | 50                              | Pool            |
| GSI           | 0.6                   | 2.4                    | 3.3                 | Al              | Inside             | ~1.0                           | 3.4                | -                               | Natural         |
| ZEUS          | 1.8                   | 1.72                   | 2.9                 | Al              | Inside             | >2                             | 16                 | ?                               | Forced          |
| DELPHI        | 1.2                   | 5.2                    | 7.4                 | Al              | Inside             | 1.7                            | 103                | 46.3                            | Forced          |
| B I           | 1.2                   | 5.2                    | 6.0                 | Al              | Inside             | 1.6                            | 150                | 46                              | Forced          |
| CLEO-2        | 1.5                   | 2.0                    | 3.9                 | Al              | Inside             | 2.2                            | 25                 | 41.3                            | Natural         |
| e-3           | 1.45                  | 1.17                   | 0.18 <sup>a</sup>   | Al              | Inside             | >2                             | 5.5                | 91.8                            | Forced          |
| KEK Barrel    | 1.2                   | 0.852                  | 2.0                 | Al              | Inside             | 0.23                           | 0.815              | 261                             | Pool at End     |
| SDC Test Coil | 1.5                   | 1.7                    | 2.4                 | Al              | Inside             | 1.2                            | 45                 | 63.4                            | Forced          |
| BaBar         | 1.6                   | 2.76                   | 3.85                | Al              | Inside             | <1.4                           | 23                 | 27 & 67                         | Natural         |
| CMS           | 4.0                   | 6.0                    | 12.5                | Al              | Inside             | >2                             | 2500               | 15.4                            | Natural         |

<sup>a</sup>Beam orbit diameter, outer solenoid coil diameter = 15.1 m, inner solenoid coil diameter = 13.4 m. Total gap between the iron poles (the inner return path is C-shaped, the total gap in the iron is about 0.23 m).

B factory at SLAC, the ATLAS (42.43) toroidal magnet detector at the LHC, and the CMS (44) solenoidal detector at the LHC are currently under development or construction. All of these magnets will use a pure aluminum matrix superconductor that will be wound on the inside of a hard aluminum support structure.

The use of thin solenoid magnet construction techniques has proven to be less costly even when thinness was not required. As a result, the thin detector solenoid construction techniques were used to build two 13.4 m diameter and one 16.1 m diameter solenoid for the g-2 experiment at the Brookhaven National Laboratory (45). These solenoids were successfully tested to full field in the summer of 1996 (46). Table 1 summarizes the design parameters for a number of the thin superconducting detector magnets.

## THE DEFINING PARAMETERS FOR THIN SOLENOIDS

In the literature, thickness is defined in terms of interaction lengths, absorption lengths, and radiation lengths. In high-energy physics detectors, there is no one universal definition of thickness. Thus, discussion of interaction lengths must identify the particle, and absorption lengths must identify the particle and its energy. The most common definition of thickness uses radiation lengths as a defining parameter. One radiation length occurs when 63.2% (1 - 1/e) of the neutral particles have formed charged particle pairs. This definition is appropriate in many experiments because the calorimeters and muon detectors are the only detectors that are located outside the magnet.

The physical thickness of a material that is one radiation length thick is a function of the material atomic number  $Z$  and the material specific density  $\gamma$ . In order for a superconducting magnet to be thin, it must be made from low-density, low- $Z$  materials. The radiation thickness of a detector magnet is the sum of the radiation thicknesses of the windings, the coil support structure, the cryostat, and the intermediate temperature shields.

The radiation thickness  $X_0$  of a magnet component can be estimated using the following expression:

$$X_0 = \frac{t}{L_r \cos(\alpha)} \quad (1)$$

where  $X_0$  is the radiation thickness of the magnet component (given in radiation lengths),  $t$  is the physical thickness of the material in the magnet component,  $L_r$  is the thickness for one radiation length of the material in the magnet component, and  $\alpha$  is the particle angle with respect to a line perpendicular to the component. In most cases, radiation thickness is defined when  $\alpha = 0$ .

The value of  $L_r$  used in Eq. (1) can be obtained from Table 2 (47,48) or it can be estimated using the following expression:

$$L_r = 158 \frac{Z^{0.73}}{\gamma} \quad (2)$$

where  $Z$  is the atomic number for the heaviest element in the compound that makes up the component and  $\gamma$  is the mass specific density for the material in the component. For pure elements, Eq. (2) yields a good estimate of  $L_r$ , except for ordinary hydrogen, which has no neutrons in its nucleus. For components made from compounds, the use of the  $Z$  for the heaviest element in the compound will tend to overestimate radiation thickness whereas using an average value of  $Z$  will often underestimate the radiation thickness. For components made from alloys or composites, the method of mixtures can be applied to achieve a good estimate of  $L_r$ .

## THIN DETECTOR SOLENOID DESIGN CRITERIA

The strategy for minimizing the radiation thickness of a superconducting detector magnet requires the following steps: (1) Massive parts such as current bus bars, gas-cooled electrical leads, cold mass support structures, vacuum services, and

Table 2. The Radiation Thickness of Various Materials

| Material                               | Z  | Mass Density<br>( $\text{kg m}^{-3}$ ) | Due Radiation Length $L_A$ |       |
|--|----|--|----------------------------|-------|
|  |    |  | $\text{kg m}^{-2}$         | mm    |
| Pure elements                          |    |  |                            |       |
| Hydrogen                               | 1  | 70.9                                   | 436.5                      | 4,800 |
| Deuterium                              | 1  | 162                                    | 1,261.0                    | 7,640 |
| Helium                                 | 2  | 125                                    | 943.2                      | 7,530 |
| Lithium                                | 3  | 534                                    | 427.0                      | 1,550 |
| Beryllium                              | 4  | 1,848                                  | 651.9                      | 353   |
| Boron                                  | 5  | 2,370                                  | 553.9                      | 254   |
| Carbon                                 | 6  | 1,550                                  | 427.0                      | 275   |
| Nitrogen                               | 7  | 808                                    | 375.0                      | 470   |
| Oxygen                                 | 8  | 1,142                                  | 344.6                      | 302   |
| Neon                                   | 10 | 1,207                                  | 299.4                      | 240   |
| Magnesium                              | 12 | 1,740                                  | 254.6                      | 146   |
| Aluminum                               | 13 | 2,700                                  | 246.1                      | 89.8  |
| Argon                                  | 18 | 1,400                                  | 195.5                      | 110.0 |
| Titanium                               | 22 | 4,740                                  | 159.7                      | 27.2  |
| Chromium                               | 24 | 7,200                                  | 140.7                      | 20.4  |
| Iron                                   | 26 | 7,870                                  | 136.4                      | 17.6  |
| Nickel                                 | 28 | 6,902                                  | 131.9                      | 14.9  |
| Copper                                 | 29 | 6,060                                  | 126.6                      | 14.3  |
| Niobium                                | 41 | 6,570                                  | 101                        | 11.8  |
| Tin                                    | 50 | 7,310                                  | 99.6                       | 12.1  |
| Tungsten                               | 74 | 19,300                                 | 67.6                       | 3.5   |
| Lead                                   | 82 | 11,350                                 | 60.5                       | 3.0   |
| Uranium                                | 92 | 18,950                                 | 40.0                       | 2.2   |
| Compounds, Alloys, and Other Materials |    |  |                            |       |
| Water                                  |    | 1,000 <sup>a</sup>                     | 360.6                      | 360.4 |
| Polyethylene                           |    | ~950                                   | 447.8                      | 470   |
| Epoxy resin                            |    | 1,450                                  | 406                        | 290   |
| Glass fiber epoxy                      |    | 1,750                                  | 330                        | ~169  |
| Carbon fiber epoxy                     |    | 1,800                                  | ~118                       | ~261  |
| Aluminum nitride (45% B)               |    | 2,550                                  | ~381                       | ~149  |
| Mylar                                  |    | ~1,380                                 | 395.5                      | 267   |
| Sodium iodide                          |    | 3,670                                  | 94.9                       | 25.9  |
| Lithium fluoride                       |    | 2,640                                  | 392.5                      | 140   |
| 304 Stainless steel                    |    | 7,900                                  | 147.0                      | 17.4  |
| Nb-17% Ti                              |    | 6,520                                  | 132.6                      | 20.4  |

<sup>a</sup>Liquid state<sup>b</sup>Graphite or carbon fiber.

cryogenic services should be located at the ends of the magnet away from the region that is supposed to have a minimum radiation thickness. (2) The superconductor should have a minimum amount of copper and niobium titanium. The stabilizer matrix material for the superconductor should be made of a low-resistivity, low-Z material such as ultrapure aluminum. (3) The support structure on the outside of the coil, which will carry the hoop forces in the solenoid, should be made from a strong, ductile, low-Z, low-density material with a high thermal conductivity. (4) The magnet should be cooled indirectly with helium in tubes that are attached to the coil support structure. (5) Intermediate temperature shields for the cryostat should be made of a low-Z, low-density, high thermal conductivity material such as aluminum. (6) The inner cryostat vacuum vessel should be made from a strong low-Z, low-density material. (7) The outer cylinder of the cryostat vacuum vessel should be made from a material with a low-Z, a low density, and an elastic modulus that is reasonably high.

The typical physics detector solenoid is usually between two unsaturated iron poles that have an average relative permeability that is greater than 20. The magnetic flux generated by the solenoid winding is returned by an iron yoke that carries the magnetic flux from one pole to the other. The relative

permeability of the iron in the return yoke is usually above 50. Figure 1 shows a typical thin detector solenoid within an experiment located around the collision point of colliding beam storage ring. The number of ampere turns needed to generate a uniform magnetic induction within the detector solenoid can be estimated by using the following expression (49):

$$NI = \frac{B_p L_y}{\mu_0} \quad (3)$$

where  $NI$  is the total number of ampere turns in the detector solenoid coil needed to generate a magnetic induction  $B_p$  in a solenoid that has unsaturated iron poles that are a distance  $L_y$  apart.  $\mu_0$  is the permeability of air ( $\mu_0 = 4\pi \times 10^{-7} \text{ H m}^{-1}$ ).

Equation (3) underestimates the ampere turns needed to generate the magnetic induction in the solenoid bore anywhere from 3% to 30% depending on the design of the magnetic circuit and the central induction within the solenoid. The equation underestimates the required ampere turns because the relative permeability of the iron in the poles and the return yoke is not infinite and the iron in the pole pieces is often segmented, with detectors between the segments. Often the extra ampere turns are put at the ends of the solenoid so that the desired

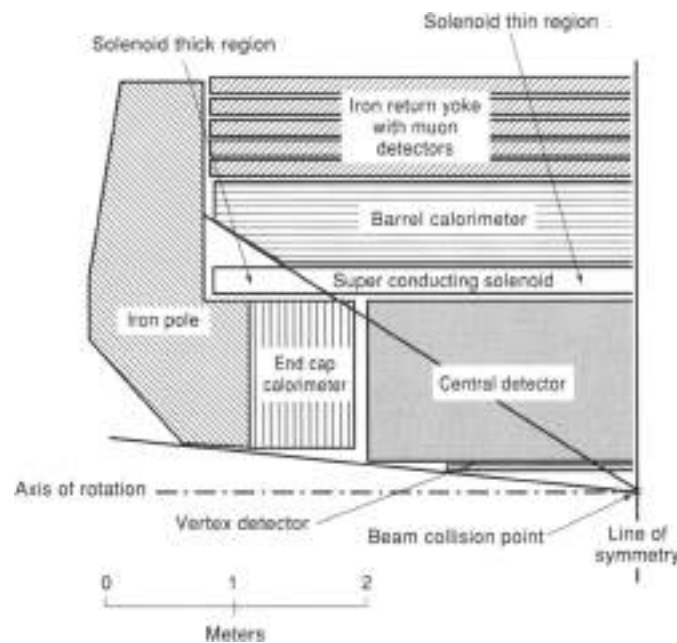


Figure 1. A quarter-section view of a typical calibrating beam physics detector with the solenoid crystal shown. (The thin and thick sections of the superconducting magnet are shown.)

field uniformity within the solenoid can be achieved. Computer codes such as BUISSON (50) and OPERA2D (51) can be used to determine the number of ampere-turns needed to generate the desired central induction and the desired field uniformity within the detector volume.

The amount of superconductor needed to generate the magnetic field is quite small. When Nb-Ti with a critical current density of 2500 A/mm<sup>2</sup> at 4.2 K and 5.0 T is used, about 0.3 mm of Nb-Ti is needed for every tesla of central magnetic induction produced (52). The copper to superconductor ratio for the conductor can be as low as 0.8. The amount of stabilizer (usually annealed 0.99999 pure RRR = 1000 aluminum) in the conductor is dictated by the type of quench protection chosen.

The physical thickness of the superconducting coil is determined by the thickness of stabilizing matrix material in the conductor. The average conductor current density  $J_c$  is determined by the safe quench condition for the coil. For safe magnet quenching through a dump resistor, the magnet  $E_0 J_c^2$  limit can be estimated using the following expression (53,54):

$$E_0 J_c^2 = V I_0 F^2 T_m \int \frac{r}{r+1} \quad (14)$$

where  $E_0$  is the magnet stored energy when it is operated at its design current  $I_0$ ,  $V$  is the discharge voltage for the magnet during the quench (for large magnets  $V$  is limited to about 500 V),  $I_0$  is the magnet design current ( $I_0$  is typically greater than 3000 A),  $r$  is the matrix to superconductor ratio, and  $F T_m$  is the integral of  $J_c^2 dr$  needed to raise the stabilizer adiabatic hot spot temperature from 4 K to a maximum hot spot temperature  $T_m$ . (For RRR = 1000 aluminum,  $F T_m = 6 \times 10^{10} \text{ A}^2 \text{ m}^{-2}$  when  $T_m = 300 \text{ K}$ .)

The magnet stored energy  $E_0$  can be estimated if one knows the solenoid coil diameter  $D_c$ , the central induction  $B_0$ , and the gap between the iron poles  $L_c$ . An approximate expression for

the magnet stored energy is as follows:

$$E_0 = \frac{\pi D_c^2 B_0^2 L_c}{8 \mu_0} \quad (15)$$

If the  $E_0 J_c^2$  limit for the magnet is increased, then the magnet design current  $I_0$  or the magnet discharge voltage  $V$  must be increased as well. Quench back from the coil support structure can be helpful in improving the quench protection for the magnet. Magnets that employ quench back (22,55) as the primary means for quench protection can be operated at a much higher  $E_0 J_c^2$  limit, but the typical solenoid that is protected with a dump resistor across the leads has the  $E_0 J_c^2$  limit given by Eq. (14).

From Eqs. (14) and (15), one can determine the thickness of the superconducting coil  $t$ , using the following expression:

$$t = \left( \frac{\pi D_c^2 B_0^2 I_0^2}{8 J_c^2 V L_c F^2 T_m \left( \frac{r}{r+1} \right)} \right)^{1/2} \quad (16)$$

In order for the coil thickness to be thinner than the value given by Eq. (16), quench back must turn the whole coil normal in a time that is significantly faster than the  $L$  over  $R$  time constant of the coil and dump resistor circuit.

In virtually all of the large detector solenoids, the superconducting coil is wound inside the support cylinder (56). When the coil is inside the support cylinder, the joint between the coil and the support structure is in compression as the magnet is charged. An additional advantage is that the coil package cools down from the outside. Thus, the support cylinder shrinks over the coil. A few of the smaller detector magnets were wound with the coil on the outside of a hobbin or support cylinder. In all these cases, the conductor was designed to carry all of the magnetic hoop forces and the helium cooling tubes were attached to the outside of the coil. The superconducting solenoid coil can be wound in one or two layers. A two-layer coil has the advantage of having both leads from the coil come out at the same end of the coil package. There are a number of accepted ways of winding coils so that they have more current per unit length at the solenoid ends than in the center. One approach is to make the matrix current density higher at the ends by making the conductor thinner along the coil axis.

The thickness of the support shell outside the superconducting coil is governed by the magnetic pressure on the coil windings (49). Total strain of the coil should be limited to prevent plastic deformation of the conductor matrix. If the conductor has a pure aluminum matrix, the strain limit for the coil should be set to about 0.1% (57,58). A conservative view assumes that virtually all of the magnetic forces are carried by the support shell, and the calculated shell thickness is given by the following expression:

$$t_s = 250 \frac{B_0^2 D_c}{\mu_0 E_s} \quad (17)$$

where  $t_s$  is the design thickness for the support shell,  $D_c$  is the inside diameter of the support shell, and  $E_s$  is the modulus of elasticity of the material in the support shell. If the superconductor is included in the overall strain calculation, the thickness of the support shell can be reduced.

Table 3. A Comparison of Four Thin Solenoids

| Component                                   | Case 1 | Case 2 | Case 3 | Case 4 |
|---|--------|--------|--------|--------|
| Central induction $B_z$ (T)                 | 1.5    | 1.5    | 0.75   | 1.5    |
| Solenoid coil diameter (m)                  | 2.0    | 4.0    | 4.0    | 4.0    |
| Gap between the iron poles (m)              | 3.3    | 6.6    | 6.6    | 3.3    |
| Length of the solenoid thin section (m)     | 3.3    | 6.6    | 6.6    | 3.3    |
| Cryostat inside diameter (m)                | 1.84   | 3.80   | 3.80   | 3.80   |
| Cryostat outside diameter (m)               | 2.24   | 4.29   | 4.22   | 4.26   |
| Cryostat overall length (m)                 | 3.85   | 7.30   | 7.30   | 3.85   |
| Magnet ampere turns (kAt)                   | 3.94   | 7.88   | 2.94   | 3.94   |
| Magnet stored energy (MJ)                   | 0.29   | 74.25  | 18.56  | 37.13  |
| Magnet design current (A)                   | 5000   | 5000   | 5000   | 5000   |
| Magnet self inductance (H)                  | 0.74   | 5.04   | 1.48   | 2.97   |
| Number of conductor layers                  | 2      | 2      | 2      | 2      |
| Number of coil turns                        | 738    | 1576   | 796    | 798    |
| Quench discharge voltage (V)                | 500    | 500    | 500    | 500    |
| Matrix current density (A/mm <sup>2</sup> ) | 127.1  | 61.9   | 89.9   | 69.6   |
| Nb-Ti plus copper thickness (mm)            | 0.90   | 0.90   | 0.45   | 0.90   |
| Total coil thickness (mm)                   | 9.39   | 20.58  | 13.26  | 16.75  |
| Coil support structure thickness (mm)       | 12.97  | 25.95  | 6.49   | 25.95  |
| Inner cryostat thickness (mm)               | 1.24   | 2.57   | 2.57   | 2.57   |
| Outer cryostat thickness (mm)               | 13.11  | 25.23  | 25.23  | 19.12  |
| Magnet radiation thickness (Rad Len)        | 0.495  | 0.971  | 0.589  | 0.829  |
| Magnet cold mass (metric tons)              | 2.12   | 14.4   | 6.18   | 7.00   |
| Magnet overall mass (metric tons)           | 4.00   | 23.5   | 15.3   | 11.8   |

The coil cryostat is primarily the vacuum vessel that provides the insulating vacuum for the magnet. The two primary cryostat elements are the outer cryostat vacuum vessels and the warm bore tube. The multilayer insulation and shields make up only a minor part of the cryostat's radiation thickness. A design thickness of a solid outer cryostat wall can be calculated using the following expression, which has been derived from the equation for elastic buckling of a cylinder under external pressures (58,60):

$$t_0 = 1.04 \left( \frac{P_0 L_0 D_0^3}{E_0} \right)^{1/4} \quad (8)$$

where  $t_0$  is the thickness of the outer cryostat wall,  $P_0$  is pressure on the outer wall of the cryostat (usually  $P_0 = 1 \text{ atm} = 1.013 \times 10^5 \text{ Pa}$ ),  $L_0$  is the length of the thin unsupported section of the outer cryostat wall,  $D_0$  is the diameter of the outer cryostat wall, and  $E_0$  is the elastic modulus of the material in the outer wall of the cryostat.

The minimum thickness of the inner wall of the cryostat can be derived if one knows the design ultimate stress for the material in the inner wall (58,61). The margin of safety normally applied to a pressure vessel wall, such as the cryostat inner wall, is usually four (62). An expression for the minimum inner cryostat wall thickness is given as follows:

$$t_1 = 2 \frac{P_1 D_1}{\sigma_1} \quad (9)$$

where  $t_1$  is the minimum wall thickness for the inner cryostat wall,  $P_1$  is the design internal pressure on the inner cryostat wall,  $D_1$  is the diameter of the inner cryostat wall, and  $\sigma_1$  is the ultimate stress for the material used in the inner cryostat wall. Sometimes, the cryostat inner wall thickness is greater than the thickness given by Eq. (9) so that one can mount particle detectors and other equipment on this wall.

The material thicknesses calculated using Eqs. (6)–(9) can be used to estimate the radiation thickness of the detector solenoid. Table 3 compares four cases where the coil diameter, the gap between the iron poles, and the central induction are varied. In all four cases, the cryostat walls and coil support structure are made from solid aluminum. The superconductor is Nb-Ti with a thick aluminum stabilizer. The assumed insulation system inside and outside the cold mass consists of 60 layers of aluminumized mylar and netting with a single 1 mm thick aluminum shield on other side of the coil. Figure 2 shows a cross section of a coil and cryostat for CASE 2 given in Table 3. In order to make a significant reduction in the radiation thicknesses shown in Table 3, quench back must be the primary mode of quench protection and the outer cryostat vacuum vessel must be made from a cellular (honeycomb) composite structure that is physically thicker than a solid aluminum vessel (63,64).

## MAGNET POWER SUPPLY AND COIL QUENCH PROTECTION

The power supply parameters are set by the coil charge time  $t_c$  and the design operating current  $I_c$  for the solenoid. The charge time for a detector solenoid is rarely an issue. Charge times as long as one hour are acceptable. The charge voltage  $V = L_1 di_c/dt$ , where  $L_1$  is the self-inductance of the magnet circuit, and  $di_c/dt$  is the magnet current charge rate. For a typical magnet,  $di_c/dt = I_c/t_{ch}$ . To determine the power supply voltage, one must add the  $IR$  voltage drop across the gas-cooled electrical leads and the cables connecting the power supply to the magnet. In addition, a voltage drop of 0.9 V should be allocated to the power supply back wheeling diodes and a current shunt.

### The Quench Protection Dump Resistor

Most large detector magnets are protected by a dump resistor across the gas-cooled electrical leads. When a quench is detected, the power supply is disconnected and the dump resistor

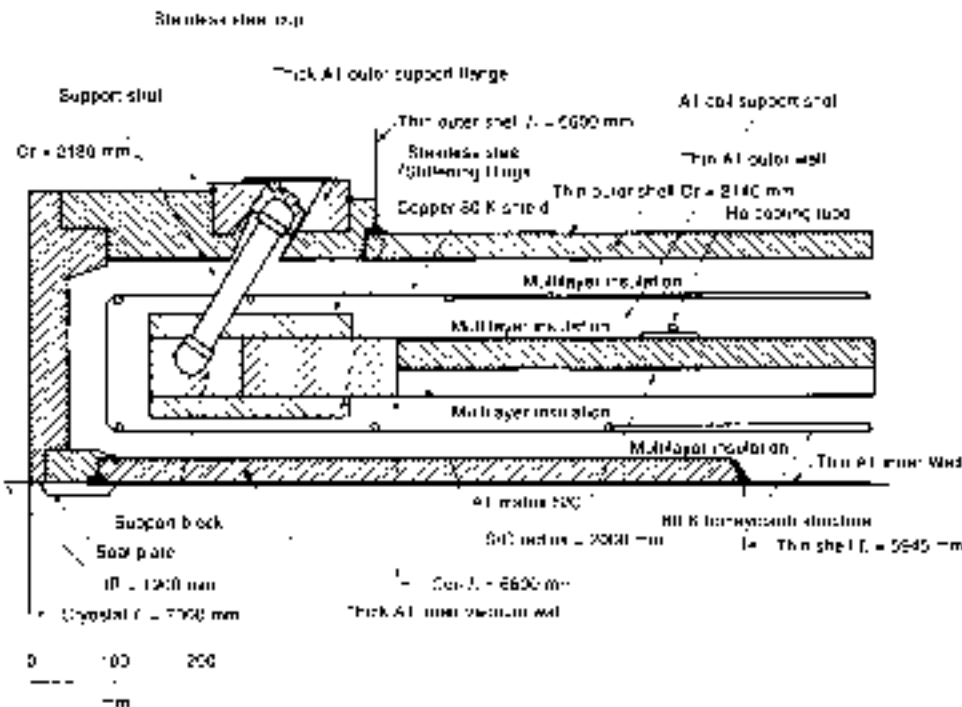


Figure 2. A cross section through the end of a 1.5 T thin solenoid with a 4.0 m coil diameter. A self-centering support strut is shown along with the stiff end ring for the superconducting coil package. (See Case 2 in Table 3.)

is put across the leads. The design of a magnet dump resistor circuit is determined by the following relationship (54):

$$F(T_n) = \int_0^{T_n} j(t)^2 dt = \frac{r}{(r-1)} \int_{T_0}^{T_n} \frac{C(T)}{\rho(T)} dT \quad (10)$$

where  $j(t)$  is the current density in the magnet superconductor cross section as a function of time  $t$ ,  $C(T)$  is the superconductor volume specific heat as a function of temperature  $T$ ,  $\rho(T)$  is the superconductor matrix material electrical resistivity as a function of temperature, and  $r$  is the ratio of matrix material to superconductor in the magnet conductor.  $T_0$  is the starting temperature of the magnet (about 4 K), and  $T_n$  is the maximum allowable hot spot temperature for the magnet conductor (usually 300–350 K). For a conductor with a very pure aluminum matrix with an RRR of 1000, the value of  $F(T_n)$  is around  $6.0 \times 10^{16} \text{ A}^2 \text{ m}^{-2} \text{ s}$  when  $T_n$  is 300 K.

When the magnet is discharged through a dump resistor, the current decay is exponential with a decay time constant  $\tau$  ( $\tau = L_0/R_0$ , where  $R_0$  is the resistance of the external dump resistor). The value of  $F(T_n)$  at the magnet coil hot spot is given as follows:

$$F(T_n) = j_0^2 \frac{(r+1)}{r} \left( \frac{1}{2} + t_{d0} \right) \quad (11)$$

where  $t_{d0}$  is the time needed to detect the quench and switch the resistor across the magnet coil (in most cases  $t_{d0}$  is less than one second) and  $j_0$  is the starting current density in the coil superconductor plus matrix material ( $I_0$  divided by the conductor cross-sectional area). If a constant resistance dump resistor is used, the value of the resistance  $R_0$  that results in a hot spot temperature less than or equal to  $T_n$  can be expressed

as follows:

$$R_0 \geq \frac{j_0^2 (r+1)}{2F(T_n) r} L_0 \quad (12)$$

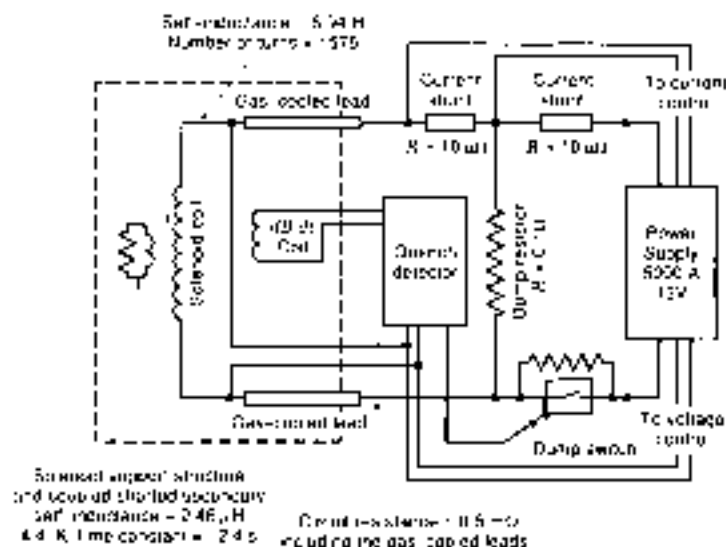
The design value of  $R_0$  should be larger than the value calculated by Eq. (12). For a constant resistance dump resistor, the maximum discharge voltage across the leads  $V = R_0 I_0$  will occur when the dump resistor is just put across the magnet.

Figure 3 shows a circuit diagram of the coil, its power supply, and the magnet dump circuit. A quench detection system is also shown. The values for inductances and  $R_0$  given in Fig. 3 would apply to CASE 2 in Table 1. The quench detection system shown in Fig. 3 compares the voltage across the superconducting coil with the  $dB/dt$  voltage due to changes in flux in the coil. If a voltage is measured across the coil and there is no corresponding  $dB/dt$  voltage, there is a normal region in the coil. The normal region detected by the quench detector will open the switch, putting the dump resistor across the coil. Other methods can also be used to detect short normal sections within a magnet (65).

#### The Role of Quench Back

It has been observed in most of the thin detector solenoid magnets that when the dump resistor is put across the electrical leads, the entire magnet becomes normal through the process of "quench back" (22). Quench back ensures that the coil current will decay faster than is predicted by the  $L$  over  $R_0$  time constant. As a result, the magnet hot spot temperature is reduced.

There is a maximum time  $t_q$  before which quench back must occur in order to have quench back be a fail-safe method of

Mutual inductance = 0.78  $\mu\text{H}$ 

**Figure 8.** A schematic circuit diagram for the coil, power supply and quench protection circuit for a large detector solenoid with a support cylinder (See Case 2 in Table 3.)

quench protection (66,67). If the resistance of the external resistor induces quench back in a time less than the time  $t_q$ , then the hot spot temperature is less than  $T_m$ , the maximum allowable hot spot temperature. The maximum allowable quench-back time  $t_q$  is the quench-back time required for fast safe quenching  $t_f$  minus the time required for a heat pulse to cross the insulation between the quench back circuit (usually the coil support structure) and magnet coil  $t_H$ . (For a layer of ground plane insulation that is two millimeters thick,  $t_H$  varies from 0.3 to 0.5 seconds depending on a number of factors.) For large detector solenoids,  $t_H$  is usually small compared  $t_f$ . The value of  $t_f$  can be determined using the following expression:

$$t_f = \frac{r}{r-1} \left[ \frac{F(T_m) - F(T_c)}{I^2} \right] \quad (13)$$

where  $r$  and  $j_H$  are previously defined.  $F(T_m)$  and  $F(T_c)$  are defined by the right-hand term in Eq. (10) for the maximum hot spot temperature  $T_m$  and the maximum temperature the coil would go to if the entire coil quenched instantaneously  $T_c$ . The value of  $T_c$  depends on how the stored energy of the magnet is split between the hard aluminum support shell and the coil. For detector solenoids it is usually safe to split the magnet stored energy between the support shell and the coil according to their masses.

Once  $t_H$  has been determined, it is possible to calculate the resistance of an external resistor needed to cause the coil to quench back from the support tube in a time that is less than  $t_q$ . The minimum resistance needed for quench back  $R_m$  can be calculated using the following relationship for a solenoid coil that is well coupled inductively to its quench-back circuit (66):

$$R_m = \left( \frac{L_1 N_2 A_2}{L_2 N_1 I} \right) \left( \frac{\Delta H_2}{\rho_2 t_q} \right)^{1/2} \quad (14)$$

where  $L_1$  is the magnet coil self-inductance,  $N_2$  is the number of turns in the quench-back circuit ( $N_2 = 1$  when the support shell is the quench back circuit),  $N_1$  is the number of turns in the magnet coil,  $A_2$  is the cross-sectional area of the quench back circuit,  $I_0$  is the coil current,  $\Delta H_2$  is the enthalpy change

per unit volume needed to raise the quench-back circuit temperature from 4 K to 10 K (for aluminum,  $\Delta H_2 = 13,200 \text{ J m}^{-3}$ ),  $\rho_2$  is the resistivity of the quench-back circuit material, and  $t_q$  is the  $L$  over  $R$  time constant for the quench back circuit (the support tube).

If the minimum quench-back resistance  $R_m$  is less than the resistance of the quench protection resistor  $R_q$ , quench back will always occur during a magnet dump. Therefore, the hot spot temperature of the coil is lower than  $T_m$ . When quench back is present, one can use a varistor (a resistor where the voltage across the resistor is nearly independent of current) as a dump resistor to speed up the quench process without increasing the coil voltage during the quench (68).

#### DESIGN CRITERIA FOR THE ENDS OF A DETECTOR SOLENOID

The previous sections have dealt primarily with the center section of a detector solenoid. Much of the engineering for a detector solenoid is in the ends of the magnet, where thickness is not an issue. For example: (1) The support system for the solenoid end mass is attached to the ends of the magnet. (2) The outside ends of the cryostat vessel are often where physical connections are made between the magnet and the rest of the detector. (3) The current leads and voltage taps into the coil will come out of the coil package at its ends. Gas-cooled electrical leads that connect the coil to the room-temperature outside world may also be located inside the magnet insulating vacuum vessel in the end region. (4) Cryogenic cooling is usually fed into the solenoid coil from the ends. Cooling should also include the intermediate temperature fluid (either liquid nitrogen at 80 K or helium gas at 50 to 80 K) used to cool the shields. (5) cryostat vacuum pumping ports will be located at the ends of the solenoid. (6) Room-temperature feed throughs for voltage taps, quench detection coils, temperature sensors, and pressure transducers will enter the magnet at the ends.

#### Cold Mass Support System

The cold mass supports to room temperature must carry gravity forces, seismic forces, magnetic forces, and shipping forces.



Most detector solenoids are designed to be at a neutral magnetic force point when the coil is at its operating temperature, so the cold mass support system must have a spring constant that is higher than the magnetic force constant.

Solenoids that are surrounded by iron are usually, but not always, in stable equilibrium in the radial direction and in unstable equilibrium in the axial direction. In the torsional direction (about the solenoid axis), there are almost no magnetic forces in a well-built solenoid, although asymmetric holes in the iron can introduce some of these forces. Stable equilibrium indicates that the magnetic forces will act in a direction that reduces a placement error; unstable equilibrium indicates that the magnetic forces will act in a direction that increases the placement error. In the direction of stable equilibrium, the spring constant of the support system is not a critical issue except when determining how the magnet responds to vibration. In the direction of unstable equilibrium (usually the axial direction) the spring constant of the support system must be larger than the force constant for the magnet at its maximum design field. In general, the magnetic force constant is linear with the location error and it increases with the magnet current squared. The magnetic force constant is a function of the design of the coil, the iron return yoke, and the pole pieces. Magnetic force constants will change as the iron in the magnetic circuit saturates.

Two types of cold mass support systems are commonly used in detector solenoids. The first is the self-centering support system where the position of the center of the solenoid coil does not change during the magnet cooldown or as the magnet is powered. The second support system carries axial forces with push-pull rods at one end of the magnet while the radial forces are carried by gravity support rods at both ends of the magnet. Both types of support systems must be designed to handle magnet shrinkage during the cooldown. A coil that is 6.6 m long and 4 m in diameter will shrink almost 28 mm in the axial direction and the radius will decrease about 8.4 mm. The external cryostat support system should be in line with the cold mass support system in order to avoid bending within the cryostat. The spring constant for the combined internal and external support systems must be greater than the magnetic force constant.

The self-centering support system has several advantages: (1) the position of the magnet center is the same both warm and cold. The PEP-4 solenoid magnetic center changed less than 0.4 mm during the magnet cooldown. (2) The radial and axial supports can be combined using either tension or compression rods. Two of the rods can also carry the torsional forces about the solenoid axis (torsional forces). The angle of the support rods can be set so that rod stress is not changed during the coil cooldown. (3) Since the axial spring constants must be high, the spring constant will be high in all directions. The self-centering support system will have a relatively high first mode vibration frequency. (4) The self-centering support system is robust in all directions, so earthquake and transportation forces should not be a problem.

The two disadvantages of the self-centering support system are as follows: (1) As the magnet coil cools down, it will move with respect to the ends of the cryostat vacuum vessel at both ends of the magnet. This movement must be considered when designing electrical leads, cryogen feed throughs, and other attachments to the coil. (2) Flexure of the coil package (at the ends of the support cylinder) will affect the spring constant of the support system. The stiffness of the ends of the coil package and the number of radial/axial supports are the determining factors for the spring constant of this type of support system.

Finite element stress and strain calculations can be used to determine the spring constant of the cold mass support system. A description of the design of a self-centering support system can be found in Ref. 69. A location of a typical self-centering support compression strut for a detector solenoid is shown in Fig. 2. The strut rotates in its sockets as the solenoid cold mass contracts. The distance between the ball sockets does not change as the solenoid cools down from room temperature to 4 K. The angle of the strut with respect to the solenoid axis changes as the coil end of the strut moves toward the center of the solenoid.

### The Solenoid Support Structure, the Cryogenic Heat Sink

The support cylinder outside the superconducting winding serves the following functions: (1) The outer cylinder carries the magnetic pressure forces that are generated by the coil. (2) The outer cylinder transfers magnetic, gravitational, and seismic forces from the coil structure to the cold mass support system. (3) The outer cylinder carries the helium cooling tubes and acts as the heat sink for the coil and all attachments to it. This means that the outer support cylinder must be made from material that conducts heat well in both the radial and axial directions.

The end ring of the support cylinder should be as stiff as possible in bending. End-ring stiffness can be increased by making the ring thicker, thus increasing its moment of inertia, or one can fabricate a laminated end ring with a high elastic modulus material such as 304 stainless steel (elastic modulus of 200 GPa as compared to 69 GPa for aluminum) on the outside and the inside of the ring with aluminum in the center. The need for stiff end rings on the support cylinder is reduced as the number of cold mass supports per end is increased for a given coil diameter (69).

### Coil Electrical Connections and Leads to the Outside World

Connections to the superconducting coil that come through the end ring should be mounted on copper bus bars that are electrically insulated from the end rings. These bus bars should be cooled in liquid helium in order to avoid heat from outside the coil being deposited directly into the superconducting windings. Heat leaks down pulsed current leads, which are usually not gas cooled, can be particularly troublesome. The cooling circuit used to cool bus bars at the ends of the coil should be part of the magnet helium cooling system. Since much of the cooling circuit is electrically grounded, in-line electrical insulators will be required in the cooling lines that cool the electrical bus bars connected to the superconducting coil.

Most detector solenoids have gas-cooled electrical leads that are fed from a liquid helium pot located somewhere near the solenoid. The current buses between the lead pot and the coil are often cooled by conduction, a practice that has led to a number of failures. All current buses should be helium-cooled. The lead pot commonly used in detector magnets can be eliminated by using gas-cooled electrical leads that are attached to the ends of the coil structure. The helium used to cool these leads comes directly from the liquid helium cooling circuit. Gas-cooled leads attached to the end of the magnet are located within the cryostat vacuum, so these leads must be completely vacuum tight and they must withstand any increase in pressure that might occur in the cooling circuit during a quench (70). The bundled nested tube leads that were used on the PEP-4 experiment (71) and the g-2 solenoids (72) can be operated at any orientation within the cryostat vacuum vessel. Properly designed

gas-cooled leads are stable and they are capable of operating for more than 30 min without gas flow.

#### CRYOGENIC COOLING OF A THIN DETECTOR SOLENOID

Most of the detector solenoids shown in Table 1 are cooled by helium in tubes attached to the superconducting coil or the support cylinder outside the coil. This technique has the following advantages over the bath cooling used for early cryostable detector magnets (73): (1) Tubular cooling eliminates the cryostat helium vessels. As a result, the solenoids are thinner and less massive. (2) The volume of helium in a tubular cooling system is small. Once this helium is evaporated during a quench, it is expelled from the tube. Large quantities of helium gas are not produced during a magnet quench. The helium expelled during a magnet quench can be returned to the refrigerator where it is recovered. (3) Tubes can withstand high pressures during a quench. Relief valves for the system can be moved from the magnet cryostat to the helium supply system, which can be outside the detector. (4) Magnet cooldown can be done directly using the helium refrigerator. (5) Recovery from a quench can be simplified using a well-designed tubular cooling system.

Detector magnets are cooled with two-phase helium rather than supercritical helium for the following reasons: (1) The operating temperature for the superconducting solenoid is lower (18). As two-phase helium flows down the cooling circuit, it gets colder as its pressure goes down. The temperature of a single-phase cooling circuit increases as one goes along the cooling circuit. (2) The mass flow through the cooling circuit is minimized. As a result, the pressure drop along the flow circuit is lower. (3) There is no need for auxiliary helium pumping in a two-phase flow circuit. Helium flow can be provided directly by

the J-T circuit of the refrigerator (4). A properly designed two-phase helium flow system can be operated at heat loads greater than the capacity of the refrigerator for a period of time. Thus fluctuations in the heat load can be tolerated by two-phase flow circuits that are designed for the average heat load. The most often stated disadvantage of two-phase cooling is the potential existence of flow and pressure fluctuations in the cooling tube. This has not been a problem in detector magnets when the two-phase helium cooling circuit is properly designed. Experiments with extensively looped cooling tubes that are hundreds of meters long have shown that proper design of the flow circuit can nearly eliminate the flow oscillation problem (73,74).

The types of two-phase helium flow circuits are commonly used in detector solenoids are the forced two-phase flow system and the natural convection two-phase flow system. Forced two-phase flow is appropriate when the flow circuits are long and when the control dewar is below the top of the magnet. Natural convection two-phase flow is appropriate when there is a large vertical head between the helium dewar and the load and when there are many parallel flow circuits so that the mass flow in any one circuit can be kept low. Either type of two-phase helium flow circuit can be made to work in most detector solenoids.

The key to stable operation of forced two-phase helium cooling circuits is the control dewar and heat exchanger (25,73). Flow for the magnet cooling circuit comes from the J-T circuit of the helium refrigerator. Cooling flow can also come from a positive displacement helium pump (75), but allowances must be made for the pump work heating generated by such a pump. Two-phase flow from the refrigerator J-T circuit flows through a heat exchanger that is cooled in a bath of liquid helium at the suction pressure of the cold end of the refrigerator. The temperature of the helium bath is the lowest temperature in the two-phase flow circuit. Within the heat exchanger, helium in

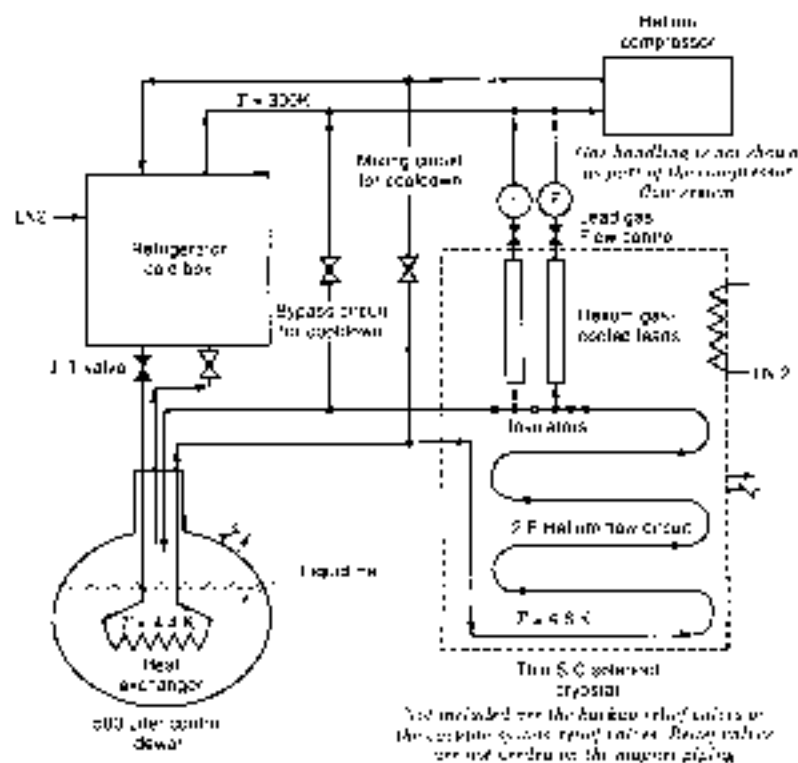


Figure 4. A schematic representation of a forced two-phase helium cooling system for a large detector solenoid while the solenoid is cold. (Note: Open valves are unshaded, closed valves are shaded. Dark lines carry flow, shaded lines carry no flow.)

the gas phase is condensed to liquid so that the helium leaving the heat exchanger either is on the saturated liquid line or is slightly subcooled. As a result, the average density of the helium in the flow circuit is maximized, which will cause the flow pressure drop through the flow circuit to be reduced a factor of two or three as compared to the same flow circuit without a heat exchanger in the helium bath. The use of the heat exchanger in the control dewar allows the operation of the cooling circuit with a heat flow into the magnet that exceeds the capacity of the refrigerator by as much as 50%. Under this condition, the magnet can be kept cold as long as the heat exchanger in the control dewar is kept covered with liquid helium. The control dewar enhances flow circuit stability. Figure 4 shows a schematic representation of a forced two-phase helium cooling circuit in its simplest form. The valves in Fig. 4 are shown as they would be when the magnet is operating at 4 K.

The advantages of forced two-phase cooling are as follows:  
 (1) The entire detector solenoid can be cooled using a single

helium flow circuit. As a result, the cooldown of the magnet is straight forward because all of the sensible heat of the helium can be employed during the cooldown process. (2) The operating temperature for a two-phase cooled magnet is lower than it would be for any supercritical helium-cooled magnet. (3) The control dewar with its heat exchange can be located flexibly with respect to the magnet coil. Transfer lines to and from the coil can be long, if desired. (4) Gas-cooled electrical leads and shields can be cooled directly from the two-phase cooling circuit. The connections for leads and shields can be made inside the magnet cryostat vacuum vessel. (5) Since the liquid helium inventory in contact with the magnet is limited to the helium in the cooling tube, the amount of helium gas produced during a quench is small. (6) The liquid helium in the control dewar can be used to speed up the recovery of the superconducting magnet after a quench. The primary disadvantage of forced two-phase cooling is that when the refrigerator stops, the magnet cooling stops. The magnet will quench within minutes after

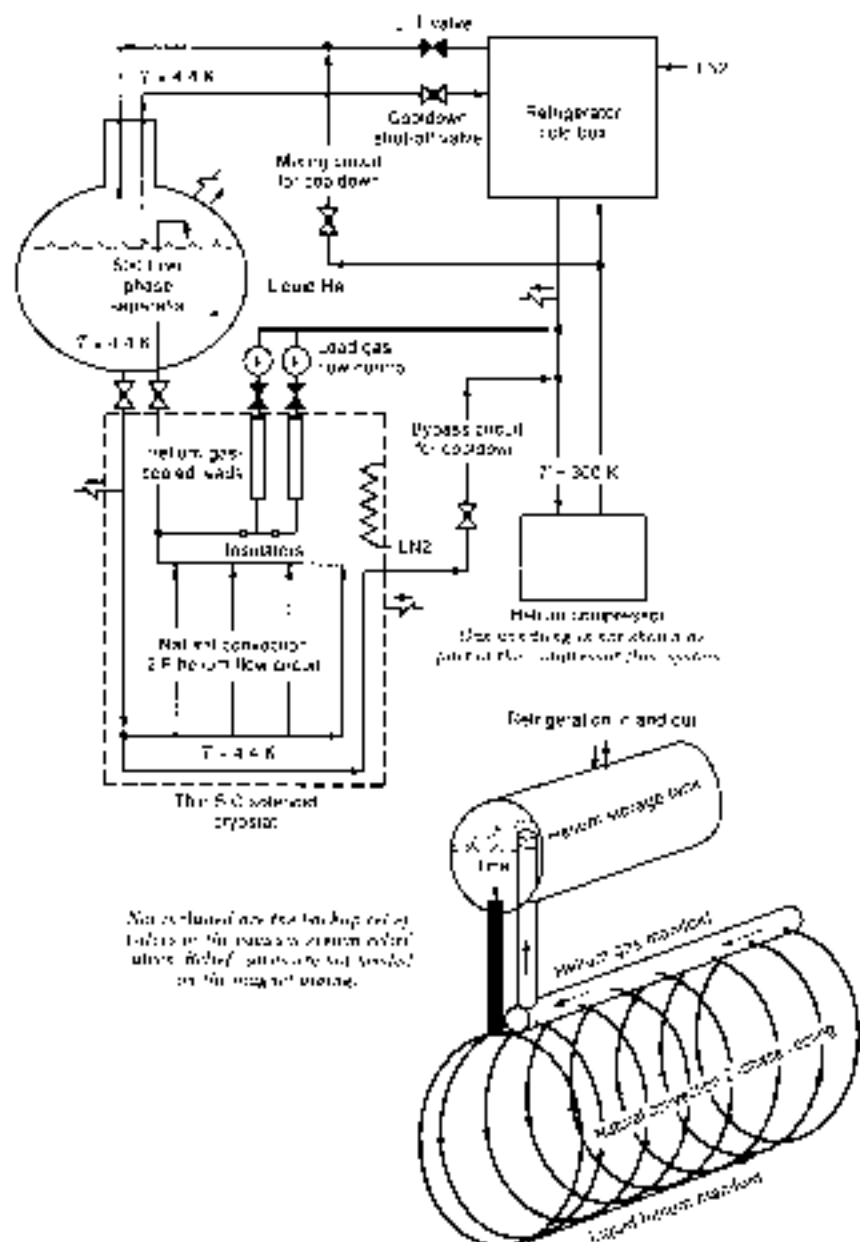


Figure 3. A schematic representation of a natural convection two-phase helium cooling system for a large detector solenoid while the solenoid is cold. (The inset at the lower right shows the physical arrangement of the natural convection cooling system.)

the refrigerator stops running. For some users this is a serious consideration.

The natural convection two-phase cooling system overcomes the primary disadvantage of a forced two-phase cooling system in that the magnet will remain cold and operating even when the refrigerator is not in operation. The cooling for a natural convection cooling system comes from the helium that is stored in a tank that is above the top of the coil. The greater the head between the storage tank and the top of the magnet, the better the natural convection two-phase flow system operates. In order for the natural convection flow system to operate effectively, the following conditions must be present: (1) The pipe from the bottom of the helium storage dewar to the manifold at the bottom of the magnet should be short and well-insulated. There should be no boiling in helium transferred to the lower manifold on the coil package. (2) In order to reduce the flow circuit pressure drop, there should be many short up-flow circuits in parallel going up and around the coil to the manifold at the top of the magnet. Boiling should occur in these tubes. This increases the helium flow through the cooling system. (3) The pipe from the manifold at the top of the coil package should dump two-phase helium into the top of the storage tank, where phase separation occurs. This pipe should be insulated from the helium that is in the storage tank. The difference in helium density in the pipe connecting the tank and the lower manifold and the two-phase helium in the cooling tubes circling the magnet provides the driving force for the flow circuit. Figure 5 shows a schematic representation of a natural convection two-phase helium cooling circuit in its simplest form. The upper part of Fig. 5 shows a schematic of how the refrigerator, its compressors, and the magnet would be hooked up. The lower part of Fig. 5 shows the physical arrangement of the helium storage tank (phase separator) and the detector solenoid. The storage tank shown in Fig. 5 is 500 liters, but that tank could be much larger if needed. The valves in Fig. 5 are shown as they would be when the magnet is operating at 4 K.

Natural convection flow has some disadvantages, which are as follows: (1) Cooledown of the magnet is not as straightforward as with the forced two-phase cooling system. This difficulty can be overcome by having a separate forced flow circuit for the magnet cooledown. (2) In a natural convection cooled magnet system, all of the helium that is in the storage tank may be boiled during a magnet quench. This can be overcome by installing an automatic shut-off valve, which is triggered by the quench detector, in the pipe between the storage tank and the liquid helium manifold on the bottom of the magnet. (3) The thin section of the solenoid has a larger radiation thickness top and bottom in the regions where the liquid helium and two-phase helium manifolds are located. (4) The helium storage tank for natural convection cooling must be located directly above the magnet coil, and the transfer lines between the storage tank and the coil should be as short as possible. The physics experiment must accommodate these transfer lines.

## BIBLIOGRAPHY

- J. E. Kunzler et al., *Phys. Rev. Lett.*, **4**, 49, 1961
- R. C. Wulgant et al., Superconducting critical currents in wire samples and some experimental coils, in *Advances in Cryogenic Engineering*, Vol. 9, New York: Plenum Press, 1962, p. 36
- C. Laverick, The performance characteristics of small superconducting coils, in *Advances in Cryogenic Engineering*, Vol. 10, New York: Plenum Press, 1964, p. 105.
- Z. J. J. Stekly and Z. L. Zar, Stable superconducting coils, *IEEE Trans. Nucl. Sci.*, **NS-12**, 367, 1965.
- J. R. Powell, The 1.6 tesla, 4.9 m bubble chamber magnet, *Proc. 1968 Summer Study Superconducting Device Accelerators*, BNL 50155 (C-55), p. 785, 1968.
- R. E. Jones et al., Construction of the 12 foot bubble chamber superconducting magnet cryostat, in *Advances in Cryogenic Engineering*, Vol. 15, New York: Plenum Press, 1969, p. 141.
- D. P. Brown, R. W. Burgess, and G. T. Mulholland, The superconducting magnet for the Brookhaven National Laboratory 7 foot bubble chamber, in *Proc. 1968 Summer Study Superconducting Device Accelerators*, BNL 50155 (C-55), p. 794, 1968.
- E. G. Pewatt, Superconductivity in high energy physics, in *Advances in Cryogenic Engineering*, Vol. 16, New York: Plenum Press, 1970, p. 19 (a review of high temperature).
- E. V. Huelbel and F. Wittenstein, *Proc. Int. Conf. Magnet Technol.*, Hamburg, Germany, 1970.
- T. H. Fields, Superconductivity applications in high energy physics, *IEEE Trans. Magn.*, **MAG-11**, 113, 1975.
- M. Marzengo, Construction of a superconducting test coil cooled by helium forced circulation, *Proc. 1968 Summer Study Superconducting Device Accelerators*, BNL 50155 (C-55), 953, 1968.
- M. Marzengo, *Proc. Int. Conf. Magnet Technol.*, Hamburg, Germany, 1970.
- C. P. Bean et al., Research Investigation of the Factors that Affect the Superconducting Properties of Material, AFML-TR-65-421, Air Force Materials Laboratory, Wright Patterson Air Force Base, Ohio, 1965.
- R. Hancock, *Phys. Lett.*, **18**, 209, 1965.
- P. F. Chester, *Rep. Prof. Phys.*, **30**(2), 361, 1967.
- P. F. Smith and J. D. Lewin, Pulsed superconducting synchrotrons, *Nucl. Instr. Methods*, **52**, 246, 1967.
- M. N. Wilson et al., Experimental and theoretical studies of filamentary superconducting composites. Part 1: Basic ideas and theory, *J. Phys. Appl. Phys.*, **3**, 1517, 1970.
- M. A. Green, Cooling intrinsically stable superconducting magnets with super-critical helium, *IEEE Trans. Nucl. Sci.*, **NS-18**, 664, 1971.
- H. Desportes, Superconducting magnets for accelerators, beam lines and detectors, *IEEE Trans. Magn.*, **MAG-17**, 1591, 1981.
- M. A. Green, Large superconducting solenoid for the MINIMAX experiment, in *Advances in Cryogenic Engineering*, Vol. 21, New York: Plenum Press, 1975, p. 24.
- P. H. Eberhard et al., Tests on large diameter superconducting solenoids designed for colliding beam accelerators, *IEEE Trans. Magn.*, **MAG-13**, 74, 1977.
- M. A. Green, The Development of Large High Current Density Superconducting Solenoid Magnets for Use in High Energy Physics Experiments, LBL-5050 doctoral dissertation, UC Berkeley, May 1977.
- M. A. Green, Large superconducting detector magnets with ultra thin coils for use in high energy accelerators and storage rings, *Proc. 6th Int. Conf. Magnet Technol.*, Bratislava, Czechoslovakia, p. 429, 1977.
- P. H. Eberhard et al., A magnet system for the time projection chamber at PEP, *IEEE Trans. Magn.*, **MAG-15**, 126, 1979.
- M. A. Green, W. A. Burns, and J. D. Taylor, Forced two-phase helium cooling of large superconducting magnets, in *Advances in Cryogenic Engineering*, Vol. 25, New York: Plenum Press, 1979, p. 420.
- J. Benches et al., Long term experience on the superconducting

- magnet system for the CLEA detector. *IEEE Trans. Magn.*, **MAG-17**: 1567, 1991.
27. J. M. Royer, J. B. Soudiere, and R. E. Schwall, Aluminum stabilized multifilamentary Nb-Ti conductor. *IEEE Trans. Magn.*, **MAG-19**: 761, 1993.
  28. E. Levng et al., Design of an indirectly cooled 3-m diameter superconducting solenoid with external support cylinder for the Fermilab collider detector facility. *IEEE Trans. Magn.*, **MAG-19**: 1386, 1993.
  29. A. Winkl et al., A large superconducting thin solenoid magnet TRISTAN experiment (VENUS) at KEK. *IEEE Trans. Magn.*, **MAG-21**: 494, 1995.
  30. H. Hirabayashi, Detector magnets in high energy physics. *IEEE Trans. Magn.*, **MAG-24**: 1256, 1998.
  31. K. Tsuchiya et al., Testing of a 3 tesla superconducting magnet for the AMY detector at Tristun. in *Advances in Cryogenic Engineering*, Vol. 33. New York: Plenum Press, 1987, p. 33.
  32. J. M. Beze et al., Design, construction and test of the large superconducting solenoid ALEPH. *IEEE Trans. Magn.*, **MAG-24**: 1260, 1988.
  33. R. Q. Apsley et al., Design of a 3.5 meter diameter superconducting solenoid for the DELPHI particle physics experiment at LEP. *IEEE Trans. Magn.*, **MAG-21**: 490, 1995.
  34. P. F. M. Clee, Rutherford Appleton Laboratory, Didcot, United Kingdom, private communication.
  35. A. Bounia Ollas et al., Zero thin solenoid Test results analysis. *IEEE Trans. Magn.*, **MAG-27**: 1954, 1991.
  36. D. Andrews, Oxford Technology Ltd., Oxford, United Kingdom, private communication.
  37. C. M. Monroe et al., The CLEO II Magnet—design, manufacture and tests. *Proc. Tenth Int. Cryogenic Eng. Conf.*, Southampton, United Kingdom, Guildford, United Kingdom: Butterworth, 773, 1989.
  38. A. Yamamoto et al., Development of a prototype thin superconducting solenoid magnet for the SDC detector. *IEEE Trans. Appl. Supercond.*, **5**: 649, 1995.
  39. T. Miao et al., Prototype thin superconducting solenoid for particle astrophysics in space. *IEEE Trans. Magn.*, **MAG-25**: 1665, 1989.
  40. Y. Makida et al., Ballasting of a thin superconducting solenoid for particle astrophysics. *IEEE Trans. Appl. Supercond.*, **6**: 656, 1995.
  41. P. Fabrice et al., The superconducting magnet for the BaBar detector of the PEP-II B factory at SLAC. *IEEE Trans. Magn.*, **MAG-32**: 2210, 1996.
  42. J. M. Beze, Progress in the design of a superconducting toroidal magnet for the ATLAS detector on LHC. *IEEE Trans. Magn.*, **MAG-32**: 2047, 1996.
  43. D. E. Baylham et al., Design of the superconducting end cap toroids for the ATLAS experiment at LHC. *IEEE Trans. Magn.*, **MAG-32**: 2055, 1996.
  44. F. Kircher, H. Daponte, and R. Gallet et al., Conductor developments for the ATLAS and CMS magnets. *IEEE Trans. Magn.*, **MAG-32**: 2870, 1996.
  45. G. Bunce et al., The large superconducting solenoids for the g-2 muon storage ring. *IEEE Trans. Appl. Supercond.*, **5**: 855, 1995.
  46. G. Bunce et al., Test results of the g-2 superconducting solenoid magnet system. *IEEE Trans. Appl. Supercond.*, **7**: 826, 1997.
  47. *Review of Particle Properties*, compiled by the Particle Data Group. *Rev. Mod. Phys.*, **49**(2) Part II: 50, 1978.
  48. Y. S. Tso, Pair Production and Bremsstrahlung of Charged Leptons. SLAC PUB-1365, Table III b, 1974.
  49. W. R. Smythe, *Static and Dynamic Electricity*, 2nd ed., New York: McGraw-Hill, 1950.
  50. POISSON/SUPERFISH Reference Manual, LANL Publication LA-UR-97-126, 1997.
  51. Part of the TWSCA series of programs produced by Vector Fields Ltd., Oxford, United Kingdom.
  52. M. A. Green, Calculating the J, B, T surface for muonium atomium using the reduced state model. *IEEE Trans. Magn.*, **MAG-25**: 2119, 1999.
  53. B. J. Maddock and C. B. James, Protection and stabilization of large superconducting coils. *Proc. IEE*, **115**(4): 543, 1966.
  54. P. H. Eberhard et al., Quenches in large superconducting magnets. In *Proc. 6th Int. Conf. Magnet Technol.*, Bratislava, Czechoslovakia, 854, 1977.
  55. M. A. Green, Quench back in thin superconducting solenoid magnets. *Cryogenics*, **24**: 3, 1984.
  56. A. Yamamoto et al., A thin superconducting solenoid with the internal winding method for collider beam experiments. *Proc. 8th Int. Conf. Magnet Technol.*, September 1983.
  57. A. Yamamoto et al., Design study of a thin superconducting solenoid for the SDC detector. *IEEE Trans. Appl. Supercond.*, **3**: 95, 1993.
  58. R. J. Bonk and W. C. Young, *Formulas for Stress and Strain*, 5th ed., New York: McGraw-Hill, 1975.
  59. S. Timoshenko, *Theory of Plates and Shells*. New York: McGraw-Hill, 1940.
  60. H. E. Sanders and D. E. Wittenberg, Strength of thin cylindrical shells under external pressure. *ASME Trans.*, **58**: 297, 1931.
  61. E. P. Popo, *Mechanics of Materials*. Englewood Cliffs, NJ: Prentice Hall, 1959.
  62. ASME Boiler and Pressure Vessel Code, Section 8, Division 1, ANSI/ASME BPP-VIII-1.
  63. R. Post et al., Ingrid vacuum shell for large superconducting solenoids. in *Advances in Cryogenic Engineering*, Vol. 39, New York: Plenum Press, 1993, p. 1991.
  64. H. Yamasa et al., Development of a brazed aluminum-honeycomb vacuum vessel for a thin superconducting solenoid magnet. in *Advances in Cryogenic Engineering*, Vol. 39, New York: Plenum Press, 1993, p. 1993.
  65. M. N. Wilson, *Superconducting Magnets*. Oxford, United Kingdom: Oxford Clarendon Press, 1983, p. 219.
  66. M. A. Green, The role of quench back in the quench protection of a superconducting solenoid. *Cryogenics*, **24**: 659, 1984.
  67. M. A. Green, PEP-4 TPC Superconducting Magnets, A Comparison of Measured Quench Back Time with Theoretical Calculations of Quench Back Time for Four Thin Superconducting Magnets, Lawrence Berkeley Laboratory Report LBL-771, August 1983, unpublished.
  68. J. D. Taylor et al., Quench protection for a 2 MJ magnet. *IEEE Trans. Magn.*, **MAG-16**: 835, 1979.
  69. M. A. Green, PEP-4, Large Thin Superconducting Solenoid Magnet, Cryogenic Support System Revisited, Lawrence Berkeley Laboratory Engineering Note M5855, March 1982, unpublished.
  70. M. A. Green, Calculation of the pressure rise in the cooling tube of a two phase cooling system during a quench of an indirectly cooled superconducting magnet. *IEEE Trans. Magn.*, **MAG-30**: 2427, 1994.
  71. B. G. Smith et al., Gas-cooled electrical leads for use on forced-cooled superconducting magnets. *Advances Cryogenic Eng.*, Vol. 27, New York: Plenum Press, 1981, p. 169.
  72. M. A. Green et al., A design method for multi-tube gas-cooled electrical leads for the g-2 superconducting magnets. in *Advances in Cryogenic Engineering*, Vol. 41, New York: Plenum Press, 1996, p. 573.
  73. M. A. Green et al., The TPC Magnet Cryogenic System, Lawrence Berkeley Laboratory Report LBL 10552, May 1980, unpublished.
  74. J. D. Taylor and M. A. Green, Garden Hose Test, Lawrence Berkeley Laboratory, Group A Physics Note-877, November 1978, unpublished.

75. W. A. Burns et al., The construction and testing of a double-acting bellows liquid helium pump, *Proc. 8th Int. Cryogenics Eng. Conf.*, Genoa, Italy, Guildford, United Kingdom, IPC Science and Technology Press, 303, 1980.

MICHAEL A. GREEN  
Lawrence Berkeley National  
Laboratory

## INFRARED DETECTOR ARRAYS, UNCOOLED

Infrared imaging has demonstrated itself to be a vital aspect of modern warfare. Infrared (IR) imaging has been used for surveillance, targeting, and night vision. The civilian applications of infrared imaging for security, thermography, and night vision in transportation are becoming increasingly widespread. The key factor limiting the adoption of widespread civilian applications of infrared imaging is the high cost associated with cryogenically cooled IR detector arrays. More recently, thermal imaging arrays capable of operating at or near room temperature without costly cryogenic refrigeration have been developed. These systems are less expensive than their cryogenic counterparts; however, they are not yet inexpensive enough to be feasible for mass consumer applications. However, as uncooled infrared imaging technology develops, low-cost systems will be developed, enabling night vision in automobiles, significantly improving automobile safety by allowing the driver to see beyond the range of the headlights and enabling other mass consumer applications in transportation, security, and medicine to be realized.

Infrared detectors can be generally classified as belonging to one of two types. There are photon detectors and thermal detectors. Photon detectors generally operate using the photovoltaic effect or photoconductivity. In either case, a photon detector relies upon the absorption of a quantum of light by an electron. The electron then may be excited to a state where it can be transported over a barrier; or, more commonly, it results in the generation of electron-hole pairs, allowing for increased conductivity (photoconductivity) or a shift in the quasi-Fermi levels (photovoltaic). Photon detectors do not possess a high detectivity at room temperature due to the noise associated with the dark current. To minimize the dark current, photon detectors are generally cooled to cryogenic temperatures by submersion in liquid helium or liquid nitrogen or by employing a closed cycle refrigerator. The necessity of cryogenic cooling significantly increases the system cost and the complexity of an infrared camera. Most infrared photon detectors are based upon HgCdTe devices or silicon Schottky barrier diodes.

In contrast, thermal detectors operate by utilizing the heat generated by the absorption of the photon flux to change temperature. The change in the temperature is associated with a change in a measurable electrical property of the material such as the electrical resistance (bolometer), spontaneous polarization (pyroelectric detector), or Fermi level (thermocouple). To maximize the response of a thermal detector, it is important to maximize the temperature change with respect to the energy absorbed by the photon flux. To this end, it is important that a minimal amount of heat be lost through conduction away from the detector. The detector is, therefore, usually thermally isolated from its surrounding as much as possible to achieve a

large responsivity. This can be achieved in microdetectors by microstructuring the detectors to be suspended above the substrate and placing the detector in vacuum, thereby minimizing the amount of heat lost to the substrate by conduction. Typically, a thermal conductance of  $10^{-7}$  W/K is achieved. In addition, microfabrication allows detectors with a very low thermal mass (specific heat) ( $10^{-16}$  J/K) to be fabricated, again providing large temperature changes in the detecting element with respect to the incident photon energy. Thermal detectors do not suffer from noise associated with dark current, therefore, high-detectivity detectors ( $D \approx 10^{11}$  cm-Hz $^{1/2}$ /W) operating at room temperature are possible.

## BOLOMETRIC DETECTORS

### Theory of Bolometer Operation

A bolometer operates through the temperature-dependence of the resistance of the sensitive element or thermometer. As the temperature of the bolometer changes with the energy carried by the incident photon flux, the resistance changes, thereby giving a measurable signal when current biased. The relative magnitude of the change in the electrical resistance is known as the temperature coefficient of resistance (TCR). The TCR is given by

$$\text{TCR} = \rho = \frac{1}{R} \frac{dR}{dT} \quad (1)$$

Performance improves with the magnitude of the TCR. Metals (1,2) were originally used as bolometers; however, their TCR is limited to  $0.5\% \text{ K}^{-1}$ , thereby limiting their responsivity and detectivity. At present, practical microbolometers are based upon semiconductors such as germanium (3), amorphous silicon (4), and vanadium oxide (5-9). In addition, more recently, semiconducting YBaCuO has been proposed (10-13).

The responsivity of a bolometer—that is, the output signal voltage per unit incident infrared power—is given by (14)

$$R_V = \frac{I_b R \rho \eta}{G(1 + \omega^2 \tau^2)^{1/2}} \quad (2)$$

where  $I_b$  is the bias current,  $R$  is the direct current (dc) resistance,  $\eta$  is the absorptivity,  $G$  is the thermal conductance between sensitive element and the substrate,  $\omega$  is the angular modulation frequency of the incident radiation, and  $\tau$  is the thermal response time which is given by  $C/G$ .  $C$  is the heat capacity (thermal mass) of the sensitive element. Therefore, for high responsivity, high  $dR/dT$ , low  $G$ , and low  $\omega$  (or  $\omega \ll 1/\tau$ ) are required. Silicon microstructuring techniques can be employed to create an air-gap bridge under the detector to provide low thermal conductance and heat capacity (15-17).

The detectivity  $D$  is determined by the ratio of the responsivity  $R_V$  to the noise voltage  $V_n$ :

$$D = \frac{R_V \sqrt{\Delta f A}}{V_n} \quad (3)$$

where  $\Delta f$  is the amplifier frequency bandwidth,  $V_n$  is the total noise voltage of the detector, and  $A$  is the area of the detector (thermometer).

The noise voltage,  $V_n$ , is determined by the sum of the contributions due to the background noise produced by the blackbody emissions of the surroundings, the temperature fluctuation noise due to thermodynamic fluctuations in the isolated thermal mass, and the noise generated by the thermometer or sensitive element. The noise generated by the thermometer includes Johnson noise and the low-frequency noise of the material. The respective noise contributions may be calculated from the following relations.

The temperature fluctuation noise is due to temperature fluctuations resulting from the exchange thermal conduction of isolated thermal detector with the substrate. The temperature fluctuation noise voltage,  $\Delta V_{TF}$ , is calculated from:

$$\frac{\Delta V_{TF}}{\sqrt{\Delta f}} = \frac{2k_B^2 I_b R T_D}{(1 + 1/4\pi^2 f^2 \tau^2)^{1/2}} \quad (14)$$

where  $T_D$  is the detector temperature and  $k_B$  is Boltzmann's constant. The background voltage noise,  $\Delta V_{Bg}$ , results from the radiative exchange of the detector with the surroundings. For a detector surrounded by a uniform blackbody at temperature  $T_n$ , the background voltage noise may be calculated from

$$\frac{\Delta V_{Bg}}{\sqrt{\Delta f}} = \frac{I_b R \epsilon (8A\sigma k_B (T_n^4 + T_D^4))^{1/2}}{(1 + 4\pi^2 f^2 \tau^2)^{1/2}} \quad (15)$$

where  $A$  is the total surface area of the detector and  $\sigma$  is the Stefan-Boltzmann constant.

The voltage noise generated by the sensitive material or thermometer occurs due to the Johnson noise associated with the resistance of the detector plus the  $1/f$  noise of the material. The Johnson noise voltage is given by

$$\frac{\Delta V_J}{\sqrt{\Delta f}} = \sqrt{4k_B T_D R} \quad (16)$$

The  $1/f$  noise may originate due to a variety of physical processes. The magnitude of the  $1/f$  noise must be determined empirically for a given material used as the thermometer. The  $1/f$  noise is often determined by material processing parameters because it often relates to the number of traps throughout the volume or at the surface of the thermometer layer. In general, the Hooge formula (17) provides a relation for the voltage spectral density associated with the  $1/f$  noise. The noise voltage is taken as the square root of the voltage spectral density. The Hooge formula is given by

$$S_V = \frac{\alpha V^2}{fN} = \frac{\Delta V_{1/f}^2}{\Delta f} \quad (17)$$

Here,  $\alpha$  is Hooge parameter which provides the relative magnitude of the  $1/f$  noise,  $V$  is the dc voltage, and  $N$  is the total number of independent fluctuators. The fluctuators are the scattering centers creating the noise. In bulk materials, the fluctuators are often distributed evenly throughout the volume of the material; however, in the case of thin films, as employed in microbolometers, the dominant scattering mechanism is often associated with surface states (18) and the fluctuators are distributed over the surface of the thin film.

The corner frequency occurs where the  $1/f$  noise merges with the Johnson noise floor. The power-normalized corner frequency reflects the inherent noise in the bolometer and is a

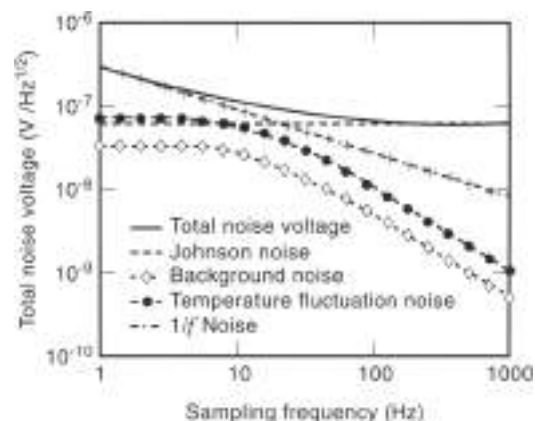
useful comparison for the noise in different types of bolometers. By equating the Hooge formula to the Johnson noise, the power-normalized corner frequency is given by

$$f_c = \frac{\alpha}{I_b^2 R} = \frac{1}{4k_B T_D N} = \frac{1}{4k_B T_D} \left( \frac{\Delta V_{1/f}}{V} \right)^2 \frac{f}{\Delta f} \quad (18)$$

The total noise voltage  $V_n$  is provided by the sum of squares of the contributions due to temperature fluctuation noise, background noise, Johnson noise, and  $1/f$  noise.

$$\frac{V_n^2}{\Delta f} = \frac{\Delta V_{TF}^2}{\Delta f} + \frac{\Delta V_{Bg}^2}{\Delta f} + \frac{\Delta V_J^2}{\Delta f} + \frac{\Delta V_{1/f}^2}{\Delta f} \quad (19)$$

The relative magnitude of the different noise components is plotted in Fig. 1. As can be seen by the plot, the  $1/f$ -noise component dominates the noise voltage at low frequencies, while at high frequencies the Johnson noise dominates. A material with lower  $1/f$  noise would see the contributions due to temperature fluctuation noise and background noise play a greater role at intermediate frequencies. It is important to note that the magnitude of the  $1/f$  noise, temperature fluctuation noise, background noise, and responsivity are proportionate to the bias current  $I_b$ . At zero bias, the noise voltage is determined by the Johnson noise. However, there is no responsivity. As  $I_b$  is increased, the responsivity is increased as well as the contributions due to background, temperature fluctuation, and  $1/f$  noise. While the magnitude of these noise components remain much less than the Johnson noise, the detectivity increases at the sampling frequency of interest, until the magnitude of these noise components starts to dominate over the Johnson noise. At this point, both the responsivity and the device noise are increasing with the current bias, thereby leading to a saturated, maximum detectivity. In imaging arrays, image quality dictates operating the bolometers in the Johnson noise regime where the contributions from the temperature fluctuation, background, and  $1/f$  noise are small, resulting in an operating detectivity less than the maximum. The background and temperature fluctuation noise components have the same cutoff frequency due to



**Figure 1.** A plot of the noise components and total noise for a microbolometer as a function of sampling frequency. The following parameters were used in constructing the plot:  $\beta = 0.5\% \text{ K}^{-1}$ ,  $R = 230 \text{ k}\Omega$ ,  $\tau = 1 \text{ f}$ ,  $C = 0.7 \text{ pF}$ ,  $K = 70 \text{ nW/K}$ ,  $\epsilon = 0.01 \text{ s}$ ,  $T_D = 300 \text{ K}$ ,  $A = 60 \mu\text{m} \times 50 \mu\text{m}$ ,  $\sigma = 10^{-12}$ , and  $I_b = 1 \mu\text{A}$ .

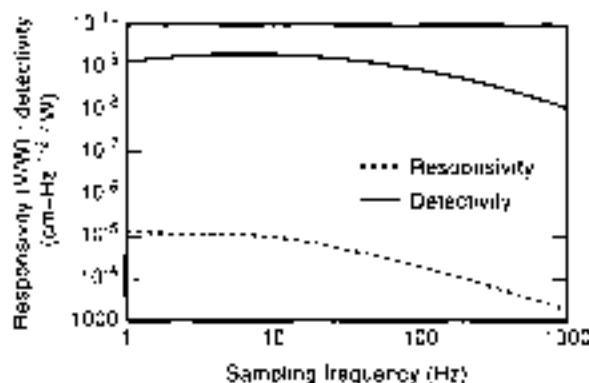


Figure 2. A plot of typical microbolometer responsivity and detectivity of sampling frequency using the same parameters described in Fig. 1.

the thermal time constant as the responsivity, while the Johnson noise continues to high frequencies. To decrease the total Johnson noise voltage contribution, a low-pass filter is used.

The corresponding responsivity and detectivity are plotted versus sampling frequency in Fig. 2. The plot shows that the responsivity is constant at low sampling frequencies and decreases at higher frequencies due to the thermal time constant of the thermal isolation structure. The detectivity is reduced slightly at low frequencies due to the presence of the  $1/f$  noise, and it reaches a maximum as the  $1/f$  noise merges with the Johnson noise before decreasing with the thermal time constant of the bridge structure. Detectivities in the range of  $10^4$   $\text{cm}\cdot\text{Hz}^{1/2}/\text{W}$  to  $10^{12}$   $\text{cm}\cdot\text{Hz}^{1/2}/\text{W}$  are achievable with state-of-the-art uncooled microbolometers.

The signal-to-noise ratio of the bolometer is often expressed in terms of a noise equivalent power (NEP), which is the incident photon power required to produce a voltage signal equal to the total noise voltage. The NEP can be calculated from the ratio of the total voltage noise divided by the responsivity. The NEP of uncooled microbolometers is typically higher than cooled IR detectors simply because of the operation at elevated temperatures:

$$\text{NEP} = \frac{V_n}{R_c} \quad (10)$$

When the microbolometers are integrated into arrays, the ability to have high detectivity is important, but the most important figure of merit is the ability to resolve small temperature differences in the field of view. This figure of merit is expressed as the noise equivalent temperature difference (NETD). The NETD is given by

$$\text{NETD} = \frac{4}{\tau} \left[ \frac{(F/\#)^2 \sqrt{\Delta f}}{D + (dL/dT)_{T_0} \sqrt{A_d}} \right] \quad (11)$$

where,  $F/\#$  is the  $F$  number of the optics,  $dL/dT$  is the rate of change of the radiance of the object with temperature, and  $A_d$  is the area of the detector. Most thermal imaging systems operate either in the  $3 \mu\text{m}$  to  $5 \mu\text{m}$  atmospheric window or the  $8 \mu\text{m}$  to  $14 \mu\text{m}$  atmospheric window, where  $(dL/dT)_{T_0} \sqrt{A_d}$  can be shown to be equal to  $2 \times 10^{-5}$   $\text{W}/\text{cm}^2\cdot\text{K}$  and  $2.6 \times 10^{-6}$   $\text{W}/\text{cm}^2\cdot\text{K}$ , respectively.

Evaluating Eq. (11) in the case of temperature fluctuation noise limited detectivity, the NETD is given by

$$\text{NETD}_{\text{TF}} = \frac{8}{\tau} \left[ \frac{T_D (F/\#)^2 \sqrt{\Delta f}}{d(L/dT)_{T_0} \sqrt{A_d}} \right]$$

The  $\text{NETD}_{\text{TF}}$  therefore decreases as the square root of the thermal conductance away from the bolometer and is directly proportional to the detector temperature. As the thermal conductance of the bolometer is decreased, the temperature fluctuation noise decreases below the background noise. In this case, the NETD would be limited by the background noise associated with the radiative energy exchange between the bolometer and its surroundings. The NETD in this case is given by

$$\text{NETD}_{\text{BG}} = \frac{8}{\pi} \frac{(F/\#)^2}{(dL/dT)_{T_0} \sqrt{A_d}} \sqrt{\frac{2\tau k_B (T_D^4 - T_B^4) \Delta f}{\eta A_d}}$$

The  $\text{NETD}_{\text{BG}}$  is independent of the thermal isolation but depends upon the temperature of the detector and background. At room temperature, the  $\text{NETD}_{\text{BG}}$  is calculated to be 5 mK in the  $3 \mu\text{m}$  to  $5 \mu\text{m}$  band and 0.4 mK in the  $8 \mu\text{m}$  to  $14 \mu\text{m}$  band. In performing these calculations, it is assumed that the system would have  $F/1$  optics,  $\eta = 1$  absorptivity,  $40 \mu\text{m} \times 40 \mu\text{m}$  pixel size, bandwidth  $\Delta f = 30$  Hz, device temperature  $T_D = 300$  K, and background temperature  $T_B = 300$  K. NETDs less than 30 mK have been measured with uncooled microbolometer imaging arrays. This result is approaching the background limited noise performance. It is an impressive performance considering the approximations made in the calculations presented here. Reference 30 provides a more detailed discussion on the calculation of NETD and the dependence upon thermal isolation and detector temperature.

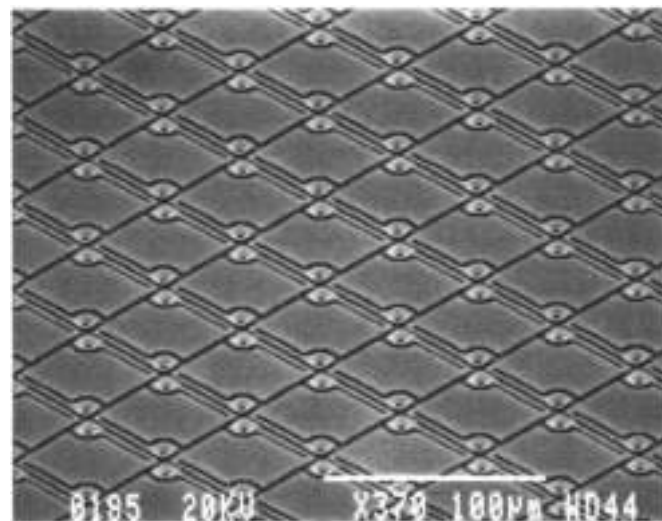
#### Vanadium Oxide Microbolometer Arrays

Vanadium oxide ( $\text{VO}_x$ ) microbolometer technology was developed by Honeywell (9) and is currently being used under license by Raytheon-Amber (8), Hughes Aerospace (5), Rockwell International (7), and Lockheed-Martin (6). The technology employed by these companies varies mainly in terms of the readout circuitry employed to operate the microbolometer arrays, the method of fabricating the thermal isolation structures, and the geometry of the detector arrays fabricated. A description of the  $\text{VO}_x$  focal plane array (FPA) technology at Hughes Santa Barbara Research Center (SBRC) follows.

SBRC has been developing uncooled microbolometer FPAs since Hughes Aircraft Company licensed the microbolometer detector technology from Honeywell (5) in 1993. SBRC has focused on developing high-performance readout integrated circuits (ROICs) and improving the microbolometer detector structure and fabrication process. These improvements have included increases in pixel thermal isolation, optical fill-factor, and temperature coefficient of resistance. Simultaneously, the fabrication process has been simplified to achieve better yields and faster cycle times.

A scanning electron micrograph of the microbolometer pixels fabricated with the dry sacrificial etch process and the optimized  $\text{Si}_3\text{N}_4$  is shown in Fig. 3. A self-aligned leg etch process is being used to reduce the microbolometer support leg width to less than  $2 \mu\text{m}$ . This has increased the thermal isolation by





**Figure 3.** Scanning electron micrograph of microbolometer pixels fabricated at SBRC using polyamide sacrificial etch process. From Hughes Santa Barbara Research Center with permission. © 1997 Santa Barbara Research Center.

a factor of two while allowing area for increased detector fill-factor. A polyamide sacrificial layer has been implemented in place of the baseline  $\text{SiO}_2$  sacrificial layer. The polyamide material allows the use of a dry sacrificial etch rather than the wet HF-based etch used in the baseline process. The elimination of the wet etch has dramatically simplified the microbolometer fabrication process. The dry etch process has an essentially zero etch rate of (a) the  $\text{Si}_3\text{N}_4$  used for the bridge and (b) the oxides and metal layers used in the readout substrates.

The dry sacrificial etch process also facilitated performance improvements in the microbolometer responsivity. The thickness of the  $\text{Si}_3\text{N}_4$  layers has been reduced without sacrificing device yield.  $\text{Si}_3\text{N}_4$  bridge thicknesses as thin as 360 nm have been successfully demonstrated using the dry etch process. The deposition process of the  $\text{Si}_3\text{N}_4$  process has been optimized for throughput and thermal properties. The nitride deposition process can complete a 20-wafer lot in about 1 h. The deposition process has also been optimized to produce films with low stress and low thermal conductivity. The thermal conductivity of the  $\text{Si}_3\text{N}_4$  is only about  $0.8 \text{ W m}^{-1} \text{ K}^{-1}$ . This represents over a factor-of-two reduction of thermal conductivity compared with the baseline material.

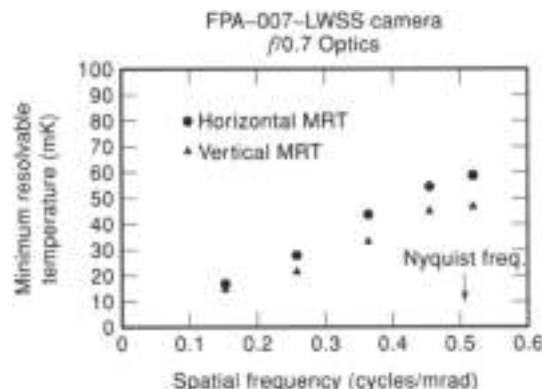
An advanced microbolometer fabrication process has been employed to build high-image quality IR FPAs. These FPAs have a  $320 \times 240$  format with  $50 \mu\text{m} \times 50 \mu\text{m}$  pixels and are based on a complementary metal oxide semiconductor (CMOS) read-out integrated circuit (ROIC). The ROIC uses on-chip clock and bias generation to provide a simple electrical interface requiring only three clocks and five bias levels. The FPA has a single video output and can operate at frame rates up to 60 Hz. The SBRC-151 ROIC operates in an electronically scanned format and pulse biases each pixel. It performs two levels of on-chip offset correction to minimize spatial nonuniformity of the FPA pixels. The on-chip correction allows the use of a high detector bias (4 V) and on-chip gain without saturation of the output range. The readout utilizes a differential architecture throughout the signal chain in order to minimize sensitivity

to bias fluctuation and external noise sources. These features give the SBRC-151 very high responsivity and good extraneous noise immunity. The ROIC also incorporates an on-chip temperature compensation capability in order to minimize temperature stabilization requirements. The various operational and test modes of the chip are controlled through a 32-bit serial programmable interface. The ROIC was specifically designed to accommodate a wide range of detector impedances without any degradation of sensitivity. The ROIC will operate with pixel impedances ranging from 10 k $\Omega$  to 200 k $\Omega$  with little degradation in performance.

The microbolometer FPAs are mounted in vacuum packages with a single-stage thermoelectric cooler for temperature stabilization. An antireflective-coated germanium (Ge) window is used to achieve high transmission in the  $8 \mu\text{m}$  to  $14 \mu\text{m}$  spectral region.

The SBRC-151 FPAs have been integrated into several camera systems including the long-wavelength staring sensor (LWSS) developed by Hughes Sensors and Communications Systems (SCS). The LWSS is a prototype military IR camera for portable applications. The camera was designed to achieve low-power consumption in order to maximize battery life. The camera utilizes a 50 mm focal length  $f/0.7$  Ge lens assembly with a broadband AR coating. The LWSS camera provides memory for the coarse on-chip correction terms as well as the gain and offset nonuniformity correction terms.

The LWSS sensor has been used for imaging demonstrations and radiometric testing of the SBRC-151 FPAs. The Hughes LWSS camera was independently evaluated at the US Army Night Vision and Electronic Sensors Directorate (NVESD). The NVESD performed measurements of NETD, minimum resolvable temperature (MRT), and three-dimensional noise. An NETD of 24 mK was measured with the  $f/0.7$  optics while an NETD of 42 mK was measured with the optics stopped down to  $f/1.0$ . Figure 4 presents MRT measurements obtained by NVESD using the  $f/0.7$  optics. The vertical MRT ranged from 48 mK at the Nyquist frequency ( $f_N$ ) to 16 mK at low spatial frequency (0.30  $f_N$ ). The horizontal MRT ranged from 59 mK at the  $f_N$  to 17 mK at 0.30  $f_N$ . The NVESD measured spatial camera noise was less than 50% of the temporal noise. Further optimization of the FPA and camera is expected to produce NETD



**Figure 4.** Minimum resolvable temperature data taken at US Army NVESD on Hughes LWSS (a portable uncooled sensor NVESD) measured an NETD value of 24 mK on the sensor. From Hughes Santa Barbara Research Center with permission. © 1997, Santa Barbara Research Center.



Figure 5. Single frame of night imagery of Mission Santa Barbara obtained with Hughes LWSS portable uncooled sensor. (From Hughes Santa Barbara Research Center with permission, © 1997, Santa Barbara Research Center.)

values of  $\sim 20$  mK for  $f/1.0$  apertures. The scene dynamic range of the camera is greater than 50 K with the  $f/0.7$  optics and greater than 100 K with the  $f/1.0$  optics.

The pixel operability for the same FPA was measured to be 99.2%. An operable pixel is defined as one with an  $f/0.7$  NETD less than 50 mK. The instantaneous scene temperature dynamic range of the pixels is typically about 50 K with  $f/0.7$  optics and about 100 K with  $f/1.0$  optics. A single-frame example of night imagery obtained from the LWSS camera is shown in Fig. 5.

#### Prospects for Semiconducting Yttrium Barium Copper Oxide (YBaCuO) Microbolometer Arrays

Although VO<sub>2</sub> technology has obtained impressive performance, there are difficulties associated with the deposition of the material across the wafer with uniform resistivity and TCR across the wafer. This, combined with the relatively low TCR which necessitates a large current bias and thereby power dissipation, has made the search for an alternative material. In general, the bolometric material should be easily deposited and patterned using standard semiconductor processing equipment. The processing of the bolometer material must also be compatible with Si micromachining processes and be compatible with the CMOS readout circuitry, which is typically fabricated underneath the bolometer array. These criteria require that the material be deposited uniformly over the wafer with only small spatial variations in the resistivity and TCR. Furthermore, all the processing would ideally be conducted at temperatures less than 300 °C so the underlying CMOS readout circuitry is not degraded. In addition, the bolometer material should have a high TCR to allow for small bias currents and, hence, low power dissipation and long battery life for the IR camera. YBaCuO is one material that satisfies these criteria.

YBaCuO is best known as a high-temperature superconductor. The optical and electronic properties of YBa<sub>2</sub>Cu<sub>3</sub>O<sub>7-x</sub> are determined by its oxygen stoichiometry. For  $x \approx 1$ , YBaCuO possesses an orthorhombic crystal structure, exhibits metallic conductivity, and becomes superconductive upon cooling below its critical temperature. As  $x$  is decreased to 0.5, the crystal

Table 1. Selected Properties of Semiconducting YBaCuO

| Parameter  | Value   |
|--|---|
| Conductivity activation energy (near room temperature)                     | $E_a \approx 0.2-0.3$ eV                            |
| Relative dielectric constant ( $T = 298$ K)                                | $\epsilon_r = 87-500$                               |
| Typical resistivity ( $T = 298$ K)   | $\rho = 1-100$ $\Omega\text{-cm}$                   |
| Hall carrier concentration $n_H$ ( $T = 298$ K)                            | $10^{18}-10^{20}$ $\text{cm}^{-3}$ p-type           |
| $dR/dT$ ( $T = 298$ K)   | $6 \times 10^3$ $\Omega/\text{K}$                   |
| Temperature coefficient of resistance ( $T = 298$ K)                       | $3.9\%$ $\text{K}^{-1}$                             |
| Voltage noise at 1 $\mu\text{A}$ current bias and 30 Hz (1/f noise regime) | $V_n/\Delta f < 0.76$ $\mu\text{V}/\text{Hz}^{1/2}$ |

\* The wide range in some of the parameters is due to the varying oxygen content.

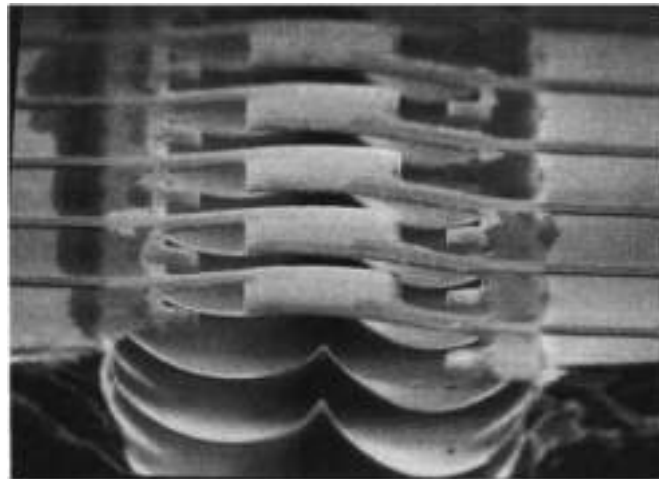
structure undergoes a phase transition to a tetragonal structure and it exhibits semiconducting conductivity characteristics because it exists in a Fermi glass state. As  $x$  is decreased further below 0.3, YBaCuO becomes a Hubbard insulator with a well-defined energy gap on the order of 1.5 eV (20). The unit cell consists of three CuO planes in the  $a$ - $b$  plane sandwiched between two planes containing BaO and one plane containing Y atoms along the  $c$ -direction. Each layer consists of corner-sharing CuO<sub>4</sub> polyhedra held together by the Y plane. As  $x$  is increased, O is randomly introduced to the O(1) sites creating carriers and simultaneously results in disorder, leading to formation of localized states in the CuO planes.

Several reports exist in the literature on transport, Hall effect, and dielectric measurements of the semiconducting YBaCuO thin films (21-24). A brief summary of results is presented in Table 1.

In the semiconducting state, YBaCuO exhibits a relatively large TCR ( $\sim 3$  to  $4\%$   $\text{K}^{-1}$ ) over a 60 K temperature range near room temperature. The large TCR, combined with the ease of thin film fabrication that is compatible with CMOS processing, makes YBaCuO attractive to microbolometer applications.

The YBaCuO thin films, whose data are reported here, were deposited by radio frequency (RF)-magnetron sputtering in a CYC-601 sputtering system at ambient temperature using commercially available YBa<sub>2</sub>Cu<sub>3</sub>O<sub>7-x</sub> sputter targets (10-13). Microbolometer array test structures were fabricated by depositing the YBaCuO IR-sensitive element onto a suspended bridge. A scanning electron micrograph of a typical array is shown in Fig. 6. In this case, the bridge structures were fabricated removing the Si underneath a 1500 nm thick SiO<sub>2</sub> by etching with a HF:HNO<sub>3</sub> solution through trench cuts in the oxide layer. A 40 nm thick MgO buffer layer was deposited by RF-magnetron sputtering, followed by the 200 nm thick YBaCuO film. The MgO buffer layer was used to act as a barrier against any interaction between YBaCuO and the substrate because some evidence of Cu diffusing into silicon had been observed earlier. In addition, lower noise characteristics have been observed in YBaCuO films deposited on MgO. The effect of encapsulating YBaCuO in MgO or other passivating films has yet to be tried. The YBaCuO and MgO was then patterned to define IR-sensitive areas by the wet etching. A 300 nm thick Au film was sputter-deposited and etched to form the contact pads and leads.

The two-probe resistance values for array pixels were found to vary roughly from 2 M $\Omega$  to 10 M $\Omega$  depending on geometry for these devices. The TCR (or  $\beta$ ) was calculated from the resistance versus temperature characteristics. For pixel #2 of the same array,  $\beta$  was found to vary from  $\sim 2.99\%$  to  $3.37\%$  in the

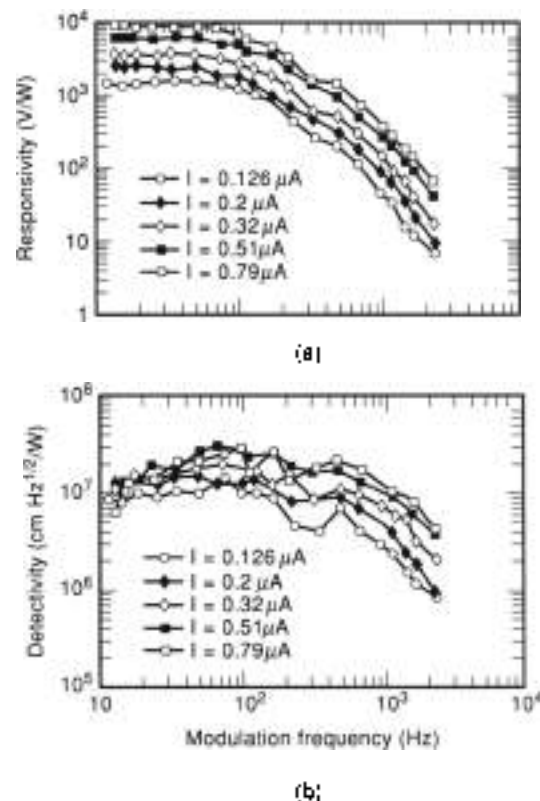


**Figure 6.** Scanning electron micrograph of a one-dimensional YBaCuO microbolometer array fabricated on an oxidized Si substrate. The YBaCuO IR-sensitive element is suspended on the micro-machined SiO<sub>2</sub> bridge and is contacted by gold leads running along the arms of the bridge. The pixel size is approximately 40  $\mu\text{m} \times 40 \mu\text{m}$ .

202 K to 312 K range with a room temperature value of 3.25%. The wafer average value of  $\beta$  was 3.5% and varied less than 10% across the wafer.

The optical response of the YBaCuO microbolometer arrays was measured by using a 1450 K blackbody source. The net usable range of this broad-band system was  $\sim 0.8 \mu\text{m}$  to  $12 \mu\text{m}$ . Narrow-band spectral analysis was performed with an Orlic MS-267 monochromator/spectrograph. The samples were characterized under both front-side and backside (through the substrate) illumination with mechanically chopped infrared light. Room temperature measurements of the responsivity,  $R_V$ , and detectivity,  $D$ , were performed in air. The temperature-dependent measurements were performed in a cryostat evacuated to a pressure of 30 mTorr. The response of the YBaCuO devices was calibrated against a pyroelectric detector with  $R_V = 1000 \text{ V/W}$ . In Fig. 7(a), typical room-temperature responsivity  $R_V$  is displayed at different bias levels of 0.126  $\mu\text{A}$  to 0.79  $\mu\text{A}$ . The responsivity was linear with respect to the bias current implying a true bolometric behavior. At high frequencies,  $R_V$  decreased in accordance with Eq. (2). A thermal time constant,  $\tau = 0.7 \text{ ms}$ , was obtained by fitting this relation to the measured response. A room-temperature value of  $G/\eta = 2.6 \times 10^{-6} \text{ W/K}$  was computed using Eq. (1) and the values of measured  $\beta$  and  $R_V$  computed  $\tau$  and the applied bias current  $I_b$ . The thermal conductance  $G$  of the suspended structure was measured by the resistive or Joule heating method to be  $7.41 \times 10^{-6} \text{ W/K}$ . From the measurement of the thermal conductance and time constant the thermal mass  $C$  was calculated to be  $\sim 10^{-6} \text{ J/K}$ . Using the value of  $G/\eta$  calculated from the responsivity data and  $G$  from Joule heating, a value of absorptivity  $\eta$  for the YBaCuO film was estimated to be about 20%. However, in order to obtain a more realistic figure, a direct measurement of absorptivity and reflectivity characteristics for this material needs to be performed.

Noise characteristics were also investigated to evaluate the performance of the bolometer in terms of detectivity  $D$ .  $D$  is displayed in Fig. 7(b). At zero-bias or very low currents ( $I_b < 0.1 \mu\text{A}$ ), the noise spectrum showed essentially the Johnson (or  $4kTR$ ) level. A  $1/f$  noise was observed at low frequencies



**Figure 7.** (a) Responsivity and (b) detectivity of 40  $\mu\text{m} \times 40 \mu\text{m}$  YBaCuO microbolometer pixel as a function of IR chopping frequency at different current biases.

( $\sim 2 \mu\text{V/Hz}^{1/2}$  at 30 Hz for 0.79  $\mu\text{A}$ ), which caused  $D$  to decline slightly. At higher frequencies (above  $\sim 300 \text{ Hz}$  for 0.79  $\mu\text{A}$ ), the excess noise spectral density merged with the Johnson noise floor of  $\sim 0.4 \mu\text{V/Hz}^{1/2}$ . From the point of view of detectivity, the optimum range for chopping frequency was found to be about 70 Hz to 200 Hz, above which the cutoff due to  $\tau$  caused  $R_V$  and, hence,  $D$  to fall below their maxima. A plausible cause of  $1/f$  noise was the contact resistance between gold metallization and the YBaCuO film. If this is true, then decreasing the contact resistance would bring the low-frequency noise down to the Johnson noise level, improving the performance of these devices by an order of magnitude.

It is important to note that the measured thermal conductance of the YBaCuO microbolometer test structures is approximately two orders of magnitude larger than the state-of-the-art obtained elsewhere for amorphous-Si and VO<sub>x</sub> bolometers. This would imply that detectivities in the range of  $10^9$  to  $5 \times 10^9 \text{ cm Hz}^{1/2}/\text{W}$  are achievable if state-of-the-art thermal isolation structures were employed. This is comparable to VO<sub>x</sub> and amorphous-Si devices. It is also important to note that the top surface of the YBaCuO bolometers was not passivated and likely suffers from a larger  $1/f$  noise component than if the YBaCuO film was totally encapsulated. VO<sub>x</sub> technology uses Si<sub>3</sub>N<sub>4</sub> encapsulation, in part, to reduce the  $1/f$  noise present in the film. For YBaCuO, it has been observed that utilizing the MgO buffer layer reduces the  $1/f$  noise present in the film, and it is likely that MgO encapsulation would further reduce the  $1/f$  noise and lead to higher detectivities. The different microbolometer technologies are compared in Table 2 with respect to the TCR and corner frequency.

**Table 2. A Comparison Between Uncooled Bolometer Technologies**

| Material              | TKR<br>(% K <sup>-1</sup> ) | $\rho/4kT$<br>(Ea/W) |
|-----------------------|-----------------------------|----------------------|
| VSI                   | 1.5-4                       | $1.2 \times 10^{-4}$ |
| Amorphous Si          | 1.8                         | $4 \times 10^{-4}$   |
| Semiconducting YBaCuO | 3-4                         | $2 \times 10^{-4}$   |

\* From Refs. 6, 10-11, and 12.

## PYROELECTRIC IR DETECTORS

### Theory of Pyroelectric Detector Operation

A pyroelectric detector operates through the temperature dependence of the spontaneous polarization of the material. As the polarization changes with temperature, the surface charge on a pyroelectric capacitor changes, resulting in a measurable pyroelectric current,

$$I = pA dT/dt \quad (12)$$

where  $p$  is the pyroelectric coefficient,  $A$  is the capacitor area,  $T$  is temperature, and  $t$  is time. The pyroelectric capacitor is connected to a preamplifier. A typical pyroelectric detector element of area  $A$  and dielectric thickness  $d$  connected to a read-out amplifier is shown in Fig. 8. Radiation with power  $\Phi(t)$  modulated at frequency  $\omega$  is incident on surface of the element. The absorptivity of the detector is  $\eta$ . The detector element has a thermal capacity  $C$  and a thermal conductance to the substrate  $G$ . Thus, the temperature change,  $\Delta T$ , resulting from the incident radiant flux can be expressed as

$$C \frac{d\Delta T}{dt} + G\Delta T = \eta A \Phi(t) \quad (13)$$

where the temperature of the detector element is  $T(t) = T_0 + \Delta T(t)$ , modulated by the radiation flux

$$\Phi(t) = \Phi_{av} + \Delta\Phi(t) = \Phi_{av} + \Phi_0 e^{i\omega t} \quad (14)$$

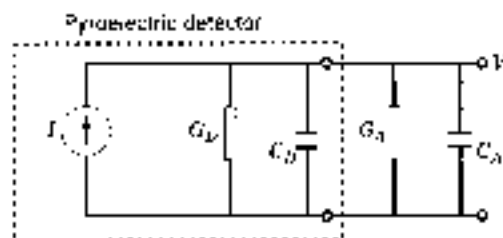
The solution to Eq. (13) is

$$\Delta T = T_0 e^{i\omega t} \quad (15)$$

with

$$T_0 = \frac{\eta}{G} \frac{1}{\sqrt{1 + \omega^2 \tau_{th}^2}} \Phi_0 \quad (16)$$

The thermal time constant  $\tau_{th}$  is defined as  $\tau_{th} = C/G$ .



**Figure 8.** The equivalent circuit depicting the detector and the read-out amplifier.

The response is usually read out by an amplifier that can be shown as a combination of conductance  $G_A$  and capacitance  $C_A$  in parallel with the detector conductance  $G_D$  and capacitance  $C_D$  (Fig. 8). If the induced pyroelectric current  $I$  is read out as voltage  $V$ , then

$$C_E \frac{dV}{dt} + G_E V = I = pA \frac{dT}{dt} \quad (17)$$

Here,  $C_E = C_D + C_A$ , and  $G_E = G_D + G_A$  are the equivalent electrical capacitance and conductance of the circuit in Fig. 1. Solving Eq. (17) with Eq. (15), we find

$$\Delta V(t) = V_0 e^{i\omega t} \quad (18)$$

and

$$V_0 = \frac{pA}{C_E} \frac{\omega \tau_E}{\sqrt{1 + \omega^2 \tau_E^2}} T_0 \quad (19)$$

The electrical time constant,  $\tau_E$ , is expressed by  $\tau_E = C_E/G_E$ . Combining Eqs. (16) and (19), voltage responsivity  $R_V$  can be obtained as (26-29)

$$R_V = \frac{V_0}{\Phi_0} = \eta \frac{pA}{GG_E} \frac{\omega}{\sqrt{1 + \omega^2 \tau_{th}^2} \sqrt{1 + \omega^2 \tau_E^2}} \quad (20)$$

It should be noted that in solving Eqs. (13) and (17), phase terms were ignored.

To achieve a high responsivity, the detector is usually fabricated in a microfabricated thermal isolation structure, minimizing  $G$ . As discussed before, the detectivity,  $D^* = R_V / (A \Delta f)^{1/2} \cdot V_N$ , gives the area-normalized signal-to-noise ratio in the frequency bandwidth  $\Delta f$  for the detector with  $V_N$ , the total noise voltage. Pyroelectric detectors experience Johnson noise, background noise in the photon flux, and temperature fluctuation noise. In practical cases, temperature fluctuation noise includes the background noise contributions and should be ideally the dominant noise source (30). The respective noise contributions may be calculated from the following relations (31).

**Temperature or Radiation Noise.** Even if the pyroelectric detector is in thermal equilibrium with its surroundings, it will experience agitation of charges due to (a) fluctuations in the incoming radiation and (b) exchange of heat due to convection and conduction. The temperature noise can be expressed as (26)

$$\frac{\Delta V_{TE}}{\sqrt{\Delta f}} = \frac{R_V \sqrt{4kT^2 G}}{\eta} \quad (21)$$

If the temperature fluctuations due to convection and conduction are eliminated so that the only interchange of energy with the surroundings is by radiation, the detector is said to be background-limited, representing the highest achievable detectivity  $D_{BLIP}^*$  at a particular temperature (26,28)

$$D_{BLIP}^* = \left( \frac{\eta^2}{16\pi kT^2} \right)^{1/2} \quad (22)$$

Here,  $\sigma$  is Stefan's constant. At room temperature, for  $n = 1$ , the background-limited detectivity is  $1.8 \times 10^{10} \text{ cm Hz}^{-1/2} \text{ W}^{-1}$ .

**Johnson Noise.** Johnson noise includes the thermal noise of the parallel conductances of the amplifier resistance and the alternating current (ac) conductance of the detector and is given by (26):

$$\frac{\Delta V_J}{\sqrt{\Delta f}} = \frac{\sqrt{4kT G_A}}{G_A \sqrt{1 + \omega^2 \tau_A^2}} \quad (27)$$

Since  $G_A = G_A + G_D = G_A + \omega C_D \tan \delta$ , depending on the frequency-dependence of the detector conductance, the spectral density of the Johnson noise in pyroelectric detection systems can take several shapes including  $1/f$  and Lorentzian forms. The term  $\tan \delta$  refers to the loss tangent of the material.

Johnson noise frequently dominates in pyroelectric detectors. For high frequencies such that ac conductance of the detector is much higher than the amplifier conductance,  $\omega \gg (C_D \tan \delta / G_A)^{-1}$ ,  $\omega \gg \tau_A^{-1}$ , and  $C_D \gg C_A$ , the Johnson noise can be written as:

$$\frac{\Delta V_J}{\sqrt{\Delta f}} = \left( 4kT \frac{\tan \delta}{C_D} \right)^{1/2} \omega^{-1/2} \quad (28)$$

Consequently, for this specific case, the detectivity is given through Eqs. (3), (20), and (24):

$$D^* = \frac{\omega l}{\sqrt{4kT}} \omega^{1/2} \frac{1}{\epsilon} \frac{p}{\sqrt{\epsilon_0 \tan \delta}} \quad (29)$$

Here,  $\rho$  is the volume specific heat,  $l$  is the thickness of the pyroelectric element,  $\epsilon$  is the relative permittivity of the pyroelectric material, and  $\epsilon_0$  is the permittivity of free space. Therefore, in order to maximize the detectivity for this region, it is desirable to maximize the last term in the above expression, which is sometimes referred to as  $F_{\text{eff}}$ , one of the figures of merit for a pyroelectric detector (26,32).

The reader should be cautioned, however, that the conditions assumed in calculating Eq. (24) are frequently not applicable to small-area imaging detectors operating at camera frame frequencies (50 Hz or 60 Hz). As an example, consider a pyroelectric detector with  $C_D = 5 \text{ pF}$  and with  $\tan \delta = 0.01$ , coupled to an amplifier with  $\tau_A = 10 \text{ pF}$  and  $R_A = 10^9 \Omega$ . In this case, the Johnson noise is determined by the amplifier circuit capacitance and has a value of about  $0.4 \mu\text{V} \cdot \text{Hz}^{-1/2}$  at 30 Hz.

**Amplifier Noise.** An amplifier noise is produced in the electronic amplifier used in the read-out circuitry. It can be due to  $1/f$  noise, generation recombination noise, or shot noise arising from the field effect transistor (FET) gate leakage current. At camera frame frequencies, we expect this component to be less than the detector noise (20).

The total equivalent noise voltage is given as the squared sum of individual noise components,  $V_{\text{eq}}^2 / \Delta f = \Delta V_{\text{NE}}^2 / \Delta f + \Delta V_{\text{FET}}^2 / \Delta f + \Delta V_{\text{JN}}^2 / \Delta f$ . One of the most important issues in systems integration of pyroelectric detectors is to minimize the noise sources by careful matching of the amplifier to the detector, adjusting pyroelectric and dielectric properties of the pyroelectric material, and careful designing of the suspended structure and therefore thermal characteristics of the detector element.

A more relevant figure of merit for a complete IR camera, which takes into account the optics, focal plane array performance, and read-out electronics, is the noise equivalent temperature difference (NETD), as discussed earlier (30).

A list of commonly used materials for pyroelectric application include triglycine sulfate (TGS), lithium tantalate (LiTaO<sub>3</sub>), Ba<sub>1-x</sub>Sr<sub>x</sub>TiO<sub>3</sub> (BST), Pb<sub>1-x</sub>La<sub>x</sub>Zr<sub>1-x</sub>(Ti<sub>1-x</sub>)<sub>2</sub>O<sub>7</sub> (PLZT), and PbTiO<sub>3</sub>. In addition, the semiconducting phase of yttrium barium copper oxide (YBaCuO) is a new promising pyroelectric material. Pyroelectric detectors can be operated without a bias below their Curie temperature (normal pyroelectrics) or at Curie temperature with bias (phase transition materials). The latter requires stringent bias and temperature stabilization requirements. Next generation uncooled pyroelectric detectors are required to operate without bias and temperature stabilization in a wide range of environmental conditions, and therefore should belong to the normal pyroelectrics family. In addition, it is desired to have a thin-film pyroelectric detector to be able to fully utilize the state-of-the-art microfabricating technology for fabrication of focal plane arrays. Table 3 lists some of the commonly used pyroelectric materials and their pyroelectric coefficients.

The following sections will summarize some of the most commonly used pyroelectric detector technology for IR focal plane arrays.

**Lead Titanate-Based Detectors**

Research groups led by Polla (36) and Takayama (41) independently developed techniques for depositing PbTiO<sub>3</sub> thin films with high pyroelectric coefficient, detectivity, and fast response time. PbTiO<sub>3</sub> is a pyroelectric material with a Curie temperature of about 490 C. It is usually operated with no applied electric field and well below its transition temperature exploiting

**Table 3. Room Temperature Pyroelectric Coefficients of Most Commonly Used Pyroelectric Materials**

| Material   | $p$<br>( $\mu\text{C}/\text{cm}^2/\text{K}$ ) | Comments  | Reference |
|--|---|---|-----------|
| TGS  | 0.028   | Single crystal, bulk, normal pyroelectric             | 33        |
| LiTaO <sub>3</sub>   | 0.19  | Single crystal, bulk, normal pyroelectric             | 34        |
| KTN (KTa <sub>1-x</sub> Nb <sub>x</sub> O <sub>3</sub> )   | 0.01  | Polycrystalline, thin film, with bias                 | 35        |
| PbTiO <sub>3</sub> sol-gel   | 0.005   | Polycrystalline, thin film, normal pyroelectric       | 36, 37    |
| PScT (Pb <sub>1-x</sub> Sc <sub>x</sub> Ta <sub>1-x</sub> O <sub>3</sub> )                                   | 0.5-4.6                                       | Polycrystalline, thin film, phase transition material | 38, 39    |
| BST (Ba <sub>1-x</sub> Sr <sub>x</sub> TiO <sub>3</sub> )  | 2:  | Ceramic, bulk, phase transition material              | 32, 40    |
| PLT (Pb <sub>1-x</sub> La <sub>x</sub> TiO <sub>3</sub> )  | 0.065   | Polycrystalline, thin film, normal pyroelectric       | 41        |
| PLZT (Pb <sub>1-x</sub> La <sub>x</sub> Zr <sub>1-x</sub> (Ti <sub>1-x</sub> ) <sub>2</sub> O <sub>7</sub> ) | 0.13-0.18                                     | Ceramic, bulk, normal pyroelectric                    | 42, 43    |
| YBaCuO   | 16  | Polycrystalline, thin film, normal pyroelectric       | 44, 45    |

the normal pyroelectric effect where the rate of change in the spontaneous polarization with respect to temperature is measured. The latest-generation pyroelectric devices are thin films deposited on Si air-gap bridge structures for low thermal mass, low thermal conductivity, and hence increased responsivity. Since these devices are made on silicon, they can be directly integrated into silicon signal processing circuitry without the need of pump-bonds or wires.  $\text{PbTiO}_3$  is typically deposited using sol-gel processing techniques with titanium-platinum (Ti-Pt) or Pt electrodes. Investigations carried out by Polla's group revealed a pyroelectric coefficient  $p$  of  $95 \text{ nC/cm}^2 \text{ K}$ , a relative dielectric coefficient  $\epsilon_r$  of 200, responsivity  $R_p$  of  $10^4 \text{ V/W}$ , and noise voltage  $V_n$  of  $10^{-2} \text{ V/Hz}^{1/2}$  at 50 Hz for thin film  $\text{PbTiO}_3$ . These values compare favorably over the ones for other pyroelectric thin films like ZnO and  $\text{PbZr}_{0.54}\text{Ti}_{0.46}\text{O}_3$  (PZT). Test arrays of  $64 \times 64 \text{ PbTiO}_3$  elements, each  $30 \times 30 \mu\text{m}^2$  in size, have been fabricated by surface microfabrication techniques, on an NMOS readout circuitry, resulting in  $1.2 \times 10^4 \text{ V/W}$  responsivity and  $2 \times 10^6 \text{ cm Hz}^{1/2}/\text{W}$  normalized detectivity at 30 Hz (37). The measured response for a single pixel element was  $90 \text{ nC/cm}^2 \text{ K}$ , while the response due to the combined effect was  $60 \text{ nC/cm}^2 \text{ K}$ . Although this is far from background-limited operation, it shows promising progress toward lead titanate-based FPAs. There are, however, some problems. Reproducibility of  $\text{PbTiO}_3$  thin-film deposition is low, requiring special attention to the deposition conditions and the chemical stability of electrode interfaces.

#### Barium Strontium Titanate Pyroelectric Detectors

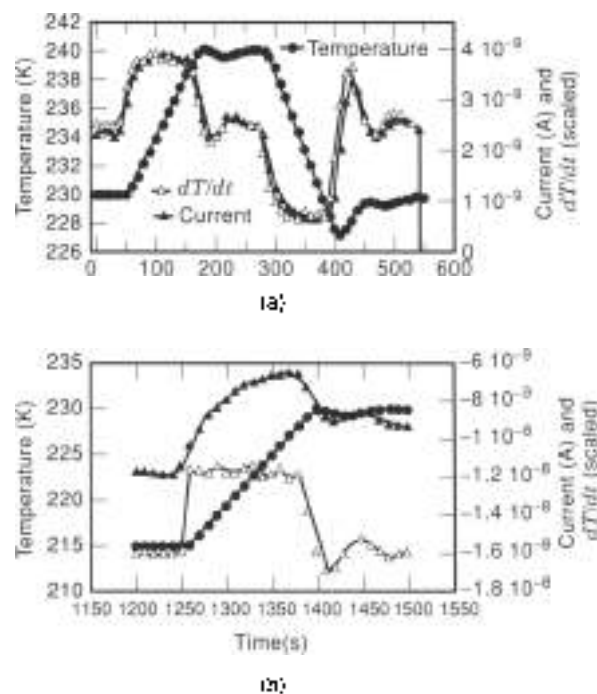
Another state-of-the-art pyroelectric detector technology is  $\text{Ba}_{1-x}\text{Sr}_x\text{TiO}_3$  thick films. Unlike  $\text{PbTiO}_3$ ,  $\text{Ba}_{1-x}\text{Sr}_x\text{TiO}_3$  is operated near the paraelectric-ferroelectric phase transition which can be adjusted to a temperature between 20°C and 30°C, depending on the value of  $x$ . The rate of change in the dielectric permittivity with respect to temperature is measured with an applied electric field. These detectors are also sometimes referred to as dielectric bolometers. Research groups at Raytheon-TI Systems fabricate IR cameras made of  $245 \times 328$  pixel arrays with an array-average NETD of 0.070°C with  $f/1.0$  optics (29,40). However, the present  $\text{Ba}_{1-x}\text{Sr}_x\text{TiO}_3$  technology is a cumbersome bulk ceramic technology which requires grinding and polishing of ceramic wafers sliced from a boule, laser reticulation of pixels, multiple thinning, and planarization steps. The array is connected to the silicon readout circuit by compression bonds. The process suffers from (a) thermal isolation problems due to the thick mesa structure and (b)  $\text{Ba}_{1-x}\text{Sr}_x\text{TiO}_3$  surface degradation due to the thinning procedure.

#### Prospects for Semiconducting Yttrium Barium Copper Oxide Pyroelectric Detectors

As discussed in the bolometer section,  $\text{YBa}_2\text{Cu}_3\text{O}_{6-x}$  belongs to a class of copper oxides well known for their superconducting properties. Its conduction properties can be changed from metallic ( $0.5 < x < 1$ ) to insulating ( $0 \leq x \leq 0.5$ ) by suitably decreasing the oxygen content (20). Pyroelectric  $\text{YBaCuO}$  thin films have been fabricated by RF sputtering from a commercially available superconducting target on top of suspended  $\text{SiO}_2$  bridges, using an  $\text{HF}/\text{HNO}_3$  etch to remove the underlying Si (44). Detectivities as high as  $10^6 \text{ cm-Hz}^{1/2}/\text{W}$  have been measured for a thermal conductance ( $G \approx 10^{-7} \text{ W/K}$ ). Typically,

$G \approx 10^{-7} \text{ W/K}$  is achievable with the state-of-the-art surface microfabrication techniques, indicating that detectivities of  $10^{10} \text{ cm-Hz}^{1/2}/\text{W}$  are possible using semiconducting  $\text{YBaCuO}$ . For a focal plane array made of  $40 \times 40 \mu\text{m}^2$  pixels, operating in the  $6 \mu\text{m}$  to  $14 \mu\text{m}$  atmospheric window [ $(\Delta P/\Delta T)_{\lambda_1-\lambda_2} = 2.6 \times 10^{-4} \text{ W/cm}^2 \text{ K}$ ], employing  $f/1$  optics with a transmittance close to 1, and clipped at 80 Hz clipper frequency, NETD values of 2 mK are achievable (Eq. (11)). This performance is comparable to the background noise limited performance calculated for the microbolometer.  $\text{YBaCuO}$  combines an ease of fabrication that should lead to a reduction in cost. Pyroelectric coefficients as high as  $18 \mu\text{C/cm}^2 \text{ K}$  have been measured (45). The ability to pole the devices to obtain pyroelectric coefficients in this range was also demonstrated.

Strong pyroelectric effect in nonmetallic  $\text{YBaCuO}$  has also been confirmed by many researchers (46–48). Although the exact origin of the observed pyroelectric effect is not fully known at this time, it can be attributed to the noncentrosymmetric nature of the material with sufficiently long-range Coulomb ordering to give a net dipole moment (47). Pyroelectricity has been shown to exist in both metallic and insulating phases of  $\text{YBa}_2\text{Cu}_3\text{O}_{6-x}$  (44,49). Despite some controversy (49) with regard to the origin of pyroelectricity in  $\text{YBaCuO}$ , it is believed to be associated with symmetry-breaking in the unit cell due to the anharmonicity of the apex O(4) oxygen site which is present in the branch connecting the  $\text{CuO}_2$  and basal planes of the  $\text{YBa}_2\text{Cu}_3\text{O}_{6-x}$  molecule (50). The cause of this behavior is probably due to the randomly filled O(5) defect site. The nonsymmetric displacement of O(4) mode results in a net polarization in the molecule with domains of dipoles scattered throughout the bulk of the material. This microscopic polarization appears at the macroscopic level when either strain by an externally applied electric field or mechanical strain acquired during the fabrication process provides the impetus for the domains to be



**Figure 9.** (a) Pyroelectric response of  $\text{YBCO}$  planar capacitor from 230 K to 240 K.  $p = 400 \text{ nC/cm}^2 \text{ K}$ . (b) Pyroelectric response of the same device after poling at 5000 V/cm at 320 K for 1 h.  $p = 18 \mu\text{C/cm}^2 \text{ K}$ .

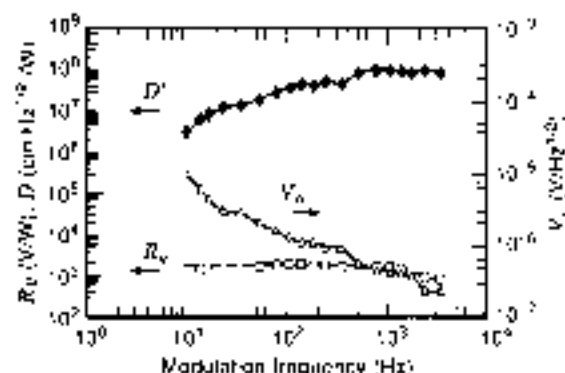


Figure 10. Voltage responsivity  $R_v$ , detectivity  $D$ , and noise  $V_n$  for a suspended Nb:YBaCuO/Nb pyroelectric detector versus chopper frequency at room temperature. Theoretical fit to  $R_v = \omega \alpha A R_p / G(k) [1 + \omega^2 \tau^2 / 31 + \omega^2 \tau_p^2 / 4]^{1/2}$  is shown as the dashed line.

lined up in one direction. In the *strain-poled* case, a sample would show pyroelectric behavior without the application of any poling bias.

Pyroelectricity is often associated with changes in individual crystal symmetry, leading to a net polarization. Mihalovic et al. (47) suggest this as a possible mechanism in their measurements on different stoichiometries of single-crystal YBCO. They found that the pyroelectric response increases with added oxygen content, applied electric field, and prior poling. Another report by Kumar et al. (51) examined the effects of temperature cycling on the properties of metallic YBCO thin films as measured by the photopyroelectric effect. Most relevant to this experiment is the changing of the charge carrier properties with respect to temperature, caused by thermal annealing at 325 K. At this temperature, the oxygen organization in the CuO planes changes as does the density of carriers near the Fermi surface. Furthermore, the changes in these properties do not reverse with respect to temperature for an extended period of time. The result is annealing-induced hysteresis in the thermal and electronic characteristics with respect to temperature.

Figure 9(a) shows the pyroelectric response of a YBaCuO planar capacitor from 230 K to 240 K,  $p \approx 400 \text{ nC/cm}^2 \text{ K}$ . Figure 9(b) is the pyroelectric response of the same device after poling at 5000 V/cm at 320 K for 1 h,  $p \approx 18 \text{ } \mu\text{C/cm}^2 \text{ K}$ .

Voltage responsivity  $R_v$ , detectivity  $D$ , and noise  $V_n$  versus chopper frequency at room temperature is shown in Fig. 10 for a semiconducting YBaCuO layer sandwiched between two Nb electrodes, fabricated on top of a thermally isolated  $\text{SiO}_2$  bridge. Theoretical fit to Eq. (20) is depicted as the dashed line.

## SUMMARY

The recent accomplishments in the development of microbolometer and pyroelectric detector arrays for uncooled infrared have demonstrated that high-performance cameras that are more economical for consumer applications are a reality. Although the detectivity of uncooled cameras will always be lower than their cryogenic counterparts, the achievement of NETDs of 30 mK or less is superior to current HgCdTe technology. The performance of these uncooled cameras is suitable for a large number of consumer applications for night vision in transportation, policing and security, as well as for thermal imaging in medicine. The commercial revenue for uncooled infrared

cameras was expected to grow to be in excess of \$100 million by the year 2001 (52). With the growth of the market and consumer demand, the cost of the cameras is expected to fall to be within reach of the mass consumer application of automotive night vision. As the market grows, new materials that offer fabrication and performance advantages over the current technologies will be developed.

## ACKNOWLEDGMENTS

We would like to express our thanks to William Radford of Hughes Santa Barbara Research Center, Charles Marshall of Lockheed Martin Infrared Systems, and Charles Hanson of Raytheon-TI Systems.

## BIBLIOGRAPHY

1. K. C. Liddiard, Thin-film resistance bolometer IR detectors II, *Infrared Phys.*, **26**: 43-49, 1986.
2. A. Tannik et al., Infrared focal plane array incorporating silicon IC process compatible bolometer, *IEEE Trans. Electron Devices*, **43**: 1644-1650, 1996.
3. T. A. Eureka et al., Amorphous silicon and germanium films for uncooled microbolometers, *Tech. Phys. Lett.*, **23**: 504-506, 1997.
4. V. Y. Zernov et al., Bolometric properties of silicon thin-film structures fabricated by plasma-enhanced vapor-phase deposition, *Tech. Phys. Lett.*, **23**: 481-483, 1997.
5. W. Radford et al., 320 x 240 silicon microbolometer uncooled IR FPNs with on-chip offset correction, *Proc. SPIE*, **2746**: 82-82, 1996.
6. C. Marshall, et al., Uncooled infrared sensor with digital focal plane array, *Proc. SPIE*, **2746**: 20-31, 1996.
7. R. J. Herring and P. E. Howard, Design and performance of the ULTRA 320 x 240 uncooled focal plane array and sensor, *Proc. SPIE*, **2746**: 2-12, 1996.
8. W. Meyer et al., Ambers' uncooled microbolometer LWIR camera, *Proc. SPIE*, **2746**: 13-22, 1996.
9. R. A. Wood, High performance infrared thermal imaging with amorphous silicon focal planes operating at room temperature, *Int. Electron Devices Meet., Washington, DC*, 1993, pp. 176-177.
10. P. C. Shan et al., Semiconducting YBaCuO thin films for uncooled bolometers, *J. Appl. Phys.*, **78**: 7334, 1995.
11. P. C. Shan et al., The investigation of semiconducting YBaCuO thin films: A new room temperature bolometer, *J. Appl. Phys.*, **80**: 7118-7123, 1996.
12. C. M. Travers et al., Fabrication of semiconducting YBaCuO surface micromachined bolometer arrays, *IEEE/ASME J. Microelectromech. Syst.*, **6**: 271-276, 1997.
13. A. Jahanzeb et al., A semiconductor YBaCuO microbolometer for room temperature IR imaging, *IEEE Trans. Electron Devices*, **44**: 1755-1801, 1997.
14. J. D. Vincent, *Fundamentals of Infrared Detector Operation and Testing*, New York: Wiley, 1990, chap. 3.
15. B. R. Johnson et al., YBa<sub>2</sub>Cu<sub>3</sub>O<sub>7</sub> superconducting microbolometer arrays fabricated by silicon micromachining, *IEEE Trans. Appl. Supercond.*, **3**: 2656, 1993.
16. J. P. Rice et al., High-T<sub>c</sub> superconducting antenna-coupled microbolometer on silicon, *Proc. SPIE*, **2159**: 98-109, 1994.
17. R. T. Howe, Surface micromachining for microsensors and micromanipulators, *J. Vac. Sci. Technol.*, **B6**: 1909, 1989.
18. E. N. Hooge,  $1/f$  noise is no surface effect, *Phys. Lett.*, **20A**: 119, 1969.

19. A. van der Ziel, Flicker noise in semiconductors: Not a true bulk effect. *Appl. Phys. Lett.*, **33**: 687, 1976.
20. For a review article see G. Yu and A. J. Heeger, Photoinduced charge carriers in insulating cuprates, Ferri glass insulator, metal-insulator transition and superconductivity. *Int. J. Mod. Phys. B*, **7**: 3751, 1993.
21. Z. Çelik-Butler, W. Yang, and D. P. Butler, Measurements of noise and temperature coefficient of resistance in YBaCuO thin films in magnetic field. *Appl. Phys. Lett.*, **60**: 245, 1992.
22. A. Jahanzeb et al., Studies and implications of the Hall-effect in superconducting and semiconducting  $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$  thin films. *J. Appl. Phys.*, **76**: 6656, 1995.
23. Z. Çelik-Butler et al., Charge transport in amorphous and tetragonal semiconducting YBCO films. *Solid State Electron.*, **41**: 695-699, 1997.
24. P. C. Shan et al., Hall-effect in semiconducting epitaxial and amorphous Y-Ba-Cu-O thin films. *J. Appl. Phys.*, **81**: 6666-6673, 1997.
25. D. L. Craig et al., Anisotropic excess noise within a Si:H. *Solid State Electron.*, **39** (6): 607-612, 1996.
26. K. W. Whitmore, Pyroelectric devices and materials. *Rep. Prog. Phys.*, **40**: 1335-1365, 1986.
27. A. Hassan and M. H. Rashid, Pyroelectric detectors and their applications. *IEEE Trans. Inf. Appl.*, **27**: 824-829, 1991.
28. D. E. Marshall, A review of pyroelectric detector technology. *Proc. SPIE*, **132**: 110-116, 1979.
29. C. M. Hanson, Hybrid Pyroelectric-Ferroelectric Bolometer Arrays. In D. Skatrud and P. W. Kruse (eds.), *Semiconductors and Sensors*, San Diego, CA: Academic Press, Vol. 47, 1997.
30. P. W. Kruse, A comparison of the limits to the performance of thermal and photon detector imaging arrays. *Infrared Phys. Technol.*, **36**: 669-682, 1995.
31. M. C. Foote et al., Transition edge  $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$  microbolometers for infrared staring arrays. *Proc. SPIE*, **2150**: 2, 1994.
32. B. M. Kolwiski et al., Pyroelectric imaging. *IEEE Symp. Appl. Ferroelectr.*, Aug. 1992.
33. J. M. Herbert, *Ferroelectric Transducers and Sensors*, New York: Gordon and Breach, 1982, pp. 266-290.
34. R. L. Byer and C. B. Round, Pyroelectric coefficient direct measurement technique and application to a fast response time detector. *Ferroelectrics*, **8**: 333-336, 1972.
35. J. Mantese et al., Infrared imaging using uncooled focal plane arrays of unreticulated, 10- $\mu\text{m}$  potassium tantalum niobate films. *IEEE Trans. Electron Devices*, **40**: 320-324, 1993.
36. C. Ye, T. Tamagawa, and D. J. Pata, Experimental studies on primary and secondary pyroelectric effects in  $\text{PbZr}_{0.9}\text{Ti}_{0.1}\text{O}_3$ ,  $\text{PbTiO}_3$ , and  $\text{ZnO}$  thin films. *J. Appl. Phys.*, **70**: 5536-5544, 1991.
37. I. Phan et al., Surface micromachined pyroelectric infrared imaging array with vertically integrated signal processing circuitry. *IEEE Trans. Ultrason. Ferroelectr. Prop. Control*, **41**: 552-555, 1994.
38. K. Wotton and M. A. Todri, Induced pyroelectricity in sputtered lead scandium tantalate films and their merit for IR detector arrays. *Ferroelectrics*, **118**: 279, 1991.
39. C. Hjemander, A. M. Grishin, and E. V. Han, Pyroelectric  $\text{PbSn}_{1-x}\text{Ta}_x\text{O}_3/\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$  thin film structures. *Appl. Phys. Lett.*, **61**: 58, 1995.
40. C. M. Hanson et al., Uncooled thermal imaging at Texas Instruments. *Proc. SPIE Int. Soc. Opt. Eng.*, 1992, pp. 1-10.
41. R. Takayama et al., Pyroelectric properties and application to infrared sensors of  $\text{PbTiO}_3$ ,  $\text{PbLnTiO}_3$  and  $\text{PbZrTiO}_3$  ferroelectric thin films. *Ferroelectrics*, **118**: 325-342, 1991.
42. S. T. Liu, J. D. Hepp, and O. N. Toffe, The pyroelectric properties of the lanthanum-doped ferroelectric PLZT ceramics. *Ferroelectrics*, **3**: 261-295, 1974.
43. K. K. Deb, Investigation of pyroelectric characteristics of modified  $\text{PbTiO}_3$  ceramics for improved IR detector performance. *Ferroelectrics*, **88**: 167-176, 1996.
44. A. Jahanzeb et al., Sensing pyroelectric response of semiconducting Y-Ba-Cu-O and its application to uncooled infrared detection. *Appl. Phys. Lett.*, **70**: 3495-3497, 1997.
45. J. P. Gray, Z. Çelik-Butler, and D. P. Butler, Piezoelectric and pyroelectric response in  $\text{Nb}(\text{YKCaTi})_2\text{Nb}$  heterostructures. *Ferroelectrics*, **100**: 8.
46. D. Milicevic and A. J. Heeger, Pyroelectric and piezoelectric effects in single crystals of  $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$ . *Solid State Commun.*, **75**: 319-323, 1990.
47. B. Milicevic, I. Poberig, and A. Mertelj, Characterization of the pyroelectric effect in  $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$ . *Phys. Rev. B*, **48**: 16, 634-16, 640, 1993.
48. A. I. Grachev and L. V. Pleshchikov, Pyroelectric voltages in YBCO thin films. *Solid State Commun.*, **101**: 507-512, 1997.
49. I. Poberaj and D. Michanovic, Pyroelectric effect measurements in  $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$  and  $\text{La}_2\text{CuO}_4$  materials. *Ferroelectrics*, **128**: 197-200, 1992.
50. D. Milicevic and C. M. Foster, Anharmonicity and frequency shift of the apex oxygen O(4) Raman mode in  $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$  as a function of doping. *Solid State Commun.*, **74**: 753, 1990.
51. S. Kumar et al., Effect of thermal cycling on normal state thermal properties of  $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$  films. *Physica C*, **215**: 285-290, 1991.
52. E. Cochran, Research Technology, private communication.

ZEHNER ÇELİK BUTLER  
Lawrence Berkeley National  
Laboratory  
DONALD P. BUTLER  
Southern Methodist University

## MAGNETIC SENSORS

Magnetic sensors find many applications in everyday life and in industry. They provide convenient, noncontact, simple, rugged, and reliable operations compared to many other sensors. The technology to produce magnetic sensors involves many aspects of different disciplines such as physics, metallurgy, chemistry, and electronics.

Generally, magnetic sensors are based on sensing the properties of magnetic materials, which can be done in many ways. For example, magnetization, which is the magnetic moment per volume of materials, is used in many measurement systems by sensing force, induction, field methods, and superconductivity. However, the majority of industrial sensors make use of the relationship between magnetic and electric phenomenon. A typical application of the phenomenon is the computer memory requiring the reading of the contents of a disk without making any contact between the sensor and the device. In other applications, the position of objects sensitive to magnetic fields (e.g., the metals in the ground) can be sensed magnetically. Magnetic sensors find most sensitive applications in medicine to diagnose human illnesses, as in the case of superconducting quantum interference devices (SQUID) and nuclear resonance magnetic (NMR) imaging.

The magnetic elements in sensors are used in a wide range of forms, toroids, rods, films, substrates, and coatings. Some elements are essentially free standing, whereas others are an integral part of more complex devices. In order to obtain maximum



**Table 1. List of Manufacturers**

Arisek Tech, Inc.  
18310 Buford Circle  
La Porte, CA 91744  
Pnx: 818 854 2778

Analog Devices, Inc.  
1 Technology Way  
P.O. Box 9106  
Norwood, MA 02062-9102  
Tel: 800-252-5663  
Fax: 781 326 8700

Dynalco Controls  
3690 N. W. 53rd Street  
Ft. Lauderdale, FL 33309  
Tel: 305-799-4300 & 900-363-6666  
Fax: 705 454 3376

Electric Corporation  
1845 57th Street  
Sarasota, FL 34243  
Tel: 813-356-8411 & 900-446-5762  
Fax: 813-356-3120

Homegwell  
Dept. 722  
11 West Spring Street  
Fremont, IL 61032  
Tel: 909-597-6945  
Fax: 815 235 5988

Kaman Instrument Company  
1500 Garden of the Gods Road  
Colorado Springs, CO 80907  
Tel: 719-599-1132 & 800-562-6267  
Fax: 719-599-1824

Kalher Corporation  
14501 Los Angeles Avenue  
Moorpark, CA 93021  
Tel: 805-523-2000  
Fax: 805-523-7125

Lucas  
1080 Lucas Way  
Hampton, VA 23656  
Tel: 800-745-8008  
Fax: 800-745-8004

Motion Sensors, Inc.  
786 Pitts Chapel Road  
Elizabeth City, NC 27909  
Tel: 919-331-2680  
Fax: 919-331-1686

Rechner Electronics Industries, Inc.  
4651 Buffalo Avenue  
Niagara Falls, NY 14304  
Tel: 800 544 4106  
Fax: 716-293-2127

Reed Switch Developments Company, Inc.  
P. O. Drawer 085297  
Racine, WI 53408  
Tel: 414-637-8646  
Fax: 414-637-6861

Smith Research and Technology, Inc.  
205 Sutton Lane, Dept. TR-95  
Colorado Springs, CO 80907  
Tel: 719-634-2252  
Fax: 719 634 2601

Smith Systems, Inc.  
6 Mill Creek Drive  
Box 667  
Rosedale, NC 28712  
Tel: 704-684-3490  
Fax: 704-877-3100

Standex Electronics  
4639 Chamberwell Road  
Dept. 301L  
Cincinnati, OH 45209  
Tel: 513 671 3779  
Fax: 513-671-3739

Torck, Inc.  
3000 Campus Drive  
Minneapolis, MN 55441  
Tel: 612-550-7300 & 800-544-7768  
Fax: 612-553-0709

Xulox Sensor Products  
6832 Gettysburg Pike  
Pt. Wayne, IN 46804  
Tel: 800-349-0741  
Fax: 210-492-0929

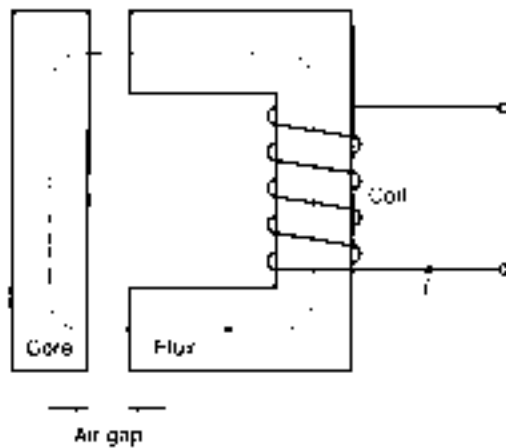
material response in magnetic sensors, the relative orientation and coupling between input measurand and magnetic properties are very important, and they are optimized at the design stages.

Many different types of magnetic sensors are available. These sensors can broadly be classified as primary or secondary. In primary sensors, also known as the magnetometers, the parameter to be measured is the external magnetic field. The primary sensors are commonly used in biological applications and geophysical and extraterrestrial measurements. In secondary sensors, the external parameter is made from other physical variables such as force and displacement. In this article, both the primary and secondary sensors will be discussed. These sensors include inductive, eddy current, transformative, magnetoresistive, Hall-effect, metal-oxide-semiconductor (MOS) magnetic field, and magneto-optical sensors; magnetotransistor

and magnetodiode sensors; magnetometers; superconductors; semiconductors; and magnetic thin films. They are offered by many manufacturers as listed in Table 1.

## INDUCTIVE SENSORS

Inductive sensors make use of the principles of magnetic circuits. They can be classified as passive sensors and self-generating sensors. The passive sensors require an external power source; hence, the action of the sensor is restricted to the modulation of the excitator signal in relation to an external stimuli. On the other hand, the self-generating types generate signals by utilizing the electrical generator principle based on Faraday's Law of Induction. That is, when there is a relative motion between a conductor and a magnetic field, a voltage is induced in the conductor. Or a varying magnetic field linking a



**Figure 1.** A basic inductive sensor consists of a magnetic circuit made up from a ferromagnetic core with a coil wound on it. The coil acts as a source of mmf, which drives the flux through the magnetic circuit and the air gap. The presence of the air gap causes a large increase in circuit reluctance and a corresponding decrease in the flux. Hence, a small variation in the air gap causes a measurable change in inductance.

stationary conductor produces voltage in the conductor, which can be expressed as

$$e = -d\Phi/dt \text{ (V)} \quad (11)$$

where  $\Phi$  is the magnetic flux.

In instrumentation applications, the magnetic field may be varying in time with some frequency, and the conductor may be moving at the same time. In many cases, the relative motion between field and conductor is supplied by changes in the measurand, usually by means of a mechanical motion.

In order to explain the operation of the basic principles of inductive sensors, a simple magnetic circuit is shown in Fig. 1. The magnetic circuit consists of a core, made from a ferromagnetic material, and a coil of a number of turns wound on it. The coil acts as a source of magnetomotive force (mmf), which drives the flux  $\Phi$  through the magnetic circuit. If we assume that the air gap is zero, the equation for the magnetic circuit may be expressed as

$$\text{mmf} = \text{Flux} \times \text{Reluctance} = \Phi \times \mathcal{R} \text{ (A-turns)} \quad (12)$$

such that the reluctance  $\mathcal{R}$  limits the flux in a magnetic circuit just as resistance limits the current in an electric circuit. By writing the magnetomotive force in terms of current, the magnetic flux may be expressed as

$$\Phi = ni/\mathcal{R} \text{ (Wb)} \quad (13)$$

In Fig. 1, the flux linking a single turn is expressed by Eq. (13). But the total flux linking by the entire  $n$  number of the turns of the coil is

$$\Psi = n\Phi = n^2 i/\mathcal{R} \text{ (Wb)} \quad (14)$$

Equation (14) leads to self-inductance  $L$  of the coil, which is described as the total flux per unit current for that particular coil.

That is,

$$L = \Psi/i = n^2/\mathcal{R} \text{ (H)} \quad (15)$$

This indicates that the self-inductance of an inductive element can be calculated by magnetic circuit properties. Expressing  $\mathcal{R}$  in terms of dimensions as

$$\mathcal{R} = l/\mu\mu_0 A \text{ (A-turns/Wb)} \quad (16)$$

where  $l$  is the total length of the flux path (meters),  $\mu$  is the relative permeability of the magnetic circuit material,  $\mu_0$  is the permeability of free space ( $= 4\pi \times 10^{-7}$  H/m), and  $A$  is the cross-sectional area of the flux path.

If the air gap is allowed to vary, the arrangement illustrated in Fig. 1 becomes a basic inductive sensor. In this case, the ferromagnetic core is separated in two parts by the air gap. The total reluctance of the circuit now is the addition of the reluctance of the core and the reluctance of the air gap. The relative permeability of air is close to unity, and the relative permeability of the ferromagnetic material can be on the order of a few thousand, indicating that the presence of the air gap causes a large increase in circuit reluctance and a corresponding decrease in the flux. Hence, a small variation in the air gap causes a measurable change in inductance. There are many different types of inductive sensors as will be discussed next.

#### Linear and Rotary Variable-Reluctance Sensors

The variable-reluctance transducers are based on change in the reluctance of a magnetic flux path. These types of devices find applications particularly in acceleration measurements. However, they can be constructed to be suitable for sensing displacements as well as velocities. They are constructed in many different forms, some of which will be described in this article.

**Single-Coil Linear Variable-Reluctance Sensor.** A typical single-coil variable-reluctance displacement sensor is illustrated in Fig. 2. The sensor consists of three elements: a ferromagnetic core in the shape of a semicircular ring, a variable air gap, and a ferromagnetic plate. The total reluctance of the magnetic circuit is the sum of the individual reluctances.

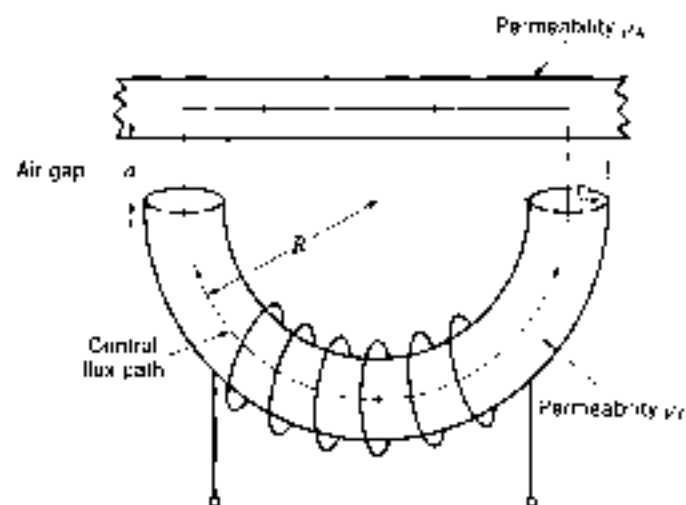
$$\mathcal{R}_T = \mathcal{R}_C + \mathcal{R}_G + \mathcal{R}_A \quad (17)$$

where  $\mathcal{R}_C$ ,  $\mathcal{R}_G$ , and  $\mathcal{R}_A$  are the reluctances of the core, air gap, and armature, respectively.

Each one of these reluctances can be determined by using the properties of materials involved as in Eq. (16). In this particular case,  $\mathcal{R}_T$  can be approximated as

$$\mathcal{R}_T = \frac{R}{\mu\mu_0\pi r^2} + \frac{2d}{\mu_0\pi r^2} + \frac{t}{\mu\mu_0\pi r^2} \quad (18)$$

In obtaining Eq. (18), the length of the flux path in the core is taken as  $\pi R$ , and the cross-sectional area is assumed to be uniform with a value of  $\pi r^2$ . The total length of the flux path in air is  $2d$ , and it is assumed that there is no fringing or bending of the flux through the air gap, such that the cross-sectional area of the flux path in air will be close to that of the cross-section of the core. The length of an average central flux path



**Figure 2.** A typical single-coil variable-reluctance displacement sensor. The sensor consists of three elements: a ferromagnetic core in the shape of a semi-circular ring, a variable air gap, and a ferromagnetic plate. The reluctance of the coil is dependent on the air gap. Air gap is the single variable, and the reluctance increases nonlinearly with the increasing gap.

in the armature is  $2\ell$ . The calculation of an appropriate cross section of the armature is difficult, but it may be approximated to  $2\ell t$ , where  $t$  is the thickness of the armature.

In Eq. (8), all the parameters are fixed except the only one independent variable, the air gap. Hence, it can be simplified as

$$\mathcal{R}_T = \mathcal{R}_0 + kd \quad (9)$$

where  $\mathcal{R}_0 = R^2(2\ell\mu_r^2\ell/\mu_0\pi^2) + \ell/\mu_r\ell$  and  $k = 2\ell/\mu_r\pi r^2$ .

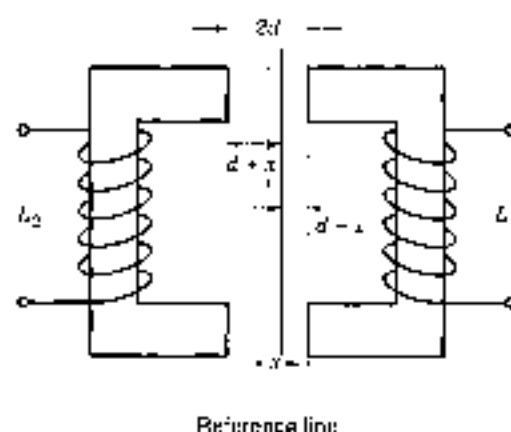
By using Eqs. (5) and (9), the inductance can be written as

$$L = \frac{\mu_0^2}{\mathcal{R}_0 + kd} = \frac{L_0}{1 + \alpha d} \quad (10)$$

where  $L_0$  represents the inductance at zero air gap and  $\alpha = k/\mathcal{R}_0$ .

The values of  $L_0$  and  $\alpha$  can be determined mathematically. They depend on the core geometry, permeability, and the like, as already explained. As it can be seen from Eq. (10), the relationship between  $L$  and  $d$  is nonlinear. Despite this nonlinearity, these types of single-coil sensors find applications in many areas, such as force measurements and telemetry. In force measurements, the resultant change in inductance can be made to be a measure of the magnitude of the applied force. The coil usually forms one of the components of an  $LC$  oscillator whose output frequency varies with the applied force. Hence, the coil modulates the frequency of the local oscillator.

**Variable-Differential Reluctance Sensor.** The problem of the nonlinearity may be overcome by modifying the single-coil system into variable-differential reluctance sensors (also known as push-pull sensors), as shown in Fig. 3. This sensor consists of an armature moving between two identical cores separated by a fixed distance of  $2d$ . Now, Eq. (10) can be written for both



**Figure 3.** A variable-differential reluctance sensor consists of an armature moving between two identical cores separated by a fixed distance. The armature moves in the air gap in response to the mechanical input. This movement alters the reluctance of coils 1 and 2 thus altering their inductive properties. This arrangement overcomes the problem of nonlinearity inherent in single-coil sensors.

coils as

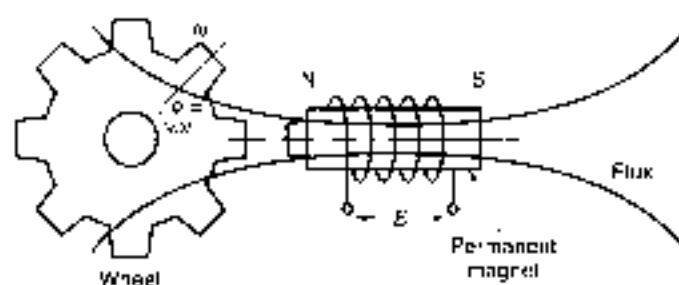
$$L_1 = \frac{L_0}{1 + \alpha(d+x)}, \quad L_2 = \frac{L_0}{1 + \alpha(d-x)} \quad (11)$$

Although the relationship between  $L_1$  and  $L_2$  is still nonlinear, the sensor can be incorporated into an ac bridge to give a linear output for small movements. The hysteresis error of these transducers is almost entirely limited to the mechanical components. These sensors respond to static and dynamic measurements. They have continuous resolution and high outputs, but they may give erratic performances in response to external magnetic fields. A typical sensor of this type has an input span of 1 mm, a coil inductance of 25 mH, and a coil resistance of 75  $\Omega$ . The resistance of the coil must be carefully considered when designing oscillator circuits. The maximum nonlinearity may be limited to 0.5%.

In typical commercially available variable-differential sensors, the iron core is located half way between the two E-shaped frames. The flux generated by primary coils depends on the reluctance of the magnetic path, the main reluctance being the air gap. Any motion of the core increases the air gap on one side and decreases it on the other side. Consequently, the reluctance changes in accordance with the principles explained previously, thus inducing more voltage on one of the coils than the other. Motion in the other direction reverses the action with a 180° phase shift occurring at null. The output voltage can be modified depending on the requirements in signal processing by means of rectification, demodulation, or filtering. In these instruments, full-scale motion may be extremely small, on the order of few thousandths of a centimeter.

In general, variable-reluctance transducers have small ranges and are used in specialized applications such as pressure transducers. Magnetic forces imposed on the armature are quite large, and this limits the application severely.

**Variable-Reluctance Tachogenerators.** Another example of the variable-reluctance sensor is shown in Fig. 4. These sensors are based on Faraday's Law of Electromagnetic Induction; therefore, they may also be referred as electromagnetic sensors.



**Figure 4.** A variable-reluctance tachogenerator is a sensor based on Faraday's Law of Electromagnetic Induction. It consists of a ferromagnetic toothed wheel attached to the rotating shaft and a coil wound onto a permanent magnet extended by a soft iron pole piece. The wheel rotates in close proximity to the pole piece, thus causing the flux linked by the coil to change. The change in flux causes an output in the coil similar to a square waveform whose frequency depends on the speed of the rotation of the wheel and the number of teeth.

Basically, the induced electromagnetic force (emf) in the sensor depends on the linear or angular velocity of the motion.

The variable-reluctance tachogenerator consists of a ferromagnetic toothed wheel attached to the rotating shaft, and a coil wound into a permanent magnet, extended by a soft iron pole piece. The wheel moves in close proximity to the pole piece, causing the flux linked by the coil to change, thus inducing an emf in the coil. The reluctance of the circuit depends on the width of the air gap between the rotating wheel and the pole piece. When the tooth is close to the pole piece, the reluctance is at a minimum, and it increases as the tooth moves away from the pole. If the wheel rotates with a velocity  $\omega$ , the flux may mathematically be expressed as

$$\psi(t) = \psi_m + \psi_1 \cos n\omega t \quad (12)$$

where  $\psi_m$  is the mean flux,  $\psi_1$  is the amplitude of the flux variation, and  $n$  is the number of teeth.

The induced emf is given by

$$E = - \frac{d\psi(t)}{dt} = - \frac{d\psi_1 \cos n\omega t}{dt} \times \frac{d\theta}{dt} \quad (13)$$

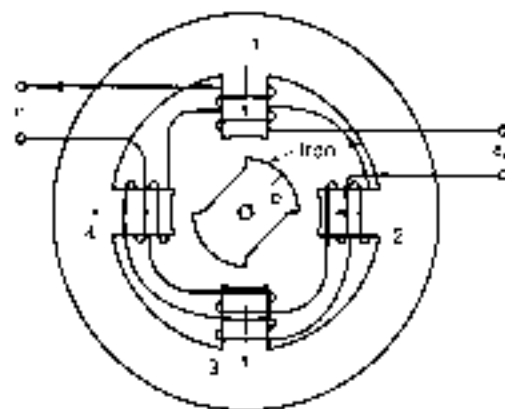
or

$$E = \psi_1 n \omega \sin n\omega t \quad (14)$$

Both the amplitude and the frequency of the generated voltage at the coil are proportional to the angular velocity of the wheel. In principle, the angular velocity  $\omega$  can be found from either the amplitude or the frequency of the signal. In practice, the amplitude measured may be influenced by loading effects and electrical interference. In signal processing, the frequency is the preferred option because it can be converted into digital signals easily.

The variable-reluctance tachogenerators are most suitable for measuring angular velocities. They are also used for volume flow rate measurements and the total volume flow determination of fluids.

**Microsyn.** Another commonly used example of variable-reluctance transducer is the microsyn, as illustrated in Fig. 5.



**Figure 5.** A microsyn is a variable-reluctance transducer that consists of a ferromagnetic rotor and a stator carrying four coils. The stator coils are connected such that at the null position, the voltages induced in coils 1 and 2 are balanced by voltages induced in coils 3 and 4. The motion of the rotor in one direction increases the reluctance of two opposite coils while decreasing the reluctance in others resulting in a net output voltage  $e_c$ . The movement in the opposite direction reverses this effect with a 180° phase shift.

In this arrangement, the coils are connected in such a manner that at the null position of the rotary element, the voltages induced in coils 1 and 3 are balanced by voltages induced in coils 2 and 4. The motion of the rotor in the clockwise direction increases the reluctance of coil 1 and 3 while decreasing the reluctance of coils 2 and 4, thus giving a net output voltage  $e_c$ . The movement in the counterclockwise direction causes a similar effect in coils 2 and 4 with a 180° phase shift. A direction-sensitive output can be obtained by using phase-sensitive demodulators.

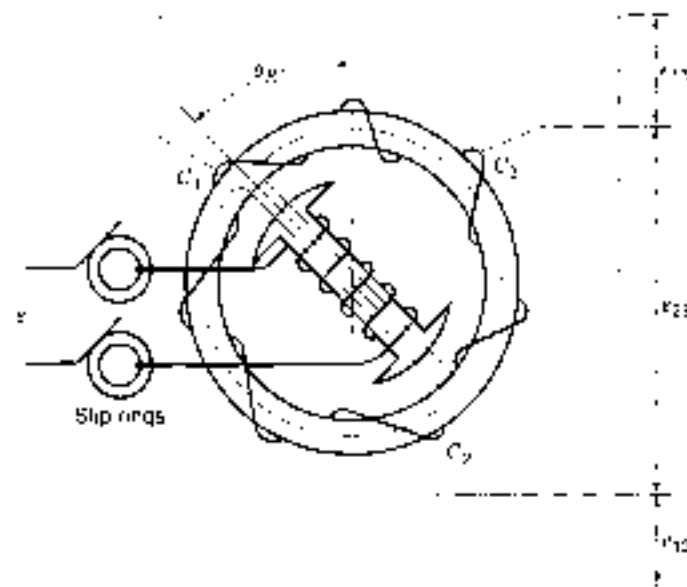
Microsyn transducers are extensively used in applications involving gyroscopes. By using microsins, very small motions can be detected giving an output signal as low as 0.01° of changes in angles. The sensitivity of the device can be made as high as 5 V per degree of rotation. The nonlinearity may vary from 0.5% to 1.0% full scale. The main advantage of these transducers is that the rotor does not have windings and slip rings. The magnetic reaction torque is also negligible.

### Synchros

The term *synchro* is associated with a family of electromechanical devices. They are primarily used in angle measurements and are commonly applied in control engineering as parts of servomechanisms, machine tools, antennas, and the like.

The construction of synchros is similar to that of wound-rotor induction motors, as shown in Fig. 6. The rotation of the motor changes the mutual inductance between the rotor coil and the three stator coils. The three voltage signals from these coils define the angular position of the rotor. Synchros are used in connection with a variety of devices, such as control transformers, Scott T transformers, resolvers, phase-sensitive demodulators, and analog-to-digital (A/D) converters.

In some cases, a control transformer is attached to the outputs of the stator coils such that the output of the control transformer produces a resultant mmf aligned in the same direction as that of the rotor of the synchro. In other words, the synchro rotor acts as a search coil in detecting the direction of the stator field of the control transformer. When the axis of this coil is



**Figure 6.** A synchro is similar to a wound rotor induction motor. The rotation of the rotor changes the mutual inductance between the rotor coil and the three stator coils. The voltages from these coils define the angular position of the rotor. They are primarily used in angle measurements and are commonly applied in control engineering as parts of servomechanisms, machine tools, antennas, and the like.

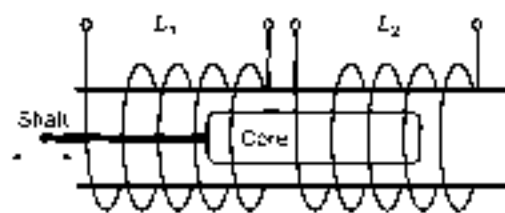
aligned with the field, the maximum voltage is supplied to the transformer.

In other cases, ac signals from the synchros are first applied to a Scott T transformer, which produces ac voltages with amplitudes proportional to the sine and cosine of the synchro shaft angle. It is also possible to use phase-sensitive demodulations to convert the output signals to make them suitable for digital signal processing.

#### Linear Variable Inductor

There is very little distinction between variable reluctance and variable-inductance transducers. Mathematically, the principles of linear variable transducers are very similar to the variable-reluctance type of transducers. The distinction is mainly in the sensing rather than principles of operations. A typical linear variable inductor consists of a movable iron core to provide the mechanical input and the two coils forming two legs of bridge network. A typical example of such a transducer is the variable coupling transducer.

The variable-coupling transducers consist of a former holding a center-tapped coil and a ferromagnetic plunger, as shown in Fig. 7. The plunger and the two coils have the same length  $l$ . As the plunger moves, the inductances of the coils change. The two inductances are usually placed to form two arms of a bridge circuit with two equal balancing resistors. The bridge is then excited with ac of 5 V to 25 V with a frequency of 50 Hz to 5 kHz. At the selected excitation frequency, the total transducer impedance at null conditions is set in the 100  $\Omega$  to 1000  $\Omega$  range. The resistors are set to have about the same value as transducer impedances. The load for the bridge output must be at least ten times the resistance  $R$  value. When the plunger is in the reference position, each coil will have equal inductances of value  $L$ . As the plunger moves by  $\delta l$ , changes in inductances



**Figure 7.** A typical linear variable inductor consists of a movable iron core inside a former holding a center-tapped coil. The core and both coils have the same length  $l$ . When the core is in the reference position, each coil will have equal inductances of value  $L$ . As the core moves by  $\delta l$ , changes in inductances  $+\delta L$  and  $-\delta L$  create voltage outputs from the coils.

$+\delta L$  and  $-\delta L$  create a voltage output from the bridge. By constructing the bridge carefully, the output voltage may be made as a linear function displacement of the moving plunger within a rated range.

In some transducers, in order to reduce power losses resulting from the heating of resistors, center-tapped transformers may be used as a part of the bridge network. In this case, the circuit becomes more inductive, and extra care must be taken to avoid the mutual coupling between the transformer and the transducer.

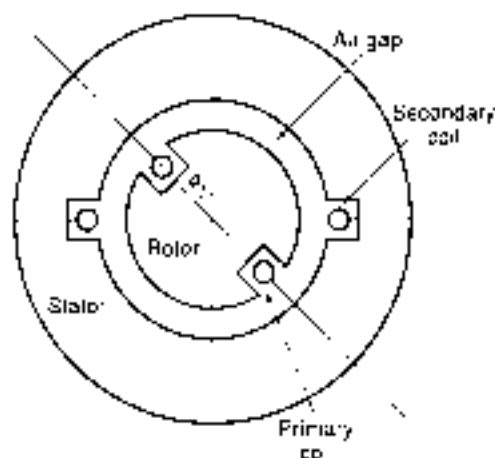
It is particularly easy to construct transducers of this type, by simply winding a center-tapped coil on a suitable former. The variable-inductance transducers are commercially available in strokes from about 2 mm to 500 cm. The sensitivity ranges between 1% full scale to 0.02% in long-stroke special constructions. These devices are also known as linear displacement transducers or LDTs, and they are available in various shapes and sizes.

Apart from linear variable inductors, rotary types are also available. Their cores are specially shaped for rotational applications. Their nonlinearity can vary between 0.5% and 1% full scale over a range of 90° rotation. Their sensitivity can be up to 100 mV per degree of rotation.

#### Induction Potentiometer

A version of rotary-type linear inductors is the induction potentiometer, as shown in Fig. 8. Two concentrated windings are wound on stator and rotor. The rotor winding is excited with an ac, thus inducing voltage in the stator windings. The amplitude of the output voltage is dependent on the mutual inductance between the two coils, where mutual inductance itself is dependent on the angle of rotation. For concentrated coil-type induction potentiometers, the variation of the amplitude is sinusoidal, but linearity is restricted in the region of the null position. A linear distribution over an angle of 180° may be obtained by carefully designed distributed coils.

Standard commercial induction pots operate in a 50 Hz to 400 Hz frequency range. They are small in size from 1 cm to 6 cm, and their sensitivity can be in the order of 1 V/1° of rotation. Although the ranges of induction pots are limited to less than 60° of rotation, it is possible to measure displacements in angles from 0° to full rotation by suitable arrangements of a number of induction pots. As in the case of most inductive sensors, the output of the induction pots may need phase-sensitive demodulators and suitable filters. In many inductive

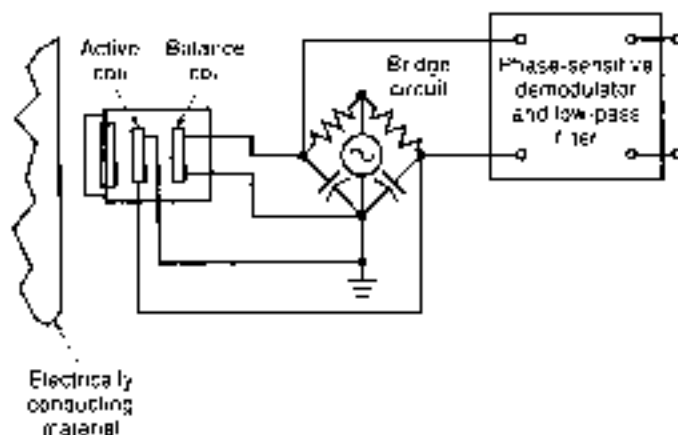


**Figure 8.** An induction potentiometer is a linear variable inductor with two concentrated windings wound on the stator and rotor. The rotor winding is excited with ac, inducing voltage in the stator windings. The amplitude of the output voltage is dependent on the relative positions of the coils—determined by the angle of rotation. For concentrated coils, the variation of the amplitude is sinusoidal but linearity is restricted in the region of the null position. Different types of induction potentiometers are available with distributed coils, which give linear voltage over an angle of 180° of rotation.

poles, additional dummy coils are used to improve linearity and accuracy.

### EDDY CURRENT SENSORS

Inductive transducers based on eddy currents are mainly probe types containing two coils, as shown in Fig. 9. One of the coils, known as the active coil, is influenced by the presence of the conducting target. The second coil, known as the balance coil, serves to complete the bridge circuit and provides temperature



**Figure 9.** Eddy current transducers are inductive transducers using probes. The probes contain one active and one balance coil. The active coil responds to the presence of a conducting target, whereas the balance coil completes a bridge circuit and provides temperature compensation. When the probe is brought close to the target, the flux from the probe links with the target, producing eddy currents within the target, which alter the inductance of the active coil. This change in inductance is detected by a bridge circuit.

compensation. The magnetic flux from the active coil passes into the conducting target by means of a probe. When the probe is brought close to the target, the flux from the probe links with the target, producing eddy currents within the target.

The eddy current density is greatest at the target surface and becomes negligibly small about three skin depths below the surface. The skin depth depends on the type of material used and the excitation frequency. Even though thinner targets can be used, a minimum of three skin depths may often be necessary to minimize the temperature effects. As the target comes closer to the probe, the eddy currents become stronger, causing the impedance of the active coil to change and altering the balance of the bridge in relation to the target position. This unbalance voltage of the bridge may be demodulated, filtered, and linearized to produce a dc output proportional to target displacement. The bridge oscillation may be as high as 1 MHz. High frequencies allow the use of thin targets and provide a good system frequency response.

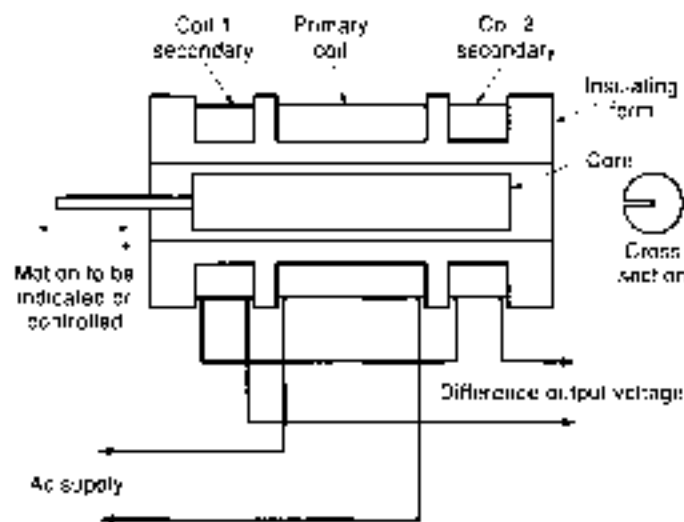
Probes are commercially available with full-scale ranges from 0.25 mm to 30 mm with a nonlinearity of 0.5% and a maximum resolution of 0.0001 mm. Targets are usually supplied by the clients, involving noncontact measurements of machine parts. For nonconductive targets, conductive materials of sufficient thickness must be attached onto the surface by means of commercially available adhesives. Because the target material, shape, and the like influence the output, it is necessary to calibrate the system statistically for a specific target. The recommended measuring range of a given probe begins at a standoff distance equal to about 20% of the stated range of the probe. In some cases, a standoff distance of 10% of the stated range for which the system is calibrated is recommended as standard. A distance greater than 10% of the measuring range can be used as long as the calibrated measuring range is reduced by the same amount.

Flat targets must be the same diameter as the probe or larger. If the target diameter is smaller than the probe diameter, the output drops considerably, thus becoming unreliable. Curved-surface targets may behave similar to flat surfaces if the diameter exceeds about three or four diameter of the probe. In this case, the target essentially becomes an infinite plane. This also allows some cross-axis movement without affecting the system output. Target diameter comparable to the sensor could result in detrimental effects from cross-axis movements.

For curved or irregularly shaped targets, the system needs to be calibrated using an exact target that may be seen in the operation. This tends to eliminate any errors caused by the curved surfaces during the applications. However, special multiprobe systems are available for orbital motions of rotating shafts. If the curved (shaft) target is about ten times greater than the sensor diameter, it acts as an infinite plane and does not need special calibrations. Special care must be exercised to deal with electrical runout resulting from factors such as inhomogeneities in hardness, particularly valid for ferritic targets. However, nonferrous targets are free from electrical runout concerns.

### TRANSFORMATIVE SENSORS

Transformative sensors make use of the principles of transformer action, that is, magnetic flux created by one coil links



**Figure 10.** A linear variable-differential transformer is a passive inductive transducer consisting of a single primary winding positioned between two identical secondary windings wound on a tubular ferromagnetic former. As the core inside the former moves, the magnetic paths between primary and secondaries alter, thus giving secondary outputs proportional to the movement. The two secondaries are made as similar as possible by having equal sizes, shapes, and number of turns.

with the other coil to induce voltages. There are many different types, such as linear variable transformers, rotary variable differential transformers, and flux-gate magnetometers.

#### Linear Variable-Differential Transformer

The linear variable-differential transformer (LVDT) is a passive inductive transducer that has found many applications. It consists of a single primary winding positioned between two identical secondary windings wound on a tubular ferromagnetic former, as shown in Fig. 10. The primary winding is energized by a high-frequency 60 Hz to 20 kHz ac voltage. The two secondaries are made identical by having an equal number of turns. They are connected in series opposition so that the induced output voltages oppose each other.

In many applications, the outputs are connected in opposing form, as shown in Fig. 11(a). The output voltages of individual secondaries  $v_1$  and  $v_2$  at null position are illustrated in Fig. 11(b). However, in opposing connection, any displacement in the core position  $x$  from the null point causes amplitude of the voltage output  $v_o$  and the phase difference  $\alpha$  to change. The output waveform  $v_o$  in relation to core position is shown in Fig. 11(c). When the core is positioned in the middle, there is an equal coupling between primary and secondaries, thus giving a null point or reference point of the sensor. As long as the core remains near the center of the coil arrangement, output is very linear. The linear ranges of commercial differential transformers are clearly specified, and the devices are seldom used outside this linear range.

The ferromagnetic core or plunger moves freely inside the former; thus altering the mutual inductance between the primary and secondaries. With the core in the center, or at the reference position, the induced emfs in the secondaries are equal, and because they oppose each other, the output voltage is zero. When the core moves, say to the left, from the center, more mag-

netic flux links with the left-hand coil than with the right-hand coil. The voltage induced in the left-hand coil is therefore larger than the induced emf on the right-hand coil. The magnitude of the output voltage is then larger than at the null position and is equal to the difference between the two secondary voltages. The net output voltage is in phase with the voltage of the left-hand coil. The output of the device is then an indication of displacement of the core. Similarly, movement in the opposite direction to the right from the center reverses this effect, and the output voltage is now in phase with the emf of the right-hand coil.

For mathematical analysis of the operation of LVDTs Fig. 11(a) may be used. The voltages induced in the secondary coils are dependent on the mutual inductance between the primary and individual secondary coils. Assuming that there is no cross coupling between the secondaries, the induced voltages may be written as

$$v_1 = M_1 \dot{i}_p \quad \text{and} \quad v_2 = M_2 \dot{i}_p \quad (15)$$

where  $M_1$  and  $M_2$  are the mutual inductances between primary and secondary coils for a fixed core position,  $s$  is the Laplace operator, and  $i_p$  is the primary current.

In the case of opposing connection, no load output voltage  $v_o$  without any secondary current may be written as

$$v_o = v_1 - v_2 = (M_1 - M_2) \dot{i}_p \quad (16)$$

writing

$$v_o = i_p (R + sL_p) \quad (17)$$

Substituting  $i_p$  in Eq. (16) gives the transfer function of the transducer as

$$\frac{v_o}{i_p} = \frac{(M_1 - M_2)s}{R + sL_p} \quad (18)$$

However, if there is a current resulting from output signal processing, then describing equations may be modified as

$$v_o = R_o i_o \quad (19)$$

where  $i_o = (M_1 - M_2) \dot{i}_p / (R + R_o + sL_p)$  and

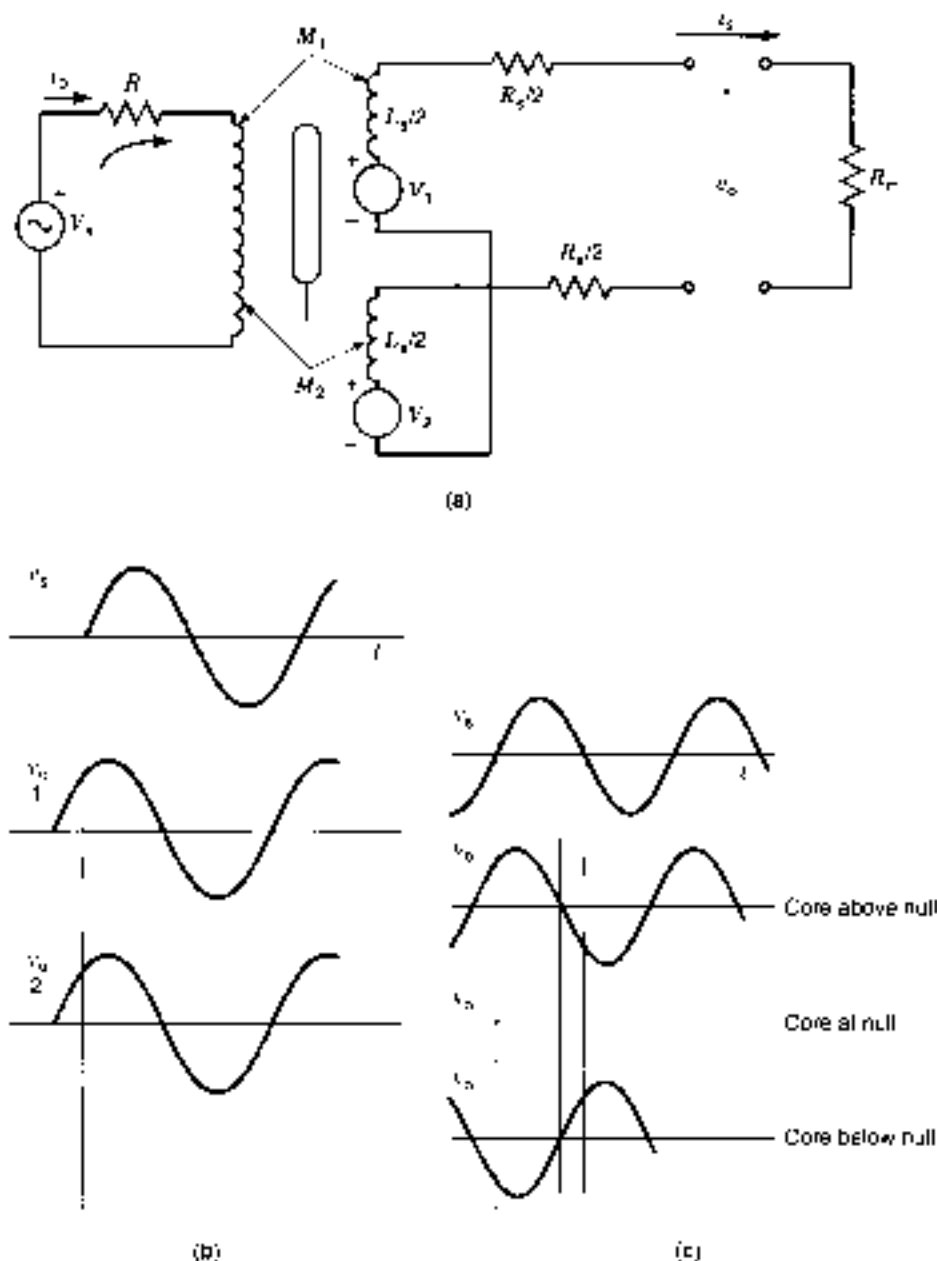
$$v_o = i_p (R + sL_p) + (M_1 - M_2) \dot{i}_o \quad (20)$$

Eliminating  $i_p$  and  $i_o$  from Eqs. (19) and (20) results in a transfer function

$$\frac{v_o}{i_o} = \frac{R_o(M_1 - M_2)s}{(M_1 - M_2)^2 + L_p L_o s^2 + (L_p R + R_o) + RL_o k + R_o - R} \quad (21)$$

This is a second-order system, which indicates that with the effect of the numerator the frequency of the system changes from  $+90^\circ$  at low frequencies to  $-90^\circ$  at high frequencies. In practical applications, the supply frequency is selected such that at null position of the core the phase angle of the system is  $0^\circ$ .

The amplitudes of the output voltages of secondary coils are dependent on the position of the core. These outputs may directly be processed from each individual secondary coils for slow



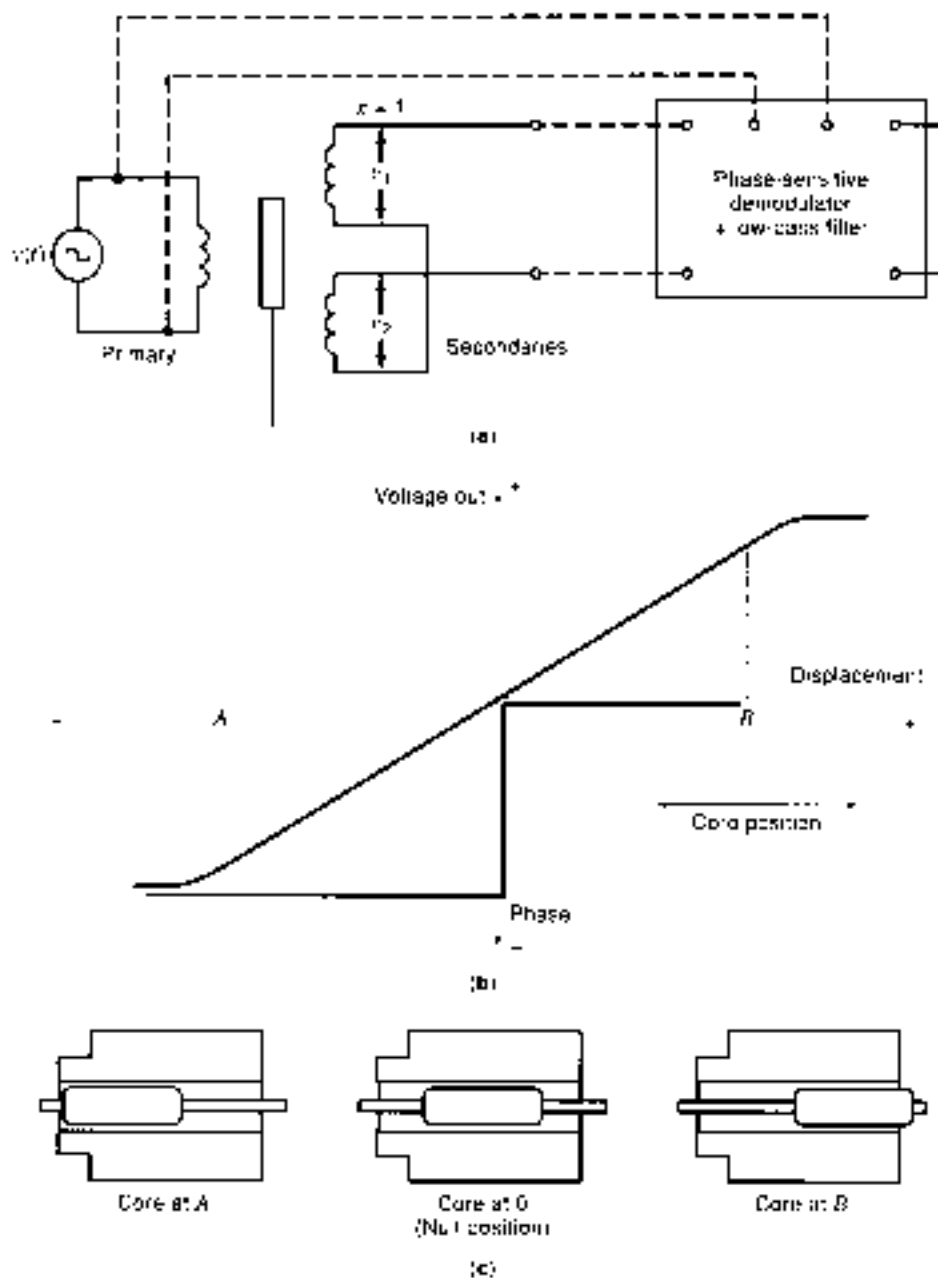
**Figure 11.** The voltages induced in the secondaries of a linear variable differential transformer (a) may be processed in a number of ways. The output voltages of individual secondaries  $v_1$  and  $v_2$  at null position are illustrated in (b). In this case, the voltages of individual coils are equal and in phase with each other. Sometimes the outputs are connected opposing each other, and the output waveform  $v_o$  becomes a function of core position  $x$  and phase angle  $\theta$  as in (c). Note the phase shift of  $180^\circ$  as the core position clamps above and below the null position.

movements of the core, if the direction of the movement of the core does not bear any importance. However, for fast movements of the core, the signals may be converted to dc, and the direction of the movement from the null position may be detected. There are many options to do this, however, a *phase-sensitive demodulator* and filter are commonly used as shown in Fig. 12(a). A typical output of the phase-sensitive demodulator is illustrated in Fig. 12(b), for core positions as in Fig. 12(c), in relation to output voltage  $v_o$ , displacement  $x$ , and phase angle  $\theta$ .

The phase-sensitive demodulators are extensively used in differential-type inductive sensors. They basically convert the ac outputs to dc values and also indicate the direction of

movement of the core from the null position. A typical phase-sensitive demodulation circuit may be constructed, based on diodes shown in Fig. 13(a). This arrangement is useful for very slow displacements, usually less than 1 or 2 Hz. In Fig. 13(a), bridge 1 acts as a rectification circuit for secondary 1, and bridge 2 acts as a rectifier for secondary 2. The net output voltage is the difference between the outputs of two bridges as in Fig. 13(b). The position of the core can be worked out from the amplitude of the dc output and the direction of the movement of the core can be determined from the polarity of the dc voltage. For rapid movements of the core, the output of the diode bridges need to be filtered, and this passes only the frequencies of the movement of the core and filters all the other frequencies





**Figure 12.** Phase-sensitive demodulator and filter (a) are commonly used to obtain displacement-proportional signals from LVDTs and other differential-type inductive sensors. They convert the ac outputs from the sensors into dc values and also indicate the direction of movement of the core from the null position. A typical output of the phase-sensitive demodulator is shown in (b). The relationship between output voltage  $v_o$  and phase angle  $\alpha$  is also shown against core position  $x$  as sketched in (c).

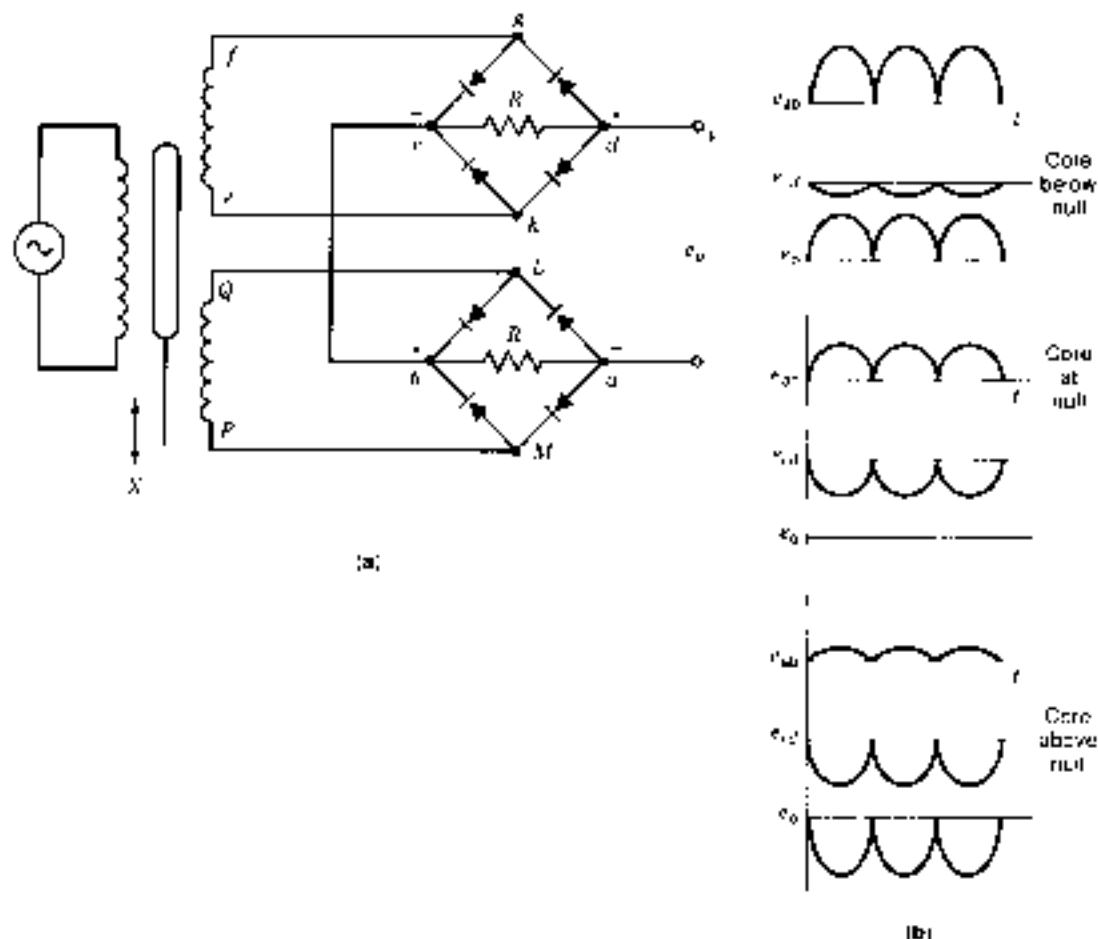
produced by the modulation process. For this purpose, a suitably designed simple RC filter may be sufficient.

In the marketplace, there are phase-sensitive demodulator chips available, such as AD598 offered by Analog Devices, Inc. These chips are highly versatile and flexible to suit particular application requirements. They offer many advantages over conventional phase-sensitive demodulation devices: for example, frequency of excitation may be adjusted to any value between 20 Hz and 20 kHz by connecting an external capacitor between two pins. The amplitude of the excitation voltage can be set up to 24 V. The internal filters may be set to required values by external capacitors. Connections to analog to digital converters are made easy by converting the bipolar output to a unipolar scale.

The frequency response of LVDTs is primarily limited by the inertia characteristics of the device. In general, the frequency of the applied voltage should be ten times the desired frequency

response. Commercial LVDTs are available in a broad range of sizes, and they are widely used for displacement measurements in a variety of applications. The displacement sensors are available to cover ranges from  $\pm 0.25$  mm to 17.5 cm. They are sensitive enough to be used to respond to displacements well below 0.0005 mm. They can have operational temperature range from  $-265$  to  $600$  C. They are also available in radiation-resistant designs for operation in nuclear reactors. For a typical sensor of range  $\pm 25$  mm, the recommended supply voltage is 4 V to 6 V, with a nominal frequency of 5 kHz and a maximum nonlinearity of 1% full scale. Several commercial models, which can produce a voltage output of 300 mV for 1 mm displacement of the core, are available.

One important advantage of the LVDTs is that there is no physical contact between the core and the coil form, hence there is no friction or wear. Nevertheless, there are radial and longitudinal magnetic forces on the core at all times. These magnetic



**Figure 15.** A typical phase-sensitive demodulation circuit based on diode bridges as in (a). The bridge 1 acts as a rectification circuit for secondary 1, and bridge 2 acts as a rectifier for secondary 2 where the net output voltage is the difference between the two bridges as in (b). The position of the core can be worked out from the amplitude of the dc output, and the direction of the movement of the core can be determined from the polarity of the voltage. For rapid movements of the core, the output of the diode bridges need to be filtered. For filters, a suitably designed simple AC filter may be sufficient.

forces may be regarded as magnetic springs that try to displace the core from its null position. This may be a critical factor in some applications.

One problem with LVDTs is that it may not be easy to make the two halves of the secondary identical; their inductance, resistance, and capacitance may be different, causing a large unwanted quadrature output in the balance position. Precision coil-winding equipment may be required to reduce this problem to an acceptable value.

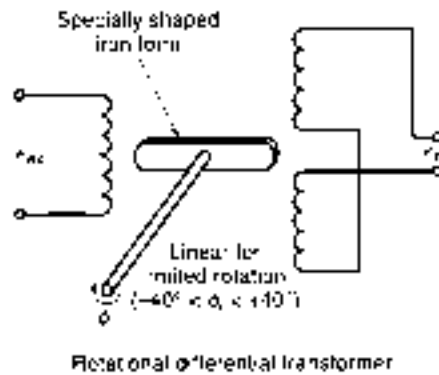
Another problem is associated with null position adjustments. The harmonics in the supply voltage and stray capacitances result in small null voltages. The null voltage may be reduced by proper grounding, which reduces the capacitive effects and the center-tapped voltage source arrangements. In center-tapped supplies, a potentiometer may be used to obtain a minimum null reading.

The LVDTs find a variety of applications, which include jet engines controls that are in close proximity to exhaust gases, and controls that measure null positions in the thickness of materials in hot-slab steel mills. After some mechanical conversions, LVDTs may also make force and pressure measurements.

### Rotary Variable-Differential Transformer

A variation from the linear variable-differential transformer is the rotary core differential transformer, as shown in Fig. 14. Here the primary winding is wound on the center leg of an E core; the secondary windings are wound on the outer legs of the E core. The armature is rotated by an externally applied force about a pivot point above the center leg of the core. When the armature is displaced from its reference or balance position, the reluctance of the magnetic circuit through one secondary coil is decreased; simultaneously the reluctance through the other coil is increased. The induced emfs in the secondary windings, which are equal in the reference position of the armature, are now different in magnitude and phase as a result of the applied displacement. The induced emfs in the secondary coils are made to oppose each other, and the transformer operates in the same manner as LVDTs. The rotating variable transformers may be sensitive to vibrations. If a dc output is required, a demodulator network may be used, as in the case of LVDTs.

In most rotary linear-variable differential transformers, the rotor mass is very small, usually less than 5 g. The nonlinearity



**Figure 14.** A rotary core differential transformer has an armature rotated by an externally applied force about a pivot point above the center leg of the core. When the armature is displaced from its reference or balance position, the reluctance of the magnetic circuit through one secondary coil is decreased; simultaneously the reluctance through the other coil is increased. The induced emfs in the secondary windings are different in magnitude and phase as a result of the applied displacement.

in the output ranges between  $\pm 1\%$  and  $\pm 3\%$ , depending on the angle of rotation. The motion in the radial direction produces a small output signal that can affect the overall sensitivity. But this transverse sensitivity is usually kept less than 1% of the longitudinal sensitivity.

#### MOS MAGNETIC FIELD SENSORS

The technology of integrated magnetic field sensors, also called semiconductor magnetic microsensors, is well developed. This technology uses either high-permeability (e.g., ferromagnetic) or low-permeability (e.g., paramagnetic) materials. The integrated circuit magnetic techniques support many sensors, such as magnetometers, optoelectronics, Hall-effect sensors, magnetic semiconductor sensors, and superconductive sensors.

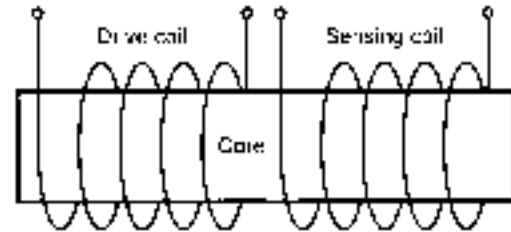
At the present, silicon offers an advantage of inexpensive batch fabrication for magnetic sensors. Most of the integrated circuit magnetic sensors are manufactured by following the design rules of standard chip manufacturing. For example, MOS and complementary metal oxide semiconductor (CMOS) technologies are used to manufacture highly sensitive Hall-effect sensors, magnetotransistors, and other semiconductor sensors. Some of the magnetic semiconductor sensors are discussed in detail in the following sections.

#### MAGNETOMETERS

Magnetometers are devices that are produced to sense external magnetic fields mainly based on Faraday's Law of Induction. There are many different types of magnetometers including search coil magnetometers, SQUIDs, flux-gates, nuclear, and optical magnetometers. The two most commonly used types are the flux-gate and the search coil magnetometers.

##### Flux-Gate Magnetometers

Flux-gate magnetometers are made from two coils wound on a ferromagnetic material, as illustrated in Fig. 15. One of the



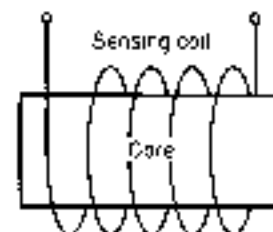
**Figure 15.** A flux-gate magnetometer consists of two coils wound on a ferromagnetic core. The driving coil is excited with a sinusoidal current, which drives the core into a saturation state. When saturated, the reluctance of the core to external magnetic field increases, thus repelling the external flux and hence reducing the effect of the external field on the second coil. The harmonics of the induced voltage in the sensing coil is an indication of the magnitude of the external magnetic field.

coil is excited with a sinusoidal current, which drives the core into a saturation state. At this stage, the reluctance of the core to external magnetic field increases, thus repelling the external flux and so reducing the effect of the external field on the second coil. As the core becomes unsaturated, the effect of the external field increases. The increase and reduction of the effect of the external field on the second coil is sensed as the harmonics of the induced voltage. These harmonics are then directly related to the strength and variations in the external field. The sensitivity of these devices is dependent on the magnetic properties and the shape of the saturation curve of the core material. The measurement range can vary from 100 pT to 10 nT operating in the 0 kHz to 10 kHz frequency range.

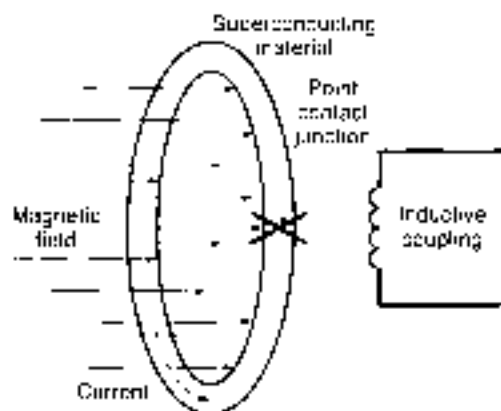
##### Search Coil Magnetometers

Search coil magnetometers operate on the principle of Faraday's Law of Induction. A typical search coil magnetometer is shown in Fig. 16. The flux through the coil changes if the magnetic field varies in time or the coil is moved through the field. The sensitivity depends on the properties of the core material, dimensions of the coil, number of turns, and rate of change of flux through the coil.

Search coil magnetometers are manufactured from 4 cm in dimensions to 100 cm. They can sense weak fields as low as 100 pT within the frequency range of 1 Hz to 1 MHz. The upper limit is dependent on the relative magnitudes of resistance and inductance of the coil. During the signal processing, they can be used as part of a bridge or resonant circuits.



**Figure 16.** A search coil magnetometer uses Faraday's Law of Induction. The changing magnetic flux induces an emf in the coil. The sensitivity of the magnetometer depends on the properties of the core material, dimensions of the coil, number of turns, and rate of change of flux through the coil. They can sense weak fields as low as 100 pT within the frequency range of 1 Hz to 1 MHz.



**Figure 17.** A SQUID consists of a superconducting ring. When subjected to an external field, the current is induced in the ring, which flows forever. The current in the ring is measured by using the Josephson effect and by creating a weak link in the ring. This weak link makes the superconducting current oscillate as a function of the external magnetic field intensity.

### SQUID Magnetometers

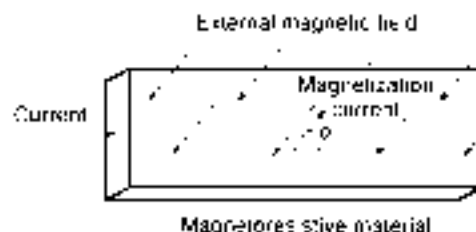
Superconducting quantum interference device, SQUID, sensors are used in many diverse applications from medicine, geophysics, and nuclear magnetic resonance to nondestructive testing of solid materials. The SQUID offers high sensitivity for the detection of weak magnetic fields and field gradients. They are made from conventional superconductors such as niobium operating at liquid helium temperatures. They are manufactured by using integrated circuit technology, as illustrated in Fig. 17.

The principle of operation is the Josephson effect. If a magnetic flux links to a ring-shaped superconducting material, a current is induced in the ring. This current flows forever due to the lack of any resistance. The intensity of current is proportional to the intensity of the field. The current in the ring is measured by using the Josephson effect wherein a weak link in the superconducting ring causes the superconducting current to oscillate as a function of magnetic field intensity. These oscillations can be sensed by many techniques such as coupling to a radio-frequency circuit and other resonance techniques.

The SQUIDs are extremely sensitive devices, with sensitivities ranging from 10 fT to 10 nT. The SQUID sensors can be arranged to measure magnetic fields in three dimensions in  $x$ ,  $y$ , and  $z$  directions of Cartesian coordinates, as in the case of high-sensitivity gradiometers. One disadvantage of SQUIDs is that they need supercooling at very low temperatures. Much research is concentrated on materials exhibiting superconductivity properties at high temperatures. SQUIDs are used in gradiometers, voltmeters and amplifiers, displacement sensors, geophysics, gravity wave detection, and nondestructive testing, to name a few.

### MAGNETORESISTIVE SENSORS

In magnetoresistive sensors, the magnetic field causes a change in resistance of some materials such as permalloys. In these materials, current passing through the material magnetizes the material in a particular magnetic orientation. An external field perpendicular to the current, as illustrated in Fig. 18,



**Figure 18.** A magnetoresistive sensor's resistance changes in response to an external magnetic field. Magnetic domain orientation of the material is a function of external field and current flowing through it. The resistance is highest when the magnetization is parallel to the current and lowest when it is perpendicular to the current. These sensors are manufactured as thin films, and they have good linearity.

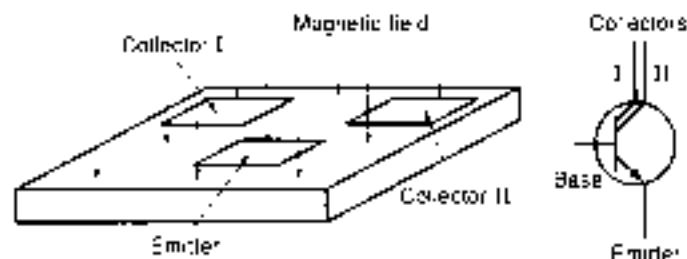
causes the magnetic orientation to be disturbed. The resistance is highest when the magnetization is parallel to the current and lowest when it is perpendicular to the current. Hence, depending on the intensity of the external magnetic field, the resistance of the permalloy changes in proportion.

Magnetoresistive sensors are manufactured as thin films and usually integrated to be part of an appropriate bridge circuit. They have good linearity and low temperature coefficients. These devices have a sensitivity ranging from 1  $\mu$ T to 50 mT. By improved electronics with suitable feedback circuits, the sensitivity can be as low as 100 pT. They can operate from dc to several gigahertz.

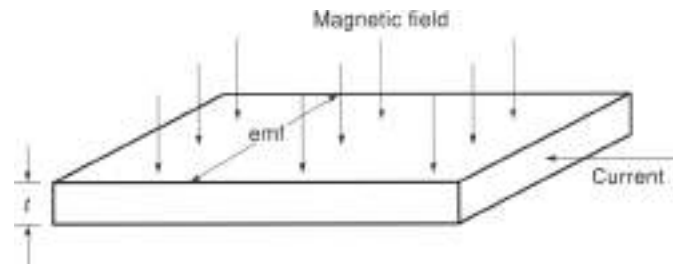
### Magnetotransistor and Magnetodiode Sensors

Magnetotransistor and magnetodiode sensors are integrated silicon devices. They contain  $n$ -doped and  $p$ -doped regions forming  $pn$ ,  $npn$ , or  $pnp$  junctions. In the case of magnetotransistors, there are two collectors, as shown in Fig. 19. Depending on the direction, an external magnetic field deflects electron flow between emitter and collector in favor of one of the collectors. The two collector voltages are sensed and related to the applied magnetic field. These devices are more sensitive than Hall-effect sensors.

In the case of magnetodiodes,  $p$  and  $n$  regions are separated by an area of undoped silicon containing the sensor. An external magnetic field perpendicular to the flow of charges deflects the holes and electrons in the opposite directions, resulting effectively in a change in the resistance of the undoped silicon layer.



**Figure 19.** A magnetotransistor is an integrated silicon device that contains  $npn$  or  $pnp$  junctions. The electron flow between emitter and collector is influenced by the external magnetic field in favor of one of the collectors. The two collector voltages are sensed and related to the applied magnetic field.



**Figure 20.** A Hall effect sensor makes use of Lorentz force. The response of electrons to Lorentz force creates the Hall voltage, which is perpendicular to both the external magnetic field and the direction of current flow. If the current is dc, the voltage has the same frequency as magnetic flux.

### HALL-EFFECT SENSORS

In Hall-effect sensors, the voltage difference across a thin conductor carrying current depends on the intensity of the magnetic field applied perpendicular to the direction of current flow, as shown in Fig. 20. An electron moving through a magnetic field experiences Lorentz force perpendicular to the direction of motion and to the direction of the field. The response of electrons to Lorentz force creates a voltage known as the Hall voltage. If a current  $I$  flows through the sensor, the Hall voltage can mathematically be found by

$$V = R_H I B_z t \quad (22)$$

where  $R_H$  is the Hall coefficient (cubic meters per degree Celsius),  $B$  is the flux density (tesla), and  $t$  is the thickness of the sensor (meters).

Therefore, for a specified current and temperature, the voltage is proportional to  $B$ . If the current is dc, the voltage has the same frequency as magnetic flux.

Hall-effect sensors can be made by using metals or silicon, but they are generally made from semiconductors with high electron mobility such as indium antimonide. They are usually manufactured in the form of probes with a sensitivity down to 100  $\mu$ T. Silicon Hall-effect sensors can measure constant or varying magnetic flux having an operational frequency from dc to 1 MHz, within the range of 1  $\mu$ T to 100 mT. They have good temperature characteristics from 200°C to near absolute zero.

### MAGNETO-OPTICAL SENSORS

In recent years, highly sensitive magneto-optical sensors have been developed. These sensors are based on fiber optics, polarization of light, Moiré effect, and Zeeman effect, among others. This type of sensors leads to highly sensitive devices and is used in applications requiring high resolution such as human brain function mapping and magnetic anomaly detection. Here, because of the availability of space, only the polarization effect will be discussed briefly. Interested readers can find further information in the Refs. 1-6.

In the polarization effect, a plane of polarized light in a strong magnetic field rotates its plane of vibration. The per-unit angular rotation is related to the per-unit magnetic field in a given length of material by the Verdet constant. For ex-

ample, terbium gallium demonstrates a Verdet constant of 50  $\text{min}/\mu\text{T}\cdot\text{cm}$ , whereas bismuth-substituted iron garnet can have up to 0.04  $\text{min}/\text{T}\cdot\text{cm}$ . This polarization effect was first noticed by Faraday in 1845; hence, it is generally known as the Faraday effect. Recently, this principle was applied to semiconductors and crystals, which have different physical properties such as interband effects, intraband free carrier effects, and absorption of magnetism by impurities.

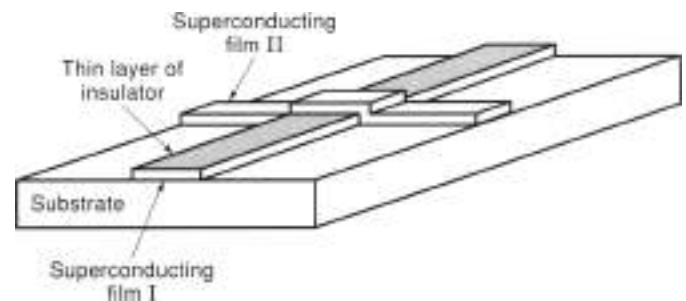
### MAGNETIC THIN FILMS

Magnetic thin films are an important part of superconducting instrumentation, sensors and electronics in which active devices are made from deposited films. The thin films are usually made from amorphous alloys, amorphous gallium, and the like. As an example of this use of thin-film technology a thin-film Josephson junction is given in Fig. 21. The deposition of thin films can be done by thermal evaporation, electroplating, sputter deposition, or chemical methods. The choice of technology depends on the characteristics of the sensors. For example, thin-film superconductors require low-temperature operations, whereas common semiconductors operate at room temperature.

The magnetic thin films find extensive applications in memory devices where high density and good sensitivities are required. In such applications, the magnetic properties of the coating are determined by the magnetic properties of the particles that can be controlled before coating. The choice of available materials for this purpose is extremely large. Thin-film technology is also developed in magneto-optics applications where erasable optical media for high-density magnetic storage is possible. The miniature magnetoresistive sensors for magnetic recording and pick-up heads, the Hall-effect sensors, and other magnetic semiconductors make use of thin-film technology extensively.

### AMORPHOUS MAGNETIC MATERIALS

The amorphous magnetic materials can be classified as amorphous alloys (Fe, Co, Ni), amorphous rare earths, and amorphous superconductors. Amorphous alloys have good soft magnetic materials and are extensively used in magnetic heads. They have high saturation magnetization, high permeability,



**Figure 21.** The magnetic thin films are made from deposited thin films from amorphous alloys such as gallium. This figure illustrates thin film technology to form a Josephson. The deposition of thin films can be done by thermal evaporation, electroplating, sputter deposition, or chemical methods.

and fast high-frequency deposition. Amorphous alloys are produced in the form of 20  $\mu\text{m}$  to 50  $\mu\text{m}$  thick ribbons using rapid solidification methods. Nevertheless, thermal stability of these alloys is the major drawback preventing the wider applications. Many of the amorphous metallic films are based on rare earth 3d transition metal alloys, such as  $\text{Gd}_2\text{Co}_2$ . They are used in bubble domain devices.

On the other hand, the amorphous superconductors are a class of superconducting materials such as niobium and gallium. They are manufactured in the form of powders, ribbons, or flakes by using evaporation methods, ion mixing or ion implantation, or liquid quenching. The vapor deposition technique is used in semiconductor-type sensors. These materials are used in high field magnets, memory devices, and other computer applications.

#### SHIELDING AND SENSITIVITY TO ELECTROMAGNETIC INTERFERENCE

Magnetic fields are produced by currents in wires and more strongly by the coils. The fields produced by coils are important as a result of magnetic coupling, particularly when there are two or more coils in the circuit. The magnetic coupling between coils may be controlled by large spacing between coils, the orientation of coils, the shape of the coils, and shielding.

Inductive sensors come in different shapes and sizes. Even though some sensors have closed cores such as toroidal shapes, others have open cores and air gaps between cores and coils. Closed cores may have practically zero external fields, except small leakage fluxes. Even if the sensors do not have closed cores, most variable-inductor sensors have a rather limited external field, as a result of two neighboring sets of coils connected in opposite directions, thus minimizing the external fields. Because the inductive sensors are made from closed conductors, a current will flow, if the conductor moves in a magnetic field. Alternatively, a magnetic change produces current in stationary closed conductor. Unless adequate measures are taken, there may be external magnetic fields linking (interference) with the sensor coils, thus producing currents and unwanted responses.

Because of inherent operations, inductive sensors are designed to have a high sensitivity to magnetic flux changes. External electromagnetic interference and external fields can affect the performance of the sensors seriously. It is known that moderate magnetic fields are found near power transformers, electrical motors, and power lines. These small fields produce current in the inductive sensors elements. One way of eliminating external effects is accomplished by magnetic shielding of the sensors and by grounding appropriately. In magnetic shielding, one or more shells of high permeability magnetic materials surround the part to be shielded. Multiple shells may be used to obtain a very complete shielding. Ends of each individual shell are separated by insulation so that the shell does not act as a single shorted turn, thus accommodating high current flows. Similarly, in the case of multiple shielding, shells are isolated from each other by proper insulation.

Alternating magnetic fields are also screened by interposing highly conductive metal sheets such as copper or aluminum on the path of the magnetic flux. The eddy currents induced in the shield give a counter mmf that tends to cancel the interfering magnetic field. This type of shielding is particularly effective

at high frequencies. Nevertheless, appropriate grounding must be observed.

In many inductive sensors, stray capacitances may be a problem, especially at null position of the moving core. If the capacitive effect is greater than a certain value, say 1% of the full-scale output, this effect may be reduced by the use of a center-tapped supply and appropriate grounding.

#### BIBLIOGRAPHY

1. J. P. Bentley, *Principles of Measurement Systems*, 2nd ed., Buncell Mills, UK: Longman Scientific and Technical, 1988.
2. E. O. Doebelin, *Measurement Systems, Application and Design*, 4th ed., New York: McGraw-Hill, 1981.
3. J. P. Holman, *Experimental Methods for Engineers*, 5th ed., New York: McGraw Hill, 1989.
4. J. E. Lewis, A review of magnetic sensors, *Proc. IEE*, **79**(4): 574-589, 1990.
5. W. Göpel, J. Hesse, and J. N. Zemel, *Sensors—A Comprehensive Survey*, Weinheim, Germany: VCH, 1989.
6. J. Evers, *Concise Encyclopedia of Magnetic and Superconducting Materials*, New York: Pergamon, 1992.

HALIT ERGEN

Curtin University of Technology

#### MAGNETIC SEPARATION

Commercial applications using the principles of magnetic separation have been around for more than one hundred years. Magnetic separators have been used since the time of Joseph Henry (1797-1878), when electromagnets were used to remove nails from horses' feed (1). Principles of magnetic separation are widely used in commercial applications today. Typical uses range from the simple removal of coarse tramp iron and steel from garbage, to more sophisticated separations, such as the removal of weakly magnetic mineral contaminants from paper-coating clays. Technical advances in magnetic separator design have led to the commercial use of high gradient magnetic separators (HGMS). These devices are capable of removing weakly magnetic particles and processing tons of material per hour. This ongoing progress has greatly broadened and enhanced the commercial magnetic separations market. For example, it is estimated that the introduction of HGMS into the purification of kaolin clay has nearly doubled the worldwide useful kaolin reserves (2) by making lower grade ores economically attractive.

Until the development of HGMS, magnetic separation techniques had been confined to manipulating mixtures that contained one or more of the three strongly magnetic (ferromagnetic) elements: iron, nickel, and cobalt. High gradient magnetic separators are potentially applicable to many more elements, mixtures, and compounds. For example, there are more than 56 weakly magnetic elements (diamagnetic and paramagnetic) listed in the periodic table. Perhaps of even greater potential benefit is the possibility of manipulating non-magnetic substances, for example, pollutants in water—using appropriate magnetic "seeding" techniques. (The term *raw magnetic* is typically reserved for materials that display extremely weak paramagnetic or diamagnetic properties.) Recently HGMSs, fabricated with superconducting coils, have

been introduced into the kaolin industry. Superconducting HGMS devices can produce even higher magnetic fields and operate with less than one fifth the total power consumption of conventional resistive units. Higher magnetic fields translate to better particulate selectivity and increased productivity by processing more material per unit container volume. For a more complete description of the history and development of high gradient magnetic separation, see Ref. 3.

## PRINCIPLES OF MAGNETIC SEPARATION

### Magnetic Phenomena

All materials possess magnetic properties to some extent. Their relative strength varies widely among materials. Materials are typically classified into four categories based on the strength of their magnetic properties: (1) ferromagnetic, (2) strongly magnetic, (3) weakly magnetic, and (4) nonmagnetic. With the exception of the ferromagnetic, the boundary defining the difference between strongly magnetic, weakly magnetic, and nonmagnetic materials is arbitrary and typically application-dependent. Any particle exposed to a magnetic field will become magnetized. The magnetic moment of an atom results from the electron spins, their orbital angular momentum, and the change induced in the orbital angular momentum (4). Materials that exhibit positive magnetization when placed in an external magnetic field are described as paramagnetic. Materials that exhibit negative magnetization are described as diamagnetic. All materials will display some degree of diamagnetism owing to the moment induced by an applied magnetic field; however, this type of diamagnetism is relatively weak and can be negligible if other forms of magnetism are present (e.g., paramagnetism or ferromagnetism). Paramagnetism results from the electron spins and their orbital angular momentum. Atoms with unpaired electrons typically exhibit paramagnetism. Paramagnetic materials are further classified as strongly or weakly paramagnetic depending on the strength of the magnetization (magnetic moment per unit volume) when placed in an external magnetic field. Many elements and their compounds exhibit paramagnetism; some common examples are hematite ( $\text{Fe}_2\text{O}_3$ ) and pyrite ( $\text{FeS}_2$ ). Ordered arrays of magnetic moments result in phenomena such as ferromagnetism, antiferromagnetism, and ferrimagnetism. When domains of paramagnetism are created in some materials such that long-range order is established, the magnetization can be quite large, and the materials are described as ferromagnetic. These materials include iron, nickel, and cobalt and a relatively small number of compounds of these elements. An example of a ferromagnetic ore is magnetite ( $\text{Fe}_3\text{O}_4$ ).

### Basic Electrodynamics

A magnetic field and magnetic field gradient can be produced in a variety of ways, typically by current-carrying conductors (electromagnets) or by residual magnetization of ferromagnetic materials (permanent magnets) or a combination of these. The magnetic induction  $\mathbf{B}$  in teslas (T) is a vector quantity defined in terms of the sideways force  $F_{\perp}$  in newtons (N) acting on a test particle with charge  $q$ , in coulombs (C) and moving with a velocity  $v$  in meters per second (m/s) such that

$$F_{\perp} = qv \times B \quad (1)$$

where  $v \times B$  is the outer (or cross) product between the vectors  $v$  and  $B$  (5). The space around a magnet or current-carrying conductor is defined as the size of the magnetic field. The size of the magnetic field is the space through which its influence extends and is mapped by lines of magnetic force. A magnetic field is said to be uniform and homogeneous when these lines are parallel and equally spaced. Using the International System of Units (SI), the magnetic induction (which is sometimes referred to as the magnetic flux density or simply the magnetic field) in teslas is calculated by

$$B = \mu_0(H + M) \quad (2)$$

where  $\mu_0$  is the permeability of free space equal to  $4\pi \times 10^{-7}$  N/A<sup>2</sup>,  $H$  is the magnetic field intensity measured in amperes per meter (A/m), and  $M$  is the induced magnetization measured in amperes per meter (A/m). The magnetic quantities  $B$ ,  $H$ , and  $M$  are often described in many textbooks using centimeter-gram-second (cgs) units. In cgs units, the magnetic induction  $B$  is expressed in gauss (G), the magnetic field intensity  $H$  in oersteds (Oe), and the magnetization  $M$  in electromagnetic units per cubic centimeter (emu/cm<sup>3</sup>). To convert between the two systems of units, several reference tables have been published elsewhere (6). One convenient conversion factor to recall is 1 T = 10,000 G.

The magnetic susceptibility  $\chi_m$  (volume susceptibility) of a material is a dimensionless parameter and is often used to describe the relative magnetic strength of the material. The magnetic susceptibility is defined as the ratio of the induced magnetization to the magnetic field intensity and is given by

$$\chi_m = M/H \quad (3)$$

The magnetic induction can now be expressed in terms of the magnetic susceptibility

$$B = \mu_0 H(1 + \chi_m) \quad (4)$$

Table 1 shows the magnetic susceptibility of some common minerals and elements. Paramagnetic materials have positive susceptibilities, and induced moments enhance the magnetic induction. Diamagnetic materials have negative susceptibilities and reduce the magnetic induction. Ferromagnetic materials have very large positive susceptibilities but become magnetically saturated above certain applied magnetic field intensities; thus, a further increase in magnetic field intensity does not produce a further increase in magnetic induction. Figure 1 shows the magnetic induction of some ferromagnetic materials as a function of the applied magnetic field intensity (4).

### Magnetic Separation Dynamics

**Magnetic Forces.** Several magnetic separation concepts have been proposed over the years, but they all rely on the same electromagnetic principle: a particle exposed to an external spatially varying magnetic field (i.e., a magnetic field gradient) will experience a force in newtons (N) equal to

$$F_M = VM(B) \text{ grad } B \quad (5)$$

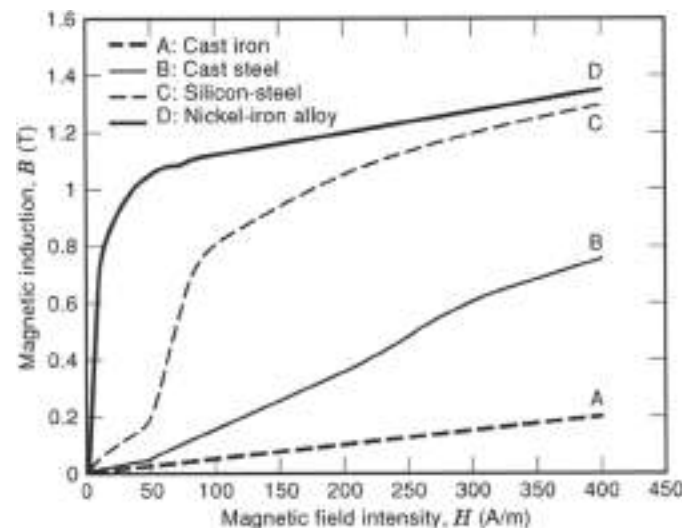
**Table 1. Magnetic Susceptibility of Some Common Elements and Minerals**

| Substance                      | Susceptibility ( $10^{-6}$ cgs) | Substance   | Susceptibility ( $10^{-6}$ cgs) |
|--------------------------------|---------------------------------|-------------|---------------------------------|
| Alumina                        | 110.6                           | Feberite    | 139.3                           |
| Al <sub>2</sub> O <sub>3</sub> | 237                             | Galena      | 20.4                            |
| Apatite                        | 11.0 to 118.0                   | Garnierite  | 130.7                           |
| Aragonite                      | 20.4                            | Gold        | 288.0                           |
| Auburn                         | 1150.0                          | Ilmenite    | 116.46 to 170.0                 |
| Azurite                        | 112.2 to 118.0                  | Lead        | 228                             |
| Avantone                       | 10.98 to 15.80                  | Malachite   | 110.6 to 114.5                  |
| Beryl                          | 10.4                            | Millerite   | 10.21 to 13.9                   |
| Braunite                       | 135.0 to 1150.0                 | Molybdenite | 14.9 to 17.1                    |
| Biotite                        | 140.0                           | Molybdenum  | 189.0                           |
| Barite (pure)                  | 271.2                           | Rutile      | 10.85 to 14.78                  |
| Braunerite                     | 13.5                            | Scheelite   | 10.13 to 10.27                  |
| Chromium                       | 1180.0                          | Siderite    | 166.2 to 1103.8                 |
| Chromite                       | 1126.6 to 1450.0                | Titanium    | 1150.0                          |
| Cobalt                         | Ferromagnetic                   | Tungsten    | 169.0                           |
| Cobaltina                      | 12.0                            | Uranium     | 1395.0                          |
| Cobaltite                      | 10.34 to 10.84                  | Vanadium    | 1255.0                          |
| Columbite                      | 182.6 to 187.2                  | Vanadinite  | 20.2 to 10.27                   |
| Copper                         | 20.1                            | Wolframite  | 142.2                           |

where  $V$  is the volume of the particle in cubic meters ( $\text{m}^3$ ),  $M(B)$  is the magnetic field dependent magnetization of the particle in amperes per meter, and  $\text{grad } B$  is the gradient of the magnetic induction in teslas per meter. The magnetic force acting on a particle can be rewritten in terms of the magnetic susceptibility where  $M(B) = \chi_m(B) H$  such that

$$F_M = \chi_m(B) V H \cdot \text{grad } B \quad (16)$$

The implications of Eq. (16) are that in order to have large magnetic separation forces, not only is the particle's magnetic



**Figure 1.** Magnetization curves of four ferrimagnetic materials: (A) cast iron, (B) cast steel, (C) silicon-steel, and (D) Nickel-iron alloy. Courtesy of John Wiley & Sons.

susceptibility important, but a combination of high magnetic field intensity and magnetic field gradient determines the magnitude of the magnetic force.

**Fluid Dynamic Effects.** For many magnetic separation processing techniques (i.e., wet processing), the particles are often mixed in an aqueous fluid (slurry) that moves past the high gradient collector. Typically, the high gradient collector is a mesh fabricated from thin ribbons or wires of highly permeable stainless steel, which concentrates the magnetic field lines. In general, magnetic field gradients are large near sharp edges or corners. Thus, the magnetic force (given in Eqs. (5) and (6)) acting on the particle can be quite large. The mesh acts as a filtering mechanism by trapping magnetic particles attracted to the wire and allowing nonmagnetic particles to pass freely. While traveling within this fluid, however, each particle in the mixture is acted on by several other competing forces, including magnetic force ( $F_M$ ), fluid viscous drag force ( $F_D$ ), gravitational force ( $F_G$ ), and fluid buoyant force ( $F_B$ ). If the buoyant force of the particle is ignored, Newton's law of motion for a particle in a fluid is given by (7)

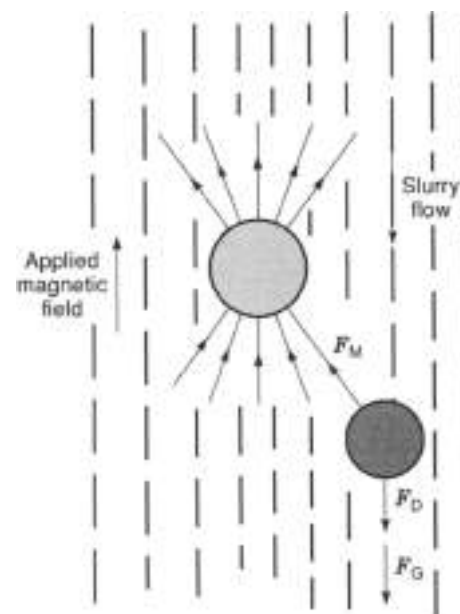
$$\rho V (d^2r/dt^2) = F_M + F_B + F_G \quad (7)$$

where  $\rho$  is the particle density in kilograms per cubic meter,  $V$  is the particle volume in cubic meters,  $r$  is the position coordinate in meters,  $t$  is the time in seconds, and the magnetic force is as defined in Eqs. (5) and (6). To a first-order approximation, the viscous drag force is given by

$$F_D = 3\pi\eta\phi v \quad (8)$$

where  $v$  = fluid velocity ( $\text{m/s}$ ) = particle diameter ( $\text{m}$ )/ $\eta$  = fluid viscosity ( $\text{N s/m}^2$ )

Figure 2 illustrates the forces acting on a magnetic particle as described in Eqs. (5) and (8). For a magnetic particle to



**Figure 2.** Schematic representation of a magnetic fiber in a high gradient magnetic separator and the forces acting on a magnetic particle. Courtesy of Plenum Publishing Company.



be collected at the surface of the wire within the magnetized volume of the separator, the following condition must be met.

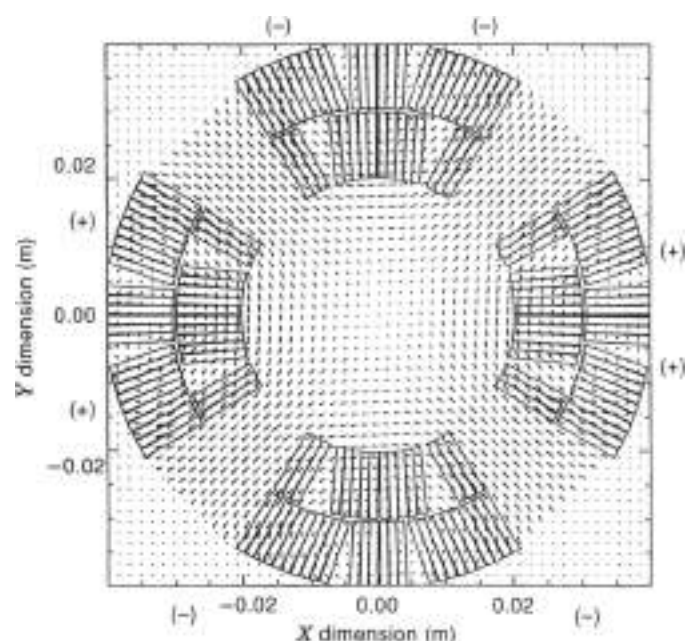
$$F_M \geq F_D - F_b \quad (9)$$

In this analysis, both the density and particle size (diameter) are extremely important in the magnitude of the magnetic and viscous drag forces. The magnetic force varies as the cube of the particle diameter ( $d^3$ ). The viscous drag force varies linearly with the particle diameter. This implies that separations between particles with similar magnetic susceptibilities are possible only when significant size and density differences exist.

**Open Gradient Magnet Systems.** The large magnetic field gradients used in separation devices are typically created using two different methods. The first method is known as the open gradient magnet system (OGMS). The second is the matrix/filter method. Open gradient devices separate particles with different magnetic susceptibilities by preferentially directing them into open spaces using magnetic forces. In the open gradient technique, the conductor geometry or suitably designed pole pieces generate magnetic field gradients and hence the magnetic forces on the particles. The magnetic forces are often in competition with other forces such as gravity, electrostatics, drag, and buoyancy. A significant shortcoming of open gradient systems is that the magnetic field gradients that can be generated are typically less than 2 T/cm. Typically, these low magnetic field gradients limit the use of OGMS devices to the separation of strongly magnetic materials and metal scrap.

Open gradient systems have an advantage in high-volume processing. No magnet matrix or filter is required to trap the magnetic particle mechanically. In standard operation, the particle stream is passed down the bore, and magnetic particles are either pushed radially inward or outward depending on whether they are diamagnetic or paramagnetic. The particles can then be physically separated further downstream. For example, a quadrupole (four-pole) magnet could be used in an open gradient system. The conductor windings in a quadrupole magnet generate a magnetic field that is near zero at the center of the bore and increases radially outward. The magnet has four distinct racetrack-shaped windings located within each quadrant at the perimeter of the bore. These racetrack-shaped windings generate the four poles of the magnet (see Fig. 3). Shown in Fig. 3 is a two-dimensional cross-sectional view of a superconducting quadrupole magnet with a 40 mm bore. In Fig. 3, the direction of the particle stream flow is into the page. The plus (+) symbol indicates current flow into the page, and the minus (-) symbol indicates current flow out of the page. The arrows in the figure represent the relative magnetic field strength generated by the conductor windings and illustrate the magnetic field gradient. The longer the arrow, the higher the magnetic field strength. The conductor windings are fabricated with a high aspect ratio (width/thickness) superconductor. In order to match the rectangular-shaped conductor with the cylindrical geometry of the magnet, V-shaped metal wedges are placed at strategic locations within the cross section. The magnetic field within the bore is periodic and fourfold symmetric and can be described mathematically by a  $\cos^2 2\theta$  function.

**Matrix/Filter Systems.** The second most often used method for generating magnetic field gradients in separation devices



**Figure 3.** Two-dimensional cross-sectional view of a quadrupole magnet with a 40 mm bore. The size of the arrows denote the relative magnetic field strength. The plus (+) symbol denotes current flow into the page, the minus (-) symbol denotes current flow out of the page.

is by use of a magnetic matrix or filter. Frantz patented the conceptual basis for the modern ferromagnetic filament-type collector matrix (8) in 1937. The Frantz magnetic filter (Frantz Ferrufilter<sup>SM</sup>) used a matrix of ferromagnetic type 430 stainless steel screens fabricated from thin sharp ribbons. In general, magnetic field gradients are large near sharp edges or corners. The introduction of a large number of spheroids, wires, or a mesh of wires will similarly create regions of high gradients near the wires. The magnetic field gradient near the vicinity of the wire can be as high as 10,000 T/cm. A magnetic particle traveling past the wire will experience two magnetic fields: one from the background ambient field ( $B_0$ ) of the magnet and the other from the collector wire ( $B_c$ ), where the total magnetic field is the vector sum of the two contributions, namely,  $B_t = B_0 + B_c$ . A more detailed analysis of how matrix shape affects the capture cross section of a particle can be found elsewhere (9).

One method to increase the separation efficiency is to increase the magnetic field gradient. To obtain the highest possible field gradient, it is important to use the finest possible matrices (filamentary type). The most cost-effective separations occur when the filament diameter of the matrix and hence the range of the magnetic field gradient is matched with the diameters of the particulates being processed. If the particle is much larger than the range of the gradient, then only a small portion of the particle will feel the effect of the force. On the other hand, if the particle is small compared with the range of the gradient, then the difference of the magnetization across the particle will not be large. If the difference in magnetization across the particle is small, then the resulting magnetic force acting on the particle will also be small.

By increasing the magnetic field gradient using a high permeability matrix, the required magnetic field intensity ( $H$ ) is

correspondingly reduced. Reducing the required background magnetic field intensity has a twofold positive impact on the cost of the separation system. First, it lowers the initial capital cost of the magnet system by requiring a lower ambient field. Second, it lowers the operating cost of the separation system by reducing the amount of electricity required to power the electromagnet. In a resistive electromagnet, the magnetic field intensity that is generated is proportional to the square root of the electric power. Therefore, as the magnetic field intensity increases, the amount of electricity required to power it increases rapidly.

A significant disadvantage of the matrix/filter system is the problem of magnetic particle build-up. A magnetic separation device that uses a matrix/filter to mechanically trap magnetic particles must be periodically cleaned to prevent clogging. This means that during the cleaning of the matrix/filter, the magnetic field must be reduced to zero (or near zero) to minimize the magnetic trapping force. Periodically cycling the magnetic field or removing the matrix/filter from the magnetic field in order to clean it reduces processing efficiency.

#### CHARACTERISTICS OF HIGH GRADIENT MAGNETIC SEPARATION

The efficiency of a magnetic separation device is typically expressed in two ways (10). First, it may be characterized by the so-called grade, which is defined as the percentage ratio of the desired component's mass relative to the total mass of the magnetic fraction (magn). Second, the separation efficiency can be characterized by the so-called recovery, which is defined as the percentage ratio of the amount of magnetic material recovered relative to the total amount of magnetic material in the feed. The two measures are independent quantities and together determine the efficiency of the separation. For most separation devices, there is a relation between the grade and the recovery of the processed material. By adjusting the operating parameters of the separator, the grade can be increased at the expense of the recovery, and vice versa. For example, Fig. 4 shows

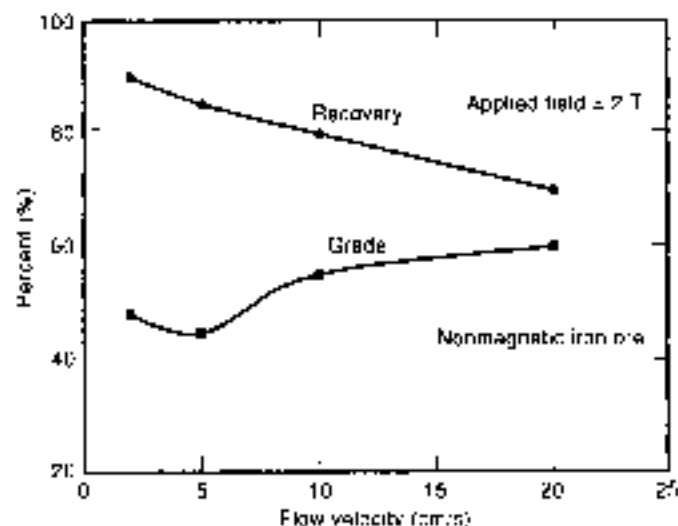


Figure 4. Grade and recovery of a nonmagnetic iron ore (manganese) versus slurry flow velocity at a constant applied field of 2 T. Courtesy of Plenum Publishing Company.

processing data for the magnetically trapped product of a concentrate. The percent recovery and grade is plotted as a function of the slurry flow rate at a constant applied field of 2 T. As the recovery of the magnetically trapped product decreases with increasing flow rate, the grade correspondingly increases.

#### KEY PROCESSING VARIABLES

The key process variables in paramagnetic mineral separations are (a) magnetic field strength, (b) high gradient matrix, (c) retention time, and (d) minimizing the competing forces acting on the particle (11, 12).

##### Magnetic Field Strength

The magnetization  $M/H$  [see Eqs. (5) and (6)] of a particle is a general function of the ambient magnetic field. Maximizing the magnetic field in a separator also increases the magnetic attractive force. For ferromagnetic materials, this effect generally saturates at fields above 2 T (see Fig. 1). For paramagnetic materials, the magnetic moment increases with increasing magnetic field well beyond 2 T. For these materials, larger magnetic fields enhance the magnetic trapping force.

##### Magnetic Field Gradient

The magnetic field gradient is the spatial variation of the magnetic field strength over the microdistance adjacent to the filament collector surface. The key factor in the generation of the magnetic field gradient is the transition of the magnetic flux from the highly permeable magnetic medium, typically type 430 stainless steel wool with a permeability  $\sim 500$ , to the surrounding medium containing the particulates, typically air or water with a permeability of 1. The geometry of the conductor (e.g., sharp points), also plays a significant role in establishing the magnetic field gradient. The greater the number of such interfaces, the more effective the collecting power of the matrix. The smaller the diameter of the filament, the higher the magnetic gradient, but the shorter the range of its influence. High collection efficiencies have been demonstrated by a variety of magnetic fibers and small magnetic spheres. Stainless steel wool is an exceptionally efficient low-cost collector and is widely used in the commercial processing industry. In particular, 430 stainless steel is often used because it is magnetically "soft" and does not have a high residual field. Having a low residual field is necessary during the cleaning/flushing process of the matrix, where the magnetic trapping force on the particle must be as small as possible. The advantage of this type of filter is that it has a high collection surface but occupies a relatively small portion of the total magnetized volume, leaving the remainder of the space for material to be processed. The amount of stainless steel per magnetized volume is application-specific but typically varies between 6% and 12%.

##### Retention Time

The retention time is a measure of how long a particle contained within a slurry is influenced by the magnetic field. The retention time is calculated by dividing the canister length by the velocity of the fluid flow or alternatively by dividing the canister volume by the flow rate per minute (11, 12). The concept of retention involves control of the flow rate in the canister

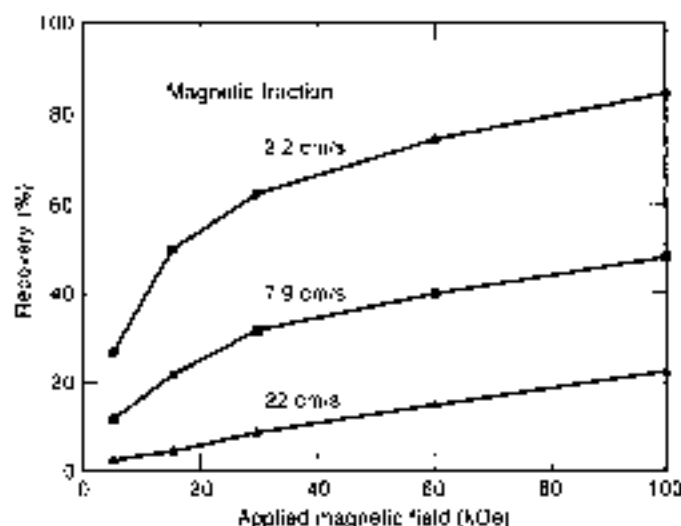


Figure 5. Recovery percent of desired component of magnetic fraction (magn) relative to feed material versus applied magnetic field for three different values of slurry flow rate.

volume to balance the magnetic force of attraction to the filament collector surface with the viscous drag of the medium of suspended particles. Retention time influences both product quality and production rate. The retention time used in commercial processes is application-specific and is usually chosen by a trial and error approach with many factors such as system back pressure and slurry flow distribution through the canister also being considered. Typically, commercial slurry feed times used in the purification of kaolin clay vary between 2 min. and 4 min. Experimental results demonstrating the effects of retention time and applied magnetic field are shown in Figs. 5 and 6. Figure 5 is a plot of the percent recovery for the magnetic fraction versus the applied magnetic field for three different values of flow rate. The test material consisted of 10% cupric oxide and aluminum oxide powder slurred in water, where the magnetic material being separated is the cupric oxide. Figure 6 shows the recovery of iron for a nonmagnetic taconite ore versus

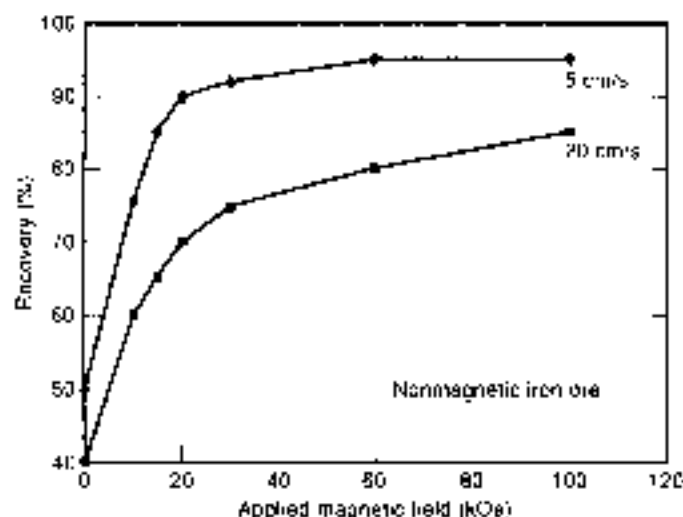


Figure 6. Recovery of nonmagnetic iron ore (magn) versus applied field at two different values of slurry flow rate.

the applied magnetic field for two different values of flow rate. In both Figs. 5 and 6, at constant flow rates the percent recovery increases with increasing applied field. This can be attributed to the increase in the magnetic force acting on the particle in the fluid [see Eq. (6)]. At constant applied magnetic field, the percent recovery decreases with increasing flow rate. This can be attributed to the increase in the viscous drag force, which is proportional to the velocity of the slurry [see Eq. (8)]. The trade-off between the magnetic field and the flow capacity is ultimately dictated by the economics of the application. Higher magnetic fields translate to higher particulate recovery, but at a higher capital equipment cost. Higher flow rates translate to increased capacity, but at lower particulate recovery.

#### Minimizing Competing Forces

Many magnetic separators that process minerals suspended in fluids try to minimize the competing gravitational and fluid dynamic forces. As seen from the right-hand side of Eqs. (7) and (9), minimizing the viscous drag force and gravitational force increases the chances for magnetic capture and correspondingly increases the separation efficiency. To minimize the unwanted gravitational effects, unprocessed slurry is typically fed upward (i.e., introduced from the bottom of the canister).

As seen from Eq. (7), the viscous drag force is proportional to the product of the viscosity and linear flow velocity. To reduce the viscous drag force, a considerable effort is made to select a suitable dispersant to add to the slurry in order to minimize the fluid viscosity. For example, kaolin clays are typically dispersed in water with 0.2% tetrasodium pyrophosphate in order to provide deflocculated slurries. To further minimize the viscous drag force, the slurry is fed at as low a linear velocity as practical while maintaining allowable commercial processing rates (see also the previous section entitled "Retention Time"). A final consideration in the reduction of the viscous drag force is the percent solids content of the slurry itself. Slurries with high solids content are more viscous, which decreases particle capture. However, slurries with low solids content process less material per canister volume and have much quicker particulate settling times, which can lead to clogging in the feed, transfer, and distribution piping. Typically, trial and error is used to find the optimal processing conditions to accommodate these two competing effects. In kaolin clays for example, the percent solids content of the slurry typically varies between 20% and 40%.

#### MAGNETIC SEPARATION EQUIPMENT

Several terms are often used to characterize magnetic separation devices. These include permanent magnet, electromagnet, high gradient, open gradient, etc., but the devices themselves basically fall into two separate categories: (1) batch type and (2) continuous type. Table 2 summarizes some common types of magnetic separators and their ranges of potential applications. The following information concerning commercial magnetic separations units can be found in more detail in Ref. 1.

##### Batch type

Batch-type separators are most useful when the unprocessed feed materials contain a relatively small amount of magnetic

Table 2. Potential Applications of Magnetic Separators

| Device         | Magnet          | Maximum Field (T) | Matrix         | Maximum Gradient (T/cm) | Required Susceptibility | Particle Size (mm) |
|----------------|-----------------|-------------------|----------------|-------------------------|-------------------------|--------------------|
| Grate          | Permanent       | 0.05              | Rada           | 0.05                    | Ferromagnetic           | .12                |
| Pulley         | Permanent       | 0.02              | -              | $0.01 \pm 0.1$          | Ferromagnetic           | .50                |
| Drum           | Permanent       | $0.05 \pm 0.1$    | -              | $0.05 \pm 0.1$          | Ferromagnetic           | $0.02 \pm 20$      |
| Belt           | Electromagnet   | $0.01 \pm 0.1$    | -              | $0.01 \pm 0.1$          | Ferromagnetic           | $0.15 \pm 30$      |
| Induced roll   | Electromagnet   | 2                 | -              | 20                      | Paramagnetic            | $0.03 \pm 3$       |
| Carpco         | Electromagnet   | 2                 | Steel balls    | 4.5                     | Paramagnetic            | $0.01 \pm 1$       |
| C-frame; Jones | Electromagnet   | 2                 | Grooved plates | 20                      | Paramagnetic            | $0.01 \pm 2$       |
| Marston Sala   | Electromagnet   | $2 \pm 5$         | Steel wool     | 2500                    | Paramagnetic weak       | $0.0001 \pm 2$     |
|                | Superconducting |                   |                |                         |                         |                    |

material to be trapped in the collection volume. This allows for a convenient duty cycle for the device. Batch-type separation devices are most often fabricated using iron-bound solenoid electromagnets surrounding a cylindrical canister. The cross section of the canister is circular in order to minimize the amount of conductor used to magnetize the collection volume. Utilizing a long coil where the height of the coil exceeds the canister height further maximizes the efficiency of a batch-type solenoid electromagnet. Iron pole pieces of the magnet are then designed to extend into the top and bottom of the solenoid. Batch-type separators operate on the principle of cycling the magnetic field from a maximum value down to zero applied field. It is important to have a magnetically soft material with a low remanent field as the collection matrix, so that when the applied field is reduced to zero the force acting on the magnetic particles is minimized. In normal operation, typical duty cycles for mineral separations vary between 10% and 80% and consist of five basic steps (see Fig. 7 for illustrative purposes): (1) The current in the magnet is increased to obtain the desired operating magnetic field level. During this step the collection matrix (typically

compressed 430 stainless steel wool) is magnetized. (2) While maintaining the desired operating magnetic field level, the appropriate valves are opened/closed and the unprocessed slurry (feed) is introduced and flows through the magnetized collection volume. The nonmagnetic material which passes through the magnetized volume is collected separately, whereas the magnetic material remains trapped on the 430 stainless steel collection matrix. (3) By shutting off the appropriate valves, the flow of the unprocessed feed material into the collection volume is stopped, and the collection volume is subsequently displaced with water. (4) The current in the magnet is decreased to zero or near zero. This step demagnetizes the collection matrix and minimizes the magnetic force acting on the captured magnetic particles. (5) Appropriate valves are opened/closed, and the (now demagnetized) collection matrix is subjected to a high velocity water rinse/flush and the magnetic by-product that is liberated from the matrix is collected separately. The steps of the process are then repeated based upon some established procedure. A schematic representing this process is shown in Fig. 7 (10).

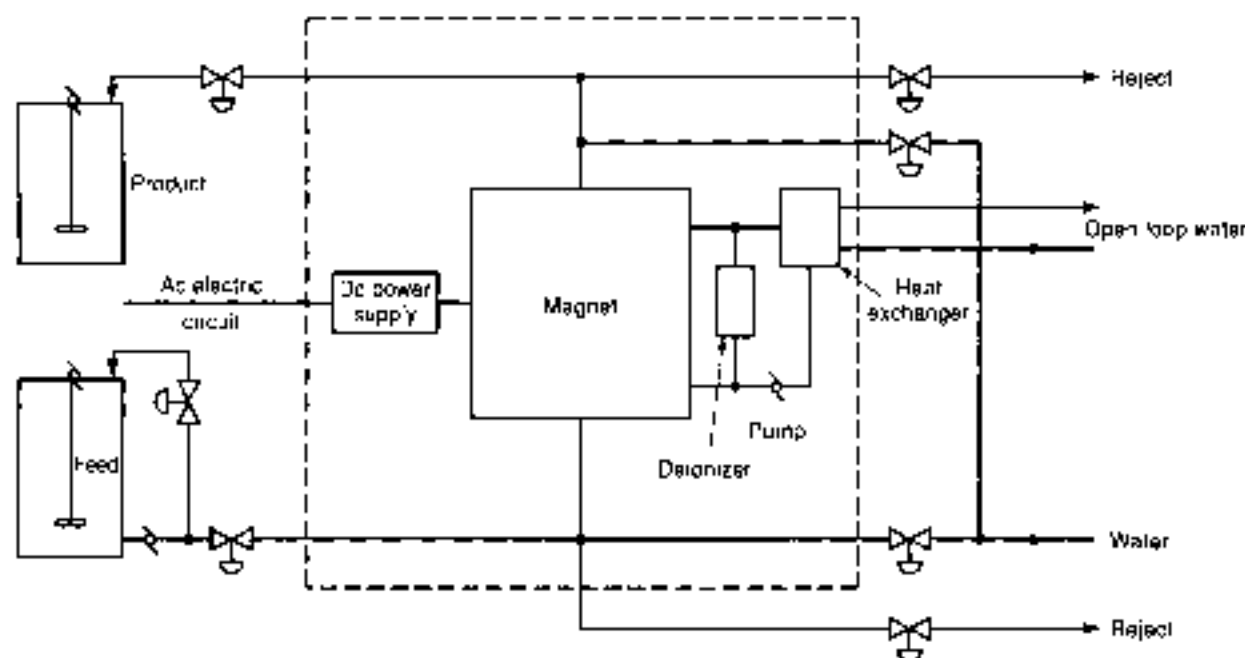


Figure 7. Schematic of batch type process used in magnetic separation. Courtesy of AquaSre Corp.

**Induced-Pole Separators.** In induced-pole separators the magnetic field gradient is generated by the application of background magnetic field to a ferromagnetic filter matrix. The word "induced" is somewhat misleading because there are no ac fields present. The name "induced-pole" refers to the magnetic poles that are induced around the ferromagnetic matrix/filter edges by the application of the background magnetic field. The orientation of the magnetic field to the edge of the matrix determines whether the device is a parallel or perpendicular field-to-flow. These batch separators are manufactured in two basic configurations: (1) C-frame and (2) solenoid. Solenoid electromagnets represent a major area of development in magnetic separation and are described in detail in the following three sections of this article.

C-frame magnets employ a ferromagnetic matrix placed between two poles of an electromagnet. The ferromagnetic material used to transfer the magnetic flux from pole to pole typically occupies between 40% and 80% of the magnetized volume. A disadvantage of this type of device is the nonuniformity of the background ambient field. With even the smallest amounts of ferromagnetic particulates present in the feed material, these devices are subject to clogging in the flush region where the fringe field is the highest.

#### Resistive Solenoid Electromagnets

**Installed Systems.** In March 1973, the first large-scale HGMS was installed at Freeport Kaolin Company in Gordon, Georgia (2). The separator was a 2.1 m diameter resistive electromagnet fabricated by Pacific Electric Motor Company. The separator could generate magnetic fields up to 2 T and process nearly 3800 liters of slurry per minute or 60 metric tons per hour (dry basis). The central magnetic field was 2 T, with an electric power consumption of nearly 500 kW. The inductor consisted of 16 hollow water-cooled copper coils surrounded by a vaultlike steel enclosure that was 3.65 m  $\times$  3.65 m  $\times$  2.44 m (13'). For its day, this magnet represented a 13-fold increase in process capacity over the largest previous commercial magnetic separator. Continued improvements in magnet design have led to similarly sized units operating at 2 T with less than 300 kW of electric power consumption. In 1982, the first 3 m diameter resistive electromagnetic separator went into operation. This enormous unit could process 130 tons of kaolin per hour with about 400 kW of electric power consumption. To date, there have been 29 resistive electromagnets (2.1 m and 3 m diameter) installed in the United States and 8 others in the rest of the world for the purification of kaolin clay.

#### Low-Temperature Superconducting Solenoid Electromagnets

**Background.** In a low-temperature superconducting (LTS) magnet, the magnetic field is generated in exactly the same way as in a conventional resistive electromagnet. The only real difference between the two is that the conductor in the LTS magnet must be maintained at a suitably low temperature in order to remain in its superconducting state. The key benefits offered by superconducting magnets are (1) very low power consumption, resulting from zero resistance in the conductor windings (see next section), and (2) much higher magnetic fields resulting in better selectivity of particles and higher separation efficiency. However, the zero resistance property of a superconductor is only true for direct current (dc). For applications where the electromagnet field is changing in time (ac),

a superconducting material no longer operates with zero electric resistance. Therefore, for batch-type magnetic separators, which periodically cycle the magnetic field, there is a practical limitation on the maximum allowable magnetic field ramp rate for these devices.

Several LTS materials have been studied, but the most prominent in terms of conductor fabrication and commercial implementation is an alloy of niobium and titanium (NbTi). In order to enhance electric and thermal stability of the superconductor, the NbTi is typically embedded in a normal metal matrix of copper or aluminum. This allows for greater heat transfer through thermal conduction and also provides a low-resistance electric path in the event that the superconductor comes out of its superconducting state. One disadvantage of the normal metal matrix is in ac applications. In ac applications, additional Joule heating is caused by the generation of induced currents in the normal metal matrix. To fabricate these composite superconductors, fine filaments of NbTi are either drawn or extruded with the aluminum or copper. The filament size of the NbTi can vary between about 5  $\mu$ m and 30  $\mu$ m depending upon the application. Copper and aluminum-clad superconducting windings using NbTi conductor are now commonplace in the research community and have been used in applications such as motors, generators, transformers, and magnets.

An important aspect of the design of an LTS magnet is the choice of refrigeration or cryogenic cooling system that will be used to cool the conductor windings. There are two basic refrigeration routes that have been used successfully. The first and most prevalent route is the use of liquid or gaseous cryogenics. The NbTi conductor windings are typically cooled by either immersion in a bath of liquid helium or by forcing cold, two-phase liquid-vapor helium gas around the conductor. Helium gas will liquefy at 4.2 K at atmospheric pressure. The helium gas that is boiled off or heated by ac-loss is recirculated through a liquefier, cooled, and recondensed into liquid or a two-phase vapor-liquid mixture. The second method used to cool superconducting windings is through indirect cooling using a refrigerator. These refrigerators are commonly referred to as cryocoolers. Cryocoolers, based on various thermodynamic cycles, allow cooling to extremely low temperatures. One such device commonly used in the cryogenic industry is the Gifford-McMahon (G-M) cryocooler. Recent advances in G-M cryocooler technology permit these units to generate temperatures of 4 K or lower. The units enable superconducting windings to be cooled indirectly and without the presence of liquid or gaseous cryogenics and are particularly advantageous in small-scale systems where the cost of a liquefier cannot be justified. The two major drawbacks to these systems are: (1) the constant supply of electrical power that is essential for reliable operation and (2) the relatively small cooling capacity of cryocoolers (typically a few watts) at 4 K.

**Energy Efficiency.** One of the major drawbacks of conventional resistive-type magnetic separators using water-cooled copper conductor windings is their high operating cost owing to their large electrical power consumption. The power consumption in a conventional copper magnetic separator can be as high as 400 kW in a 3 m diameter unit. In a superconducting unit, the primary source of power consumption is the refrigeration unit. An equivalent 3 m diameter LTS unit is rated at about 50 kW. In production environments, these units operate approximately 8000 h/yr with an expected lifetime between 5



**Figure 8.** Low-temperature superconducting batch-type magnetic separator with a 3 m bore. Note the liquid helium refrigeration system to the right of the magnet. Courtesy of Erez Magnetics.

and 10 years. In the southern United States, where most of the processing plants are located, the present average price of electricity is approximately \$80.05/kWh. This translates to an annual cost saving of about \$140,000 with a lifetime saving of \$1.4 million for a superconducting magnet over its resistive counterpart. For processing plants located in nonindustrialized countries, the savings in annual operating cost is even more substantial as price and availability of electrical power is at a premium.

**Installed Systems.** The first large-scale LTS magnetic separator went into operation in 1986. The device had a 2.1 m diameter bore and was fabricated by Erez Magnetics of Erie, Pennsylvania (14). The device was installed at the J. M. Huber Company in Georgia and is used for the beneficiation of kaolin clay (see Fig. 8). The installed cost of this device was around \$2 million. To date, 12 batch-type LTS magnetic separation systems have been installed in the United States. Eight of the 12 LTS systems installed have been retrofits, where the existing resistive electromagnets have been replaced with superconducting windings. Worldwide, five other batch-type LTS magnetic separation systems have been installed for the beneficiation of kaolin clay in Australia, Brazil, China, England, and Germany. Low-temperature superconducting magnetic separation systems appear to be displacing their conventional resistive counterparts.

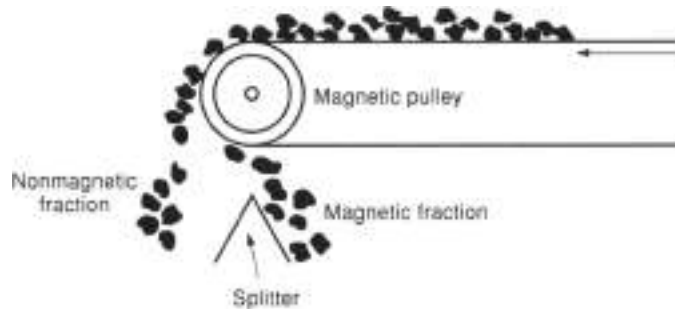
**High-Temperature Superconducting Solenoid Electromagnets.** In 1986, Bednorz and Mueller (15) discovered a ceramic oxide compound that would superconduct at higher temperatures than the previously well-standard LTS materials. Before this discovery, the highest recorded superconducting transition temperature was about 23 K. Since 1986, several more ceramic oxide materials, which superconduct at even higher temperatures, have been discovered. This class of ceramic oxide materials has been given the name high-temperature superconductors (HTS). At the time of this writing, the highest recorded superconducting transition temperature of an HTS material at atmospheric pressure was approximately 140 K. High-temperature

superconducting materials may offer substantial benefits in the practical commercialization of superconducting devices because of the higher temperatures at which these materials superconduct. Higher superconducting transition temperatures translate to higher (Carnot) refrigeration efficiencies. For example, an LTS material operating at 4 K in a bath of liquid helium would require more than 20 times the amount of electric power for every watt of cooling versus its HTS counterpart operating in a bath of liquid nitrogen at 77 K.

Presently magnetic separation represents one of only two industrial applications of large-scale superconducting devices; the other is magnetic resonance imaging (MRI). Consequently, one of the first prototype industrial devices fabricated using HTS wire was an HGMS. In 1996, the first successful demonstration of kaolin beneficiation using an HTS magnetic separator was reported (16). In this report, five different types of kaolin clays representing major worldwide deposits were processed in a 5 cm diameter warm bore HTS magnet in fields up to 2.6 T. Results indicated brightness improvements varying from one to five GE brightness units, depending on the particular clay processed. Most likely, HGMS will be one of the first industrial areas impacted by HTS technology and may represent the future migration of HGMS to this technology. However, because of the higher cost of the HTS wire per ampere-meter compared to its LTS counterpart, it is unclear at this time if the economic benefits from increased refrigeration efficiency will outweigh the additional capital cost of the HTS wire.

#### Continuous Type

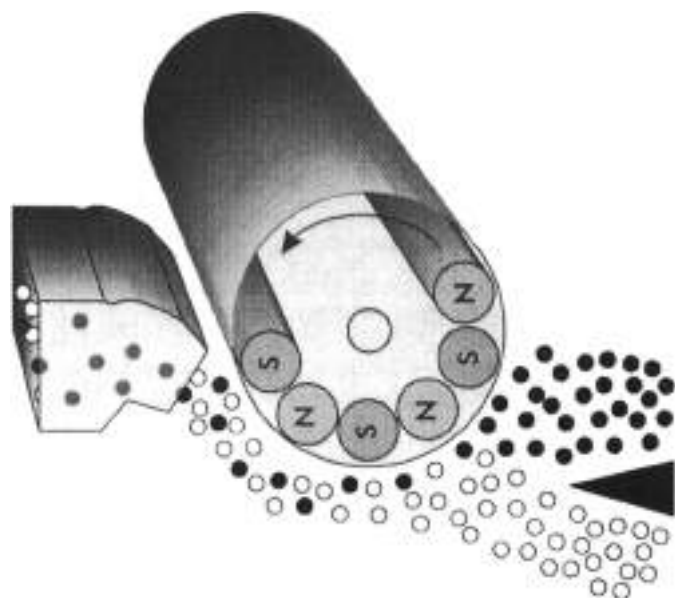
One of the primary shortcomings of the batch-type separator is processing efficiency. Batch-type separators operate on a duty cycle, so there is a period of time where material is not being processed. Continuous-type separation devices are designed to maximize the duty cycle and minimize the processing time. Continuous-type separators are advantageous when the magnetic fraction of the unprocessed feed material is relatively high.



**Figure 9.** Conventional pulley-type separator. Courtesy of McGraw-Hill Publishing Companies.

**Drum and Pulley Magnets.** In the earliest magnetic separators, and in many that are still applicable for attracting strongly magnetic materials, permanent magnets were used in an open single-surface device. These devices consisted of suspended magnets, pulleys, conveyors, or drums. These devices produced fields in the neighborhood 0.06 T and gradients of the order of 0.05 T/cm. Drum and pulley-type magnetic separation equipment has been used since the time of Thomas Edison who used a magnetic pulley for the concentration of nickel ore. These are among the most common types of magnetic separators in the world today. They can be made from either permanent or electromagnets and process either wet or dry feeds. Dry magnetic drums can be designed to perform as lifting magnets or pulleys. Magnetic drum devices have stationary magnets, whereas pulley drums rotate. Schematics for these types of devices are shown in Figs. 9 and 10.

**Grate-Type Magnets.** Grate-type magnets consist of a series of stainless steel tubes packed with ceramic magnets that are mounted in a trap perpendicular to the flow of the material. Grate magnets are used in both wet and dry processing for the



**Figure 10.** Conventional rare earth permanent magnet drum-type separator. Courtesy of Corpos, Inc.

removal of both coarse and fine tramp iron. Various types of designs such as self-cleaning grates, permanent magnets, vibrating grates, and rota-grates, have been proposed and implemented. These trap-type separation devices come in a variety of shapes and sizes and have been used in such applications as food processing, chemical processing, paper and plastic processing, and recycling. Grate magnets are most often implemented in production lines to prevent accidents or contamination from tramp iron where dry or wet pulverized material is processed.

**Lifting Magnets.** Lifting magnets are used in both continuous and batch operation depending upon the particular application. Continuous devices typically carry the unprocessed material along a conveyor belt that is located directly under the lifting magnet. The lifting magnet itself has a belt that moves over its lifting magnetic poles. The magnetic material is lifted from the first belt carrying the unprocessed material to the second belt moving over the magnetic poles of the lifting magnet. The magnetic material on the second belt is then carried away to a region of low magnetic field where it is cleaned. These devices can be either high or low intensity depending upon the design of the pole pieces. Lifting magnets are used primarily in the removal of tramp iron. The lifting magnet is typically positioned about 5 cm to 10 cm from the highest point of the conveyor transporting the unprocessed material. Typical conveyor speeds vary between 1.75 m/s and 2.5 m/s depending on the application and whether the device is self-cleaning or not.

**Plate Magnets and Magnetic Humps.** Plate magnets operate by removing tramp iron from materials moving pneumatically or by free falling gravity flow. The tramp iron is removed by attaching itself to a magnetized plate that is periodically scraped and cleaned. Chute angles of 45° are recommended, and the magnetic plate should be located as close to the feed point as possible to minimize velocity effects of the incoming stream. Plate magnets are fabricated using both permanent and electromagnets. For permanent magnet versions of this device, plate widths can extend up to 1.23 m; electromagnet versions of this device can increase plate widths to 2.85 m. The capacity of these devices varies as function of the chute angle and the size of iron particulates being removed.

**Induced Roll Separators.** Induced roll separators have been in commercial use since 1890. These devices can only handle dry granulated weakly magnetic material and are similar to drum separators with the primary difference being that the cylinder rotates in the gap of an electromagnet. An example of an induced roll separator is shown in Fig. 11. The magnetic field gradient is created by making sharply edged ridges of the surface of the cylinder or by constructing a cylinder of alternating magnetic and nonmagnetic disks. Industrial devices are built with multiple cylinders that operate in either series or parallel. By varying the configuration, these devices can be used as either concentrators or purifiers. Typical cylinder widths vary between 0.25 m and 0.75 m. Installed cost for these devices is relatively low compared with other units with similar capacity.

**Wet Drum Magnetic Separators.** Wet drum separators are used exclusively for the processing of wet feed material for the separation of strongly magnetic coarse particles. The key processing variables that determine the size and processing capacity of the device are slurry volume, percent magnetic and

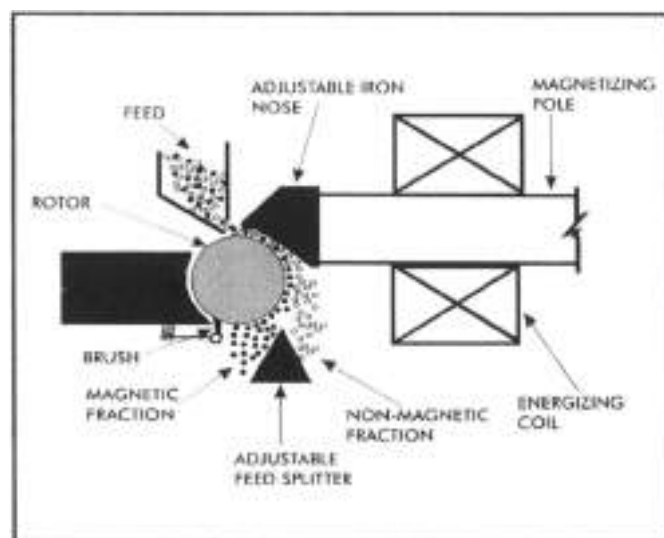


Figure 11. Conventional induced-pole separator. Courtesy of Carpa, Inc.

solids in the slurry, and the required recovery and concentration of magnetic particles. Several vendors manufacture these devices. Typical drum sizes vary from 0.623 m to 1.2 m in diameter with heights up to 3 m. Wet drum devices can come in either a single- or concurrent-type arrangement (see Fig. 12). A concurrent arrangement is where two or more drums are placed in series. Unprocessed feed material is introduced into the first drum, which is used to remove the strongly magnetic coarse particles. The processed material is then fed into the second drum for the removal of finer magnetic particles. Units with single drums can process slurry with magnetics up to 20% by weight and units with two drums can process slurry with up to 45% magnetics by weight. Recommended maximum particle size is 6 mm. Concurrent devices can process finer particles down to about 0.8 mm (20 mesh) with an optimum solids content of about 30% by weight. Wet drums separators have the advantage of being able to process material with a wide variation of particle size and throughput. The installed cost of a single wet drum-type separator can vary between \$25,000 and \$75,000 per meter of magnet width. Multiple drum cost increases in direct proportion to the number of drums required.

**Carousel Electromagnets.** One of the first high-intensity continuous-type separators was developed around 1963. This device was a rotating carousel version of the wet high-intensity magnetic separator (WHIMS) developed earlier by Jones in 1957 (18). This device consists of a cylinder subdivided into equally spaced compartments that rotate about a vertical axis. Each compartment is packed with 430 stainless steel wool. The magnetic field is generated using an electromagnet that produces a magnetic field in the axial direction perpendicular to the rotating compartments. Unprocessed feed material is introduced into the compartments in the low field region (see Fig. 13). These compartments are then rotated into the high field region. In this region, the magnetic particles are trapped in the cells, and the nonmagnetic particles are washed out. Once the compartment cell has rotated to a region of low field, the magnetic particle is then washed out with a high-velocity rinse (19).

### LTS Reciprocating Magnet System

**Background.** As early as 1975, studies were carried out on a new type of magnetic separator design that benefits from the zero resistance property of superconducting coils (20). In 1989, Carpa introduced the first commercial reciprocating magnetic separator. Unlike its LTS batch-type predecessor that cycles its magnetic field, this magnet maintains the field at a constant level and instead cycles the matrix/filter canister in and out of the active magnetic field region. This design allows for semi-continuous processing and reduces ac losses by not cycling the magnetic field. The magnet operates in what is known as the persistent mode. Below the superconducting transition temperature, a superconducting magnet can be energized and then disconnected from the power supply, and current will continue to flow without additional power input. The basic processing cycle for the reciprocating magnetic separator is shown in Fig. 14. The key feature is that while one matrix canister is processing material in the central magnetic field region, the other matrix canister is being cleaned/flushed in the low field region (21).

**Installed Systems.** In 1989, the first LTS reciprocating magnetic separations unit, used in the purification of kaolin clay, went into operation in Cornwall in the United Kingdom. This unit consists of a NbTi conductor winding operating in a bath of liquid helium. It has a warm bore diameter of about 0.28 m with a maximum central field of 4 T and can process between 2 tons and 5 tons of kaolin per hour. In 1992, the second LTS reciprocating magnet system was installed in southern Germany for the purification of kaolin clay. This unit has a 0.26 m warm bore diameter with a maximum central field of 5 T and can process up to 5 tons of kaolin per hour. To date, 15 more industrial-scale LTS reciprocating magnetic separators have been installed worldwide. Reciprocating systems with warm bore diameters up to 1 m and central fields of 5 T are presently in operation (see Fig. 15). Ten smaller-diameter LTS reciprocating units, operating in research laboratories and pilot-scale production lines, have also been installed. One might expect that the complexity of superconducting technology would restrict the commercial viability of these units to developed and industrialized areas. In reality, the simplicity and reliability of the low loss cryogen technology coupled with the reciprocating canister principle has enabled operation of these units in Manaus and Kin Chaprum (21), which are remote areas of the Amazon rain forest.

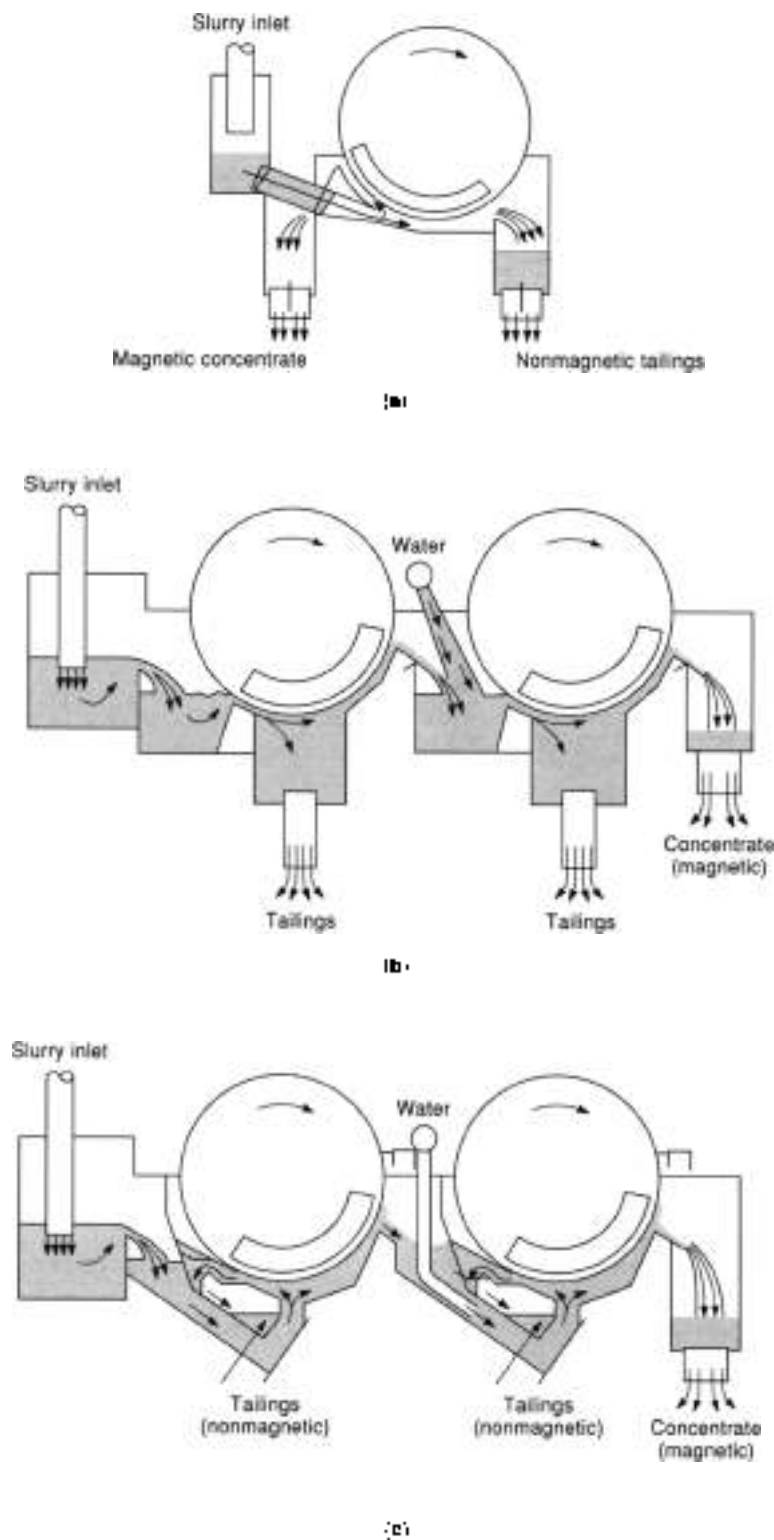
**HIS Reciprocating Magnetic Separator.** Similar to its batch-type counterpart, it is possible to fabricate a reciprocating magnetic separation system using HTS coils. The commercial advantage of such a system would be the elimination of the use of liquid and gaseous cryogens even for very large-scale systems. It is too early to speculate on the commercial viability and technical feasibility of such a device.

## APPLICATIONS OF MAGNETIC SEPARATION

### Background

There are several commercially available mineral separation technologies that use specific gravity, magnetic separation, electrostatic separation, and column flotation, among others. All of these techniques exploit various discernible properties



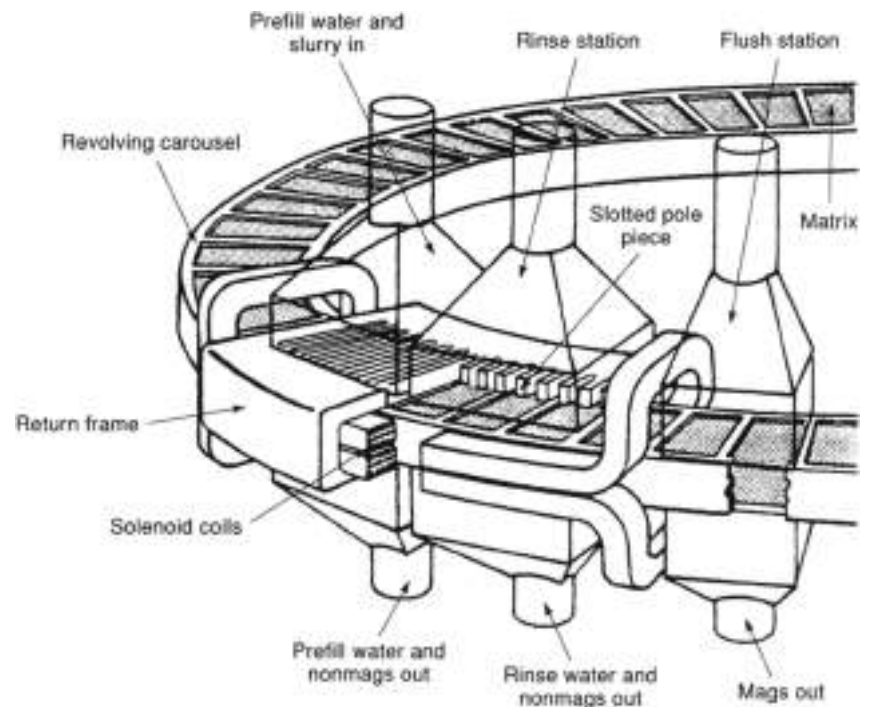


**Figure 12.** Wet drum magnetic separator arrangements: (a) counter-rotating-type, (b) concurrent double-drum-type, and (c) countercurrent double-drum-type. Courtesy of Swedala Industries, AR Pumps and Process.

among mixtures of minerals. A comparison of different separation methods as a function of particle size range is shown in Fig. 16. Each separation technology has its own particular strength; however, typically a combination of techniques provides the best industrial minerals separation processes. Commercial magnetic separators come in a variety of shapes and sizes depending upon the required application. To select the most appropriate magnetic separator for a specific application

requires an evaluation of several variables including the type of material being processed, wet or dry processing, particle size, magnetic characteristics, and processing rate.

**Kaolin Processing.** Kaolin is a naturally occurring white clay consisting of microscopic platelets of aluminum silicate. The United States is the largest producer and exporter of kaolin in the world, with over 10 million tons valued at over \$1.3 billion



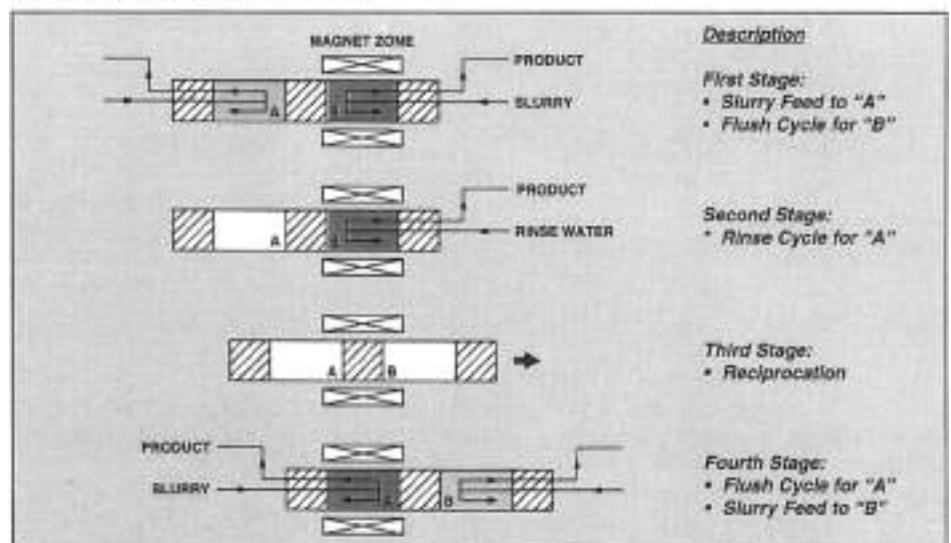
**Figure 13.** A schematic representation of a carousel type magnetic separator. Courtesy of Svedala Industries AB Pumps and Process.

produced in 1993. Georgia generates 80% of the tonnage and 90% of the value of kaolin in the United States (16). Kaolin is used in the manufacture of fine porcelain and as a base filler in the manufacture of high-grade paper.

One of the most widely recognized industrial applications of high gradient magnetic separation is in the purification of kaolin clay. Magnetic separation has been used in the kaolin industry for over 25 years. The primary benefit is increased whiteness or brightness of the kaolin product in paper or ceramic applications. Magnetic separation has offered additional benefits to kaolin processing such as improved viscosity or rheology.

Magnetic separation is used primarily to remove small paramagnetic impurities of titanium and iron compounds that discolor the clay. The impurities content typically makes up between 2% and 5% of the weight of kaolin. The quality of the processed clay is determined by the resulting brightness. The industrial association that establishes the brightness standards for the kaolin industry is the Technical Association of the Pulp and Paper Industry (TAPPI). Brightness is typically measured using a reflectance technique, and the results are compared with industry standards (e.g., TAPPI T 646 am-94 for pulverized material with 45  $\mu$ m geometry) (22). The improvement

#### CRYOFILTER OPERATING CYCLE



**Figure 14.** A schematic representation of a reciprocating-type operating cycle. Courtesy of Carper, Inc.



**Figure 15.** A 5 T low-temperature superconducting reciprocating magnet with a 0.5 m bore installed in Georgia. This magnet can process 20 tons of kaolin clay per hour. Courtesy of Caroco, Inc.

in brightness after magnetic separation depends upon many processing variables such as the initial quality of the unprocessed clay, magnetic field strength, mesh size and density, and retention time. Typical improvement between 1 and 5 brightness units can be expected as a result of processing kaolin in a magnetic field. Essentially all kaolin is processed through magnetic separators. The most common separator size in current use can process kaolin slurry at typically 70 m<sup>3</sup>/h to 120 m<sup>3</sup>/h, which translates to a production rate of approximately 23 metric ton/h to 45 metric ton/h (dry basis). Most of the magnetic separation devices used in the purification of kaolin clay are resistive magnets that operate in batch mode. However, both batch-type and reciprocating-type superconducting magnets are quickly displacing their resistive counterparts.

**Titanium Dioxide.** Rutile and anatase are naturally occurring ores containing microscopic particles of titanium dioxide (TiO<sub>2</sub>), which is the most effective white pigment used in the paint, paper, and plastics industries. It is widely used because it efficiently refracts visible light imparting whiteness, brightness, and opacity when incorporated in a huge variety of fabricated products. Titanium dioxide is chemically inert, insoluble, and thermally stable under the harshest processing conditions. Titanium dioxide is produced commercially in two crystal forms—anatase and rutile. Rutile-based pigments are preferred because they refract light more efficiently, and are more stable and less photoreactive. Worldwide, over 2.2 million metric tons of TiO<sub>2</sub> are produced annually. The United States is both the largest exporter and consumer of TiO<sub>2</sub>; however, the United States is also a net exporter of TiO<sub>2</sub>. Asia is the fastest growing segment for both production and consumption of TiO<sub>2</sub>. DuPont is the largest producer of TiO<sub>2</sub> in the world. A

typical method for the separation of titanium dioxide is shown in Fig. 17. Magnetic separation is used in both the front-end screening process as well as the final processing of the "fines" of the product. As shown in Fig. 17, both ilmenite (FeTiO<sub>3</sub>) and rutile (TiO<sub>2</sub>) are electrically conductive; however, successful separation is possible because ilmenite is strongly paramagnetic and rutile is very weakly paramagnetic.

### Chemical Processing

Reports of substantial improvements in material properties of polymers, particularly the spatial orientation of polymers, during the polymerization process by the application of a magnetic field, are intriguing to chemical and polymer manufacturers even though the reports do not represent typical magnetic separation techniques. For instance, Dow Chemical Company has reported an enhancement in the orientation of synthetic fibers extruded in the presence of a strong magnetic field (23). Significant increases in epoxy tensile strength of 80%, microhardness of 25%, and bonding strength of 35% have been reported when materials of *N*-glycidyl ether are extruded in applied magnetic fields (24). In addition, increases in electrical conductivity and polymerization rates have been reported for conducting polymers of polyacetylene processed in magnetic fields. This has the implication of improved product quality as well as productivity increases. However, the benefits and feasibility of controlling spatial orientation on a molecular level while conducting high volume manufacturing are presently unknown.

### Water Treatment

Various studies have been performed demonstrating the successful treatment and clean up of wastewater using magnetic

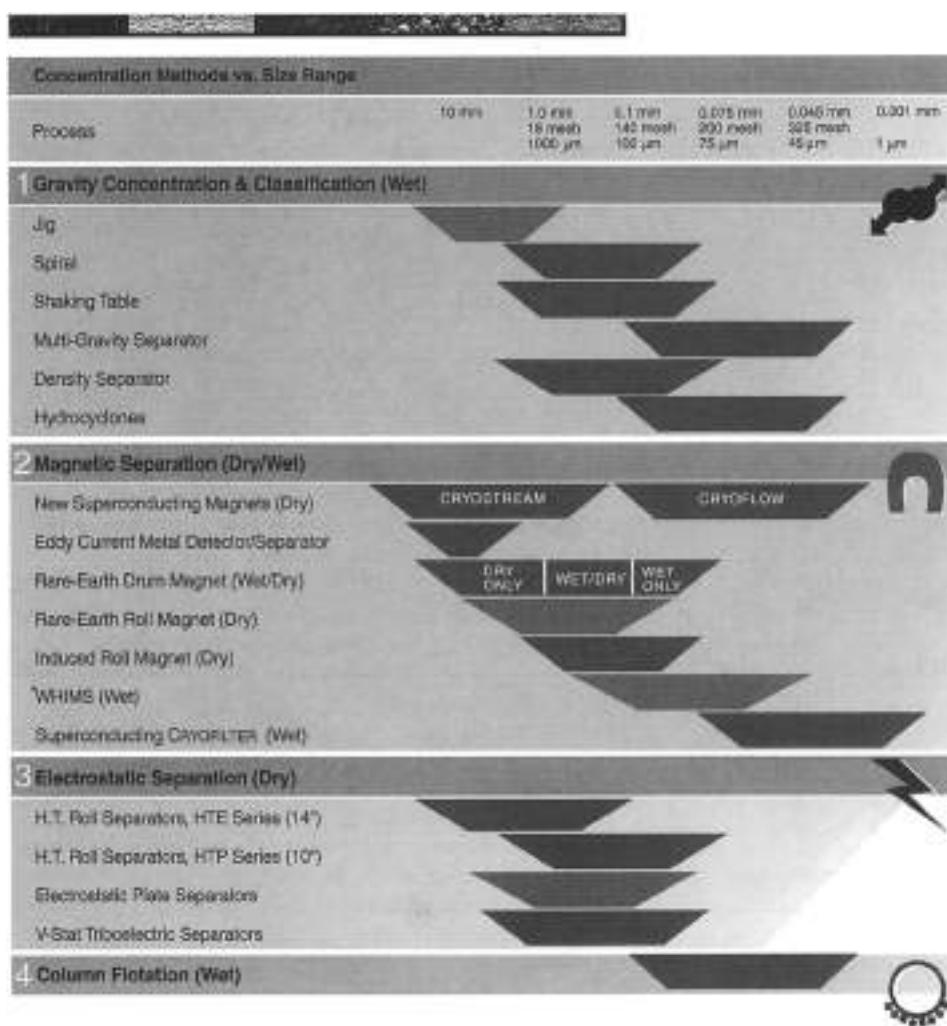


Figure 16. A comparison of separating methods as a function of particle size. Courtesy of Carcopin, Inc.

separation techniques. One study performed by Harvard University and Sala Magnetics showed that IHMS was highly effective in the removal of suspended solids, turbidity, coliform, color, viruses, and heavy metals from wastewater that had been treated with sulfidic magnetite (25). In this study, ferromagnetic magnetite ( $\text{Fe}_3\text{O}_4$ ) was chemically combined with traditionally nonmagnetic materials (i.e., pollutants in water). Using this "seeding" technique with magnetite and aluminum sulfate as a flocculent, coliform bacteria and other solids were successfully removed. For example, Table 3 shows the feed and magnetically processed data for surface and bottom samples from the Charles River and the Deer Island Sewage treatment plant in Boston, MA (10). Dissolved ions can also be separated from water systems in a similar manner. Another study (26) involved the evaluation and optimization of magnetic filters on boiler water. In this study, it was shown that magnetic separation was a highly effective technique for the removal of heavy metal contaminants of iron and copper oxide from boiler water. Results indicated that an expanded metal filter provided the highest efficiency, coupled with the lowest clogging incidences. Flow velocities up to 800 liters per minute were studied in magnetic fields of 0.5 T. Particle capture up to 98% was achieved under optimal conditions.

In terms of commercial activity, two companies in the United States, Aqua Magnetics International (27) and Fluid-Tech International Corporation (28) sell magnetic equipment for the degrading of boiler water as well as conventional filtration. Note that extensive research has been conducted in the former Soviet Union on the magnetic treatment of wastewater and that several thousand installations use the magnetic treatment of water from steam boilers.

#### Solid Waste Remediation

Magnetic separation techniques have received new interest in two specific areas of solid waste clean up. The first is in the area of ferrous metallic recovery (for recycle) from municipal solid waste. The Bureau of Mines reports that approximately 8% of municipal solid waste is comprised of ferromagnetic materials. Each year up to 11 million tons of ferrous metallics are discarded domestically. Using magnetic separation techniques, the present recovery rate is about 180,000 tons per year (34).

Another area that has generated recent interest is the removal of uranium compounds from contaminated soil. Successful removal of uranium compounds from contaminated soil

**Table 3. Effect of Magnetic Treatment on Water Quality**

|                                | Charles River   |         |                |         | Deer Island       |         |
|--------------------------------|-----------------|---------|----------------|---------|-------------------|---------|
|                                | Surface Control | Treated | Bottom Control | Treated | Control           | Treated |
| Coliform Bacteria (per 100 ml) | 16,000          | 0       | 16,000         | 300     | $2.9 \times 10^4$ | 18,000  |
| Turbidity (JTU units)          | 20              | 2       | 1,700          | 1       | 50                | 0       |
| Color (color units)            | 105             | 3       | 3,700          | 1       | 150               | 20      |
| Suspended Solids (mg/L)        | 7               | 5       | 890            | 7       | 45                | 0       |

using magnetic separation techniques in a superconducting magnetic separator has been reported (29). This research has been limited to laboratory-scale quantities. The economic feasibility of using HGMS in large-scale solid waste cleanup (e.g., a government-qualified superfund site) has not been determined.

### Coal Purification

Coal is the largest fossil fuel resource in the United States, with approximately 180 billion tons of currently recoverable reserves (30). One of the major difficulties with widespread coal use is the pollution problem, particularly with sulfur oxide emissions. Sulfur in coal occurs in three forms: pyritic, sulfate, and organic. Pyritic sulfur ( $\text{FeS}_2$ ) accounts for 40% to 80% of the total content of most coals. In the bituminous coals from Illinois, for example, the mean total sulfur content is 3.57% of which 2.06% is pyritic sulfur. Pyrite occurs in coal as discrete particles varying in size from tens of micrometers to the sub-micron level.

Coal is diamagnetic, and pyrite is paramagnetic in nature. If coal is crushed and pulverized fine enough to liberate the

pyrite, then separation by magnetic means is possible. High gradient and open gradient separators have been shown to remove between 80% and 90% of the pyritic sulfur and between 20% and 50% of the ash. The economic feasibility of magnetic separation of sulfur from coal is however highly questionable, compared with conventional processes.

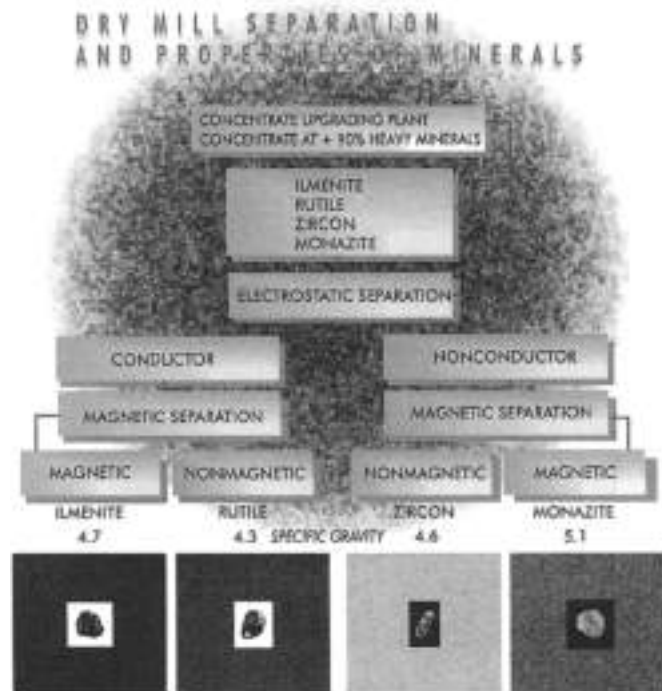
### SUMMARY

Magnetic separation has been used in the processing of materials for more than one hundred years. Magnetic separation has a variety of modern uses ranging from the simple removal of tramp iron to the highly sophisticated removal of weakly paramagnetic minerals from clays. Several different types of magnetic separation devices are available. The choice and type of magnetic separator best suited for a particular application depends on several variables such as wet or dry processing, magnetic susceptibility of the materials present, magnetic fraction of the feed materials, particle size, and processing rate. The introduction of high gradient magnetic separation greatly expanded the role that these devices play in the mining of minerals. It has been estimated that the introduction of HGMS has nearly doubled the worldwide useful reserves of kaolin clay, by allowing the mining of lower grade material. It is clear that in large-scale high gradient magnetic separators, low temperature superconducting technology is displacing conventional water-cooled copper magnets. With the recent discovery of high-temperature superconductivity, it remains to be seen if these new ceramic-oxide superconductors will replace the traditional intermetallic low-temperature superconductors.

Magnetic separation techniques are being explored in many nontraditional applications such as wastewater cleanup, the removal of pyritic sulfur and ash from pulverized coal, chemical processing, and the removal of uranium-oxide compounds from contaminated soil. The use of magnetizing techniques on nonmagnetic materials may open an entirely new area of benefits and applications. As more research and development is being performed, magnetic separation techniques continue to find new areas of potential environmental and commercial benefit.

### BIBLIOGRAPHY

1. F. Knoll, D. Green, and J. Maloney (eds.), Solid-solid operations and equipment, *Perry's Chemical Engineers' Handbook*, 17th ed., New York: McGraw-Hill, 1987, pp. 19-40-19-49.
2. J. Jannirelli and S. Malghan, Eds.: Paramagnetic separation in ultrafine industrial minerals and coal, *Ultrafine Grinding and Separation of Industrial Minerals*, New York: American Institute of



**Figure 17.** Dry mill separation process of rutile and ilmenite. Courtesy of Carpen, Inc.

- Mining, Metallurgical and Petroleum Engineers Inc., 1983, p. 165.
3. H. Kalm, J. Oberbauer, and Th. Kelland, *Sci. Am.*, 283(5): 44-54, 1975.
  4. K. M. Bezzuth, *Aberrationism*, New York: IEEE Press, 1990.
  5. D. Holliday and R. Resnick, *Fundamentals of Physics*, 2nd ed., New York: Wiley, 1970.
  6. J. D. Jackson, *Classical Electrodynamics*, 3rd ed., New York: Wiley, 1975.
  7. F. C. Moon, *Magneto Solid Mechanics*, New York: Wiley, 1984.
  8. S. G. Frantz, US Patent 2,974,086, March 16, 1937.
  9. Z. J. Szekly and J. V. Manerini, *IEEE Trans. Magn.*, 12(5): 474-479, 1976.
  10. S. Foner and B. Schwartz (eds.), *Superconducting Machines and Devices*, New York: Plenum Press, 1974, pp. 581-594.
  11. J. Iannicelli, Paramagnetic separation in ultrasonic industrial minerals and coal, presented at the *AIMME/SME Meeting*, Atlanta, GA, 1993. Reprinted by Aquafine Corporation, Brunswick, GA.
  12. J. Iannicelli, *Clays and Clay Minerals*, 20, Great Britain: Pergamon Press, 1976; pp. 64-68.
  13. Aquafine Corporation, *Products Catalog*, Brunswick, GA, 1996.
  14. Eriez Magnetics, *Products Catalog*, Erie, PA, 1997.
  15. J. G. Bednorz and K. A. Mueller, *Z. Phys. B: Condens. Matter*, 64(2): 189-183, 1986.
  16. J. Iannicelli et al., *IEEE Trans. Appl. Supercond.*, 7(2): 1061-1064, 1997.
  17. C. M. Rey and K. Sato, Conceptual design of a 0.8 m diameter HSE magnetic separator presented at the Applied Superconductivity Conference, Palm Desert, CA, Sept. 1998.
  18. G. H. Jones, US Patent 3,846,471.011, October 10, 1967.
  19. J. Iannicelli, *IEEE Trans. Magn.*, 12(5): 1976.
  20. Z. J. Szekly, *IEEE Trans. Magn.*, 11: 1594-1606, 1975.
  21. Carpac, *Products Catalog*, Jacksonville, FL, 1988.
  22. TAPPI Test Methods, Atlanta: TAPPI Press, 1987.
  23. H. J. Schneider-Muntau (ed.), *High Magnetic Fields*, River Edge, NJ: World Scientific, 1997, p. 31.
  24. S. Dale, S. Wolf, and T. Schnieder, *Energy Applications of High-Temperature Superconductivity*, 2, 1976.
  25. R. Mitchell and D. Allen, *Industrial Applications of Magnetic Separators*, IEEE Catalog No. 79CH1447-2, 1979, p. 142.
  26. J. Iannicelli and T. Webster, Electric Power Research Institute Grant No. NP 3273, research project S108-1, 1983.
  27. AQUA Magnetics International, *Products Catalog*, Safety Harbor, FL, 1998.
  28. Fluid-Tech International, *Products Catalog*, Ft. Wayne, IN, 1988.
  29. L. A. Wurl, et al., Magnetic separation for nuclear material and surveillance, to be published in *Emerging Technologies in Hazardous Waste Management*, Proc. 210th American Chemical Society Conference, Boston, MA, 1996.
  30. J. Iannicelli and H. Murray, National Science Foundation Grant No. NSF-GT 44219 and NSF-ISP 74-21921, 1969.

CHRISTOPHER M. REY  
DuPont Superconductivity

## MAGNETIC SOURCE IMAGING

More than 200 years ago, it was discovered that biological processes are accompanied by electrical currents. Since then, measurements of bioelectric signals have become widespread

procedures of great importance in both biophysical research and medical applications in clinical use. These studies include, for example, measurements of electric potential differences arising from human heart (the electrocardiogram (ECG)), brain (the electroencephalogram (EEG)), and other organs.

The same bioelectric activity that generates electrical potentials also generates weak magnetic fields. Because these biomagnetic fields measured outside the body are extremely low in magnitude ( $\sim 10$  fT to 100 pT), it was not until 1963 that the first successful detection of the magnetic field arising from human heart was performed (1). This was the beginning of magnetocardiography (MCG). Magnetoencephalography (MEG) was introduced in 1965 when magnetic signals due to the spontaneous  $\alpha$ -rhythm in the brain were detected (2). However, it was only after the development of ultrasensitive superconducting quantum interference device (SQUID) detectors in the beginning of the 1970s (3) that easier detection of biomagnetic signals became possible. In addition to the magnetocardiogram and the magnetoencephalogram, various biomagnetic fields arising from the body have been studied since then.

Biomagnetic measurements offer information that is very difficult to obtain with other imaging methods (4-8). MEG and MCG are generated by the electric currents in neurons or myocardial cells, and therefore the measurements provide direct real-time functional information about the brain or the heart, respectively. The time scale of the detectable signals ranges from fractions of a millisecond to several seconds or even longer periods. The biomagnetic measurements are totally noninvasive and the body is not exposed to radiation or high magnetic fields. Mapping of biomagnetic signals at several locations simultaneously is easy and fast to perform with multichannel systems.

The metabolic processes associated with the neural or myocardial activity can be studied with positron emission tomography (PET), but the imaging times are several minutes, and the spatial resolution is about 5 mm. Better spatial resolution is obtained from functional magnetic resonance imaging.

Estimation of bioelectric current sources in the body from biomagnetic measurements is often called magnetic source imaging (MSI). To relate the functional information provided by MSI to the underlying individual anatomy, other imaging methods are employed, such as magnetic resonance imaging (MRI), computer tomography (CT), and X ray. In this article, we focus on MEG and MCG, followed by a brief discussion of other fields of biomagnetism. Furthermore, instead of a comprehensive review of MEG and MCG applications we provide a few illustrative examples of recent MSI studies.

## MEG AND MCG STUDIES

During recent years, MEG and MCG have attained increasing interest. The ability of these methods to locate current sources combined with precise timing of events is valuable both in basic research and in clinical studies.

One common type of an MEG experiment is to record the magnetic field associated with a sensory stimulus or a movement. Since these fields are usually masked by the ongoing background activity, signal averaging is routinely employed to reveal the interesting signal component. Recordings of neuromagnetic fields have provided a wealth of new information about the organization of primary cortical areas (9).

Sensor arrays covering the whole head have made studies of complicated phenomena involving simultaneous or sequential processing in multiple cortical regions feasible. Because MEG is a unique tool to study information processing in healthy humans, several language-related studies have recently been conducted (10–12).

It is also possible to record the ongoing rhythmic spontaneous brain activity in real time and follow its changes under different conditions (13,14). In addition to the well-known 10 Hz  $\alpha$ -rhythm originating in the vision-related cortical areas, similar spontaneous signals occur, for example, in the somatosensory system. MEG measurements have provided new information about both the generation sites of these rhythmic activities and their functional significance (15).

Both evoked responses and spontaneous activity recordings can be utilized in clinical studies (16). For example, the locations of the somatosensory and motor cortices deduced from evoked MEG signals can be superimposed on three-dimensional surface reconstructions of the brain, computed from MRI data. The resulting individual functional map can be a valuable aid in planning neurosurgical operations. Encouraging results have also been obtained in locating epileptic foci in candidates for epileptic surgery.

High-resolution MCG recordings have been applied both in basic cardiac research and in clinical studies. In the first MCG studies in the 1970s and 1980s, only single-channel devices were available, which limited the use of MCG to subjects and patients with normal sinus rhythm. Introduction of multichannel recording systems in the 1990s made the technique more suitable for routine clinical studies and for analysis of beat-to-beat variations. Currently, MCG is being used at some hospitals to test and further develop its clinical use.

Multichannel MCG studies are particularly promising in two clinically important problems: (1) in locating noninvasively abnormal cardiac activity critical for the arousal of life-threatening arrhythmias and (2) in evaluating the risk of such arrhythmias in different cardiac pathologies, especially after myocardial infarction. Successful MCG results have been reported, for example, in locating abnormal ventricular preexcitation sites associated with the Wolff-Parkinson-White syndrome, the origin of ventricular extrasystolic beats, and the origin of focal atrial tachycardias (17–21).

Furthermore, MCG localization accuracy has been tested with artificial sources, such as a pacing catheter in the heart (20,22). The localization accuracy reported so far, ranging from about 5 mm to 25 mm, is sufficient to provide valuable information for preablation evaluation of the patients. In addition to localization studies, MCG has been applied to retrospective identification of patients prone to malignant arrhythmias with about 90% sensitivity and specificity (23).

## INSTRUMENTATION

### Detection of Neuromagnetic Fields

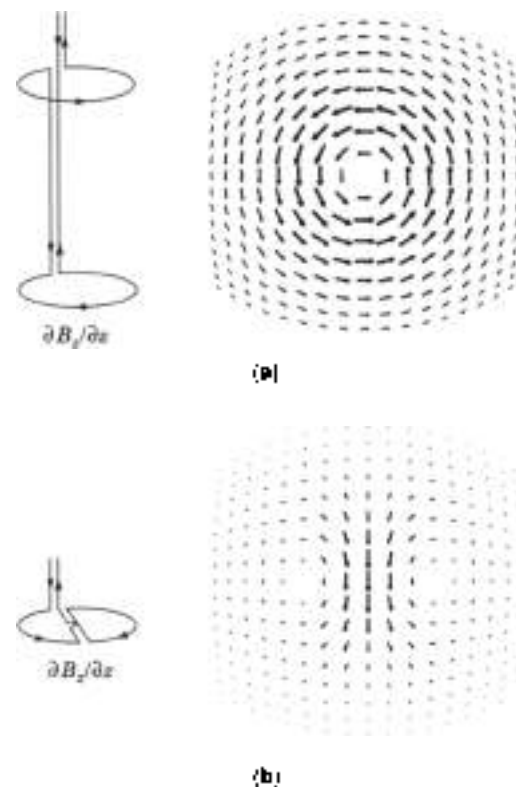
The detector that offers the best sensitivity for the measurement of these tiny fields is the SQUID (24,25), which is a superconducting ring, interrupted by one or two Josephson junctions (26). These weak links limit the flow of the supercurrent, which is characterized by the maximum critical current  $I_c$  that can be sustained without loss of superconductivity. Dried-current (dri-

SQUIDs, with two junctions, are preferred because the noise level is lower in them than in radio-frequency (RF) SQUIDs (27–29).

The magnetic signals from the body are extremely weak compared with ambient magnetic field variations (5). Thus, rejection of outside disturbances is of utmost importance. Significant magnetic noise is caused, for example, by fluctuations in the earth's geomagnetic field, by moving vehicles and elevators, and by the omnipresent powerline fields.

For rejection of external disturbances, biomagnetic measurements are usually performed in a magnetically shielded room. To make such an enclosure, four different methods exist: Ferromagnetic shielding, eddy-current shielding, active compensation, and the recently introduced high- $T_c$  superconducting shielding. Many experimental rooms have been built for biomagnetic measurements utilizing combinations of these techniques (34–36). Commercially available rooms utilized in biomagnetic measurements usually employ two layers of aluminum and ferromagnetic shielding, possibly combined with active compensation. The inside floor area is usually 3 m by 4 m, and the height around 2.5 m.

In addition, the sensitivity of the SQUID measuring system to external magnetic noise can be greatly reduced by the proper design of the flux transformer, a device normally used for bringing the magnetic signal to the SQUID. For example, an axial first-order gradiometer consists of a pickup (lower) coil and a compensation coil with identical effective area and connected in series but wound in opposition [see Fig. 1(a)]. This system of



**Figure 1.** Left: Coil configurations for (a) an axial and (b) a planar gradiometer. Right: The corresponding sensitivity patterns (lead fields) (see section entitled "Distributed Source Models"). The plots show the lead fields on a spherical surface. The gradiometer coil is located above the center of each pattern. The direction and size of the arrows indicate the magnitude and direction of the lead field at the center of the arrow.

coils is insensitive to a spatially uniform background field, but it responds to inhomogeneities. Therefore, a source near the lower coil, which will cause a much greater field in the pickup loop than in the more remote compensation coil, will thus produce a net output.

Most biomagnetic measurements have been performed with axial gradiometers. However, the off-diagonal planar configuration of Fig. 1(b) has some advantages over axial coils. The double-D construction (2K) is compact in size, and it can be fabricated easily with thin-film techniques. The locating accuracies of planar and axial gradiometer arrays are essentially the same for superficial sources (31–33). The spatial sensitivity pattern, lead field, of off-diagonal gradiometers is narrower and shallower than that of axial gradiometers. These sensors thus collect their signals from a more restricted area near the sources of interest, and there is less overlap between lead fields of adjacent sensors in a multichannel array.

Nevertheless, distant sources can often be detected more easily with axial gradiometer or magnetometer sensors. Therefore, many experimental and commercial systems include these coil configurations, possibly in combination with planar gradiometers.

#### Multichannel Magnetometers

The first biomagnetic measurements were performed with single-channel instruments. However, reliable localization of current sources requires mapping in several locations, and this is time-consuming with only one channel. Besides, unique spatial features present in, for example, brain rhythms cannot be

studied. Fortunately, during the past 15 years, multichannel SQUID systems for biomagnetic measurements have been developed to provide reliable commercial products. A detailed account of this development can be found in Ref. 5.

A state-of-the-art multichannel MEG system comprises more than 100 channels in a helmet-shaped array to record the magnetic field distribution across the brain simultaneously. The latest MEG systems contain 60 to 80 detectors in a flat or slightly curved array to cover an area about 30 cm in diameter over the subject's chest or back. The dewar containing the sensors is attached to a gantry, which allows easy positioning of the dewar above the subject's head or chest. The position of the dewar with respect to the subject's head or torso is typically determined by measuring the magnetic field arising from an ac current fed into small marker coils attached to the skin (37,38) and by calculating their locations with respect to the sensor array. The locations of the marker coils with respect to an anatomical frame of reference are determined before the biomagnetic measurement by a three-dimensional digitizer.

As an example of an MEG installation we describe the Neurmag 122 system (Neurmag Ltd., Helsinki, Finland) (33). This dewar employs planar first-order two-gradiometer units to measure the two off-diagonal derivatives,  $\partial B_x/\partial x$  and  $\partial B_y/\partial y$ , of  $B_z$ , the field component normal to the dewar location at 63 locations. The thin-film pickup coils are deposited on  $28 \times 28 \text{ mm}^2$  silicon chips; they are connected to 122 dc SQUIDs attached to the coil chip. The separation between two double-sensor units is about 43 mm. The system is depicted in Fig. 2.



Figure 2. Left: The Neurmag 122<sup>TM</sup> MEG system. Right: The 67-channel MEG system. (Photographs courtesy of Neurmag, Ltd.)



The MCG system from the same company shown in Fig. 2 comprises 67 channels arranged on a slightly curved surface with diameter about 30 cm. The magnetic-field component ( $B_z$ ) perpendicular to the sensor array surface is sensed by seven large-coil axial gradiometers and 30 two-channel planar gradiometer units identical to those described in the previous paragraph.

## GENERATION OF BIOELECTROMAGNETIC FIELDS

### Cellular Sources

To interpret the measured signals, one has to understand how electric and magnetic fields are generated by biological tissue. In this article, we consider biomagnetic signals generated by the electric currents in excitable tissue. These magnetic fields are linked to bioelectric potentials, and it is useful to consider both the magnetic field and the electric potential together.

Living cells sustain a potential difference between intra- and extracellular media. In a static situation, most cells are, as seen externally, electrically and magnetically silent. Excitable cells can produce electric surface potentials and external magnetic fields, which can be detected from outside.

### The Quasi-Static Approximation

The total electric current density in the body,  $\mathbf{J}$ , is time-dependent, and the electric field ( $\mathbf{E}$ ) and the magnetic field ( $\mathbf{B}$ ) produced by  $\mathbf{J}$  can be found from Maxwell's equations. However, the variations in time are relatively slow (below 1 kHz) (8,39), which allows treatment of the sources and the fields in a quasi-static approximation. This means that inductive, capacitive, and displacement effects can be neglected. In the quasi-static approximation Maxwell's equations thus read:

$$\nabla \cdot \mathbf{E} = \rho/\epsilon_0 \quad (1)$$

$$\nabla \times \mathbf{E} = 0 \quad (2)$$

$$\nabla \cdot \mathbf{B} = 0 \quad (3)$$

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} \quad (4)$$

where  $\mu_0$  and  $\epsilon_0$  are the magnetic permeability and electric permittivity of the vacuum, respectively.

### Primary Current

It is useful to divide the total current density in the body,  $\mathbf{J}(\mathbf{r}, t)$ , into two components. The passive volume or return current is proportional to the conductivity  $\sigma(\mathbf{r}, t)$  and the electric field  $\mathbf{E}$ :

$$\mathbf{J}(\mathbf{r}, t) = \sigma(\mathbf{r}, t) \mathbf{E}(\mathbf{r}, t) \quad (5)$$

$\mathbf{J}^p$  is the result of the macroscopic electric field on charge carriers in the conducting medium. Everything else is the primary current  $\mathbf{J}^p$ :

$$\mathbf{J}(\mathbf{r}, t) = \mathbf{J}^p(\mathbf{r}, t) + \sigma(\mathbf{r}, t) \mathbf{E}(\mathbf{r}, t) \quad (6)$$

This definition would be meaningless without reference to the length scale. Here  $\sigma(\mathbf{r}, t)$  is the macroscopic conductivity; cellular-level details are left without explicit attention. The division in Eq. (6) is illustrative in that neural or cardiac activity gives

rise to primary current mainly inside or in the vicinity of a cell, whereas the volume current flows passively everywhere in the medium.

It should be emphasized that  $\mathbf{J}^p$  is to be considered the driving "battery" in the macroscopic conductor; although the conversion of chemical gradients to current is due to diffusion, the primary current is largely determined by the cellular-level details of conductivity. In particular, the membranes, being good electrical insulators, guide the flow of both intracellular and extracellular currents.

If the events are considered on a cellular level, it is customary to speak about the impressed rather than primary current (39).

### Neurons

Signals propagate in the brain along nerve fibers called axons as a series of action potentials. During an action potential, the primary current can be approximated by a pair of current dipoles corresponding to a local depolarization of the cell membrane, followed by repolarization. This source moves along the axon as the activation propagates. Although the model is a simplified one, the experimental magnetic findings are in reasonable agreement with this concept (40,41).

The axons connect to other neurons through synapses. In a synapse, transmitter molecules are released to the synaptic cleft and attach to the receptors on the postsynaptic cells. As a result, the ionic permeabilities of the postsynaptic membrane are modified and a postsynaptic potential is generated. The postsynaptic current can be adequately described by a single-current dipole.

The dipolar field produced by the postsynaptic current flow falls off with distance more slowly than the quadrupolar field associated with the action potentials. Furthermore, the postsynaptic currents last tens of milliseconds, whereas the duration of a typical action potential is only about 1 ms. On this basis, it is believed that the electromagnetic signals observed outside and on the surface of the head are largely due to the synaptic current flow.

The two principal groups of neurons on the surface layer of the brain, the cortex, are the pyramidal and the stellate cells. The former are relatively large; their apical dendrites from above reach out parallel to each other, so that they tend to be perpendicular to the cortical surface. Since neurons guide the current flow, the resultant direction of the electrical current flowing in the dendrites is also perpendicular to the cortical sheet of gray matter.

### Myocardium

In heart tissue, there are three main types of cells: pacemaker cells in the sinus and atrioventricular nodes, cells specialized for rapid conduction of the excitation along the bundle of His and Purkinje network, and, finally, muscle cells that perform mechanical work. Cardiac muscle consists of a large number of individual cells, each about 15  $\mu\text{m}$  in diameter and 100  $\mu\text{m}$  long. The intracellular spaces of adjacent muscle cells are interconnected, which makes the three-dimensional structure very complex.

An action potential in myocardial cells lasts 300 ms to 400 ms, which is over 100 times longer than a typical neural action potential. Provided that we observe a single myocardial cell at some distance from the membrane, the depolarization and

repolarization can be modeled, respectively, by an equivalent depolarization and repolarization dipole.

Ventricular depolarization or repolarization propagates as about 1 mm thick wavefronts in the heart. A commonly used model to describe such propagating fronts is a uniform double layer (39). It consists of dipoles with equivalent strengths (assuming a constant dipole density), oriented perpendicular to the wavefront. The model is more suitable than a single current dipole in characterizing an excitation taking place simultaneously in a spatially large region, but it cannot account for possible holes in the wavefront (e.g., necrotic tissue). In addition, the classical concept of a uniform double layer is not valid if the anisotropic nature of myocardial tissue is to be included.

### Calculation of the Bioelectromagnetic Fields

In the quasi-static approximation, the electric potential  $\phi$  obeys Poisson's equation,

$$\nabla \cdot (\sigma \nabla \phi) = \nabla \cdot \mathcal{J}^P \quad (41)$$

while the magnetic field due to the total current density,  $\mathcal{J}$ , is obtained from the Ampère–Laplace law:

$$\mathcal{B}(r) = \frac{\mu_0}{4\pi} \int_V \frac{\mathcal{J}(r') \times R}{R^3} dV' \quad (42)$$

where the integration is performed over a volume  $V$  containing all active sources,  $r' \in V$ ,  $R = r - r'$ .

It can be shown that the volume currents in an infinite homogeneous volume conductor give no contribution to the electric potential or the magnetic field, which are solely due to the primary currents,  $\mathcal{J}^P$  (42).

Next, we assume that the body consists of homogeneous subvolumes  $v_k$ ,  $k = 1, 2, \dots, M$ , bounded by the surfaces  $S_k$ . The electrical conductivity within  $v_k$  is constant,  $\sigma_k$ . Usually, the body is surrounded by air, and thus the conductivity outside the body surface is zero. In this case, the surface potential,  $\phi_S$ , can be obtained from an integral equation (43)

$$(\sigma_k' + \sigma_k'')\phi_S(r) = 2\sigma_k'\phi_\infty(r) + \frac{1}{2\pi} \sum_{k=1}^M (\sigma_k' - \sigma_k'') \int_{S_k} \phi_\infty dS_k \frac{R}{R^3} \quad (43)$$

where  $\sigma_k'$  is the conductivity at the source location,  $\sigma_k''$  is the conductivity inside and  $\sigma_k''$  is the conductivity outside the surface  $S_k$ , and  $dS_k$  is the surface element vector perpendicular to the boundary. The term  $\phi_\infty$  denotes the electric potential in an infinite homogeneous medium (in the absence of the boundaries  $S_k$ ), and the surface integral accounts for the contribution of the conductivity change on the boundary  $S_k$ .

The external magnetic field is then evaluated by substituting the total current density  $\mathcal{J}$ , into Eq. (42). It can be shown (44) that the result can be transformed to the form

$$\mathcal{B}(r) = \mathcal{B}_\infty(r) + \frac{\mu_0}{4\pi} \sum_{k=1}^M (\sigma_k' - \sigma_k'') \int_{S_k} \phi_\infty dS_k \times \frac{R}{R^3} \quad (44)$$

where the term  $\mathcal{B}_\infty$  is the magnetic field in the absence of the boundaries,  $S_k$ . Again, the surface integral accounts for the contribution of the conductivity change on the boundary  $S_k$ .

**Analytic Solutions.** Analytic solutions of Eqs. (39) and (40) exist only in a few simple symmetric geometries. If we approximate the head or the torso by a layered spherically symmetric conductor, it is possible to derive a simple analytic expression for the magnetic field of a current dipole (41):

$$\mathcal{B}(r) = \frac{\mu_0}{4\pi} \frac{FQ \times r_Q + (Q \times r_Q) r / \sqrt{F(r, r_Q)}}{F(r, r_Q)^2} \quad (45)$$

where  $r_Q$  is the location of the current dipole,  $Q$  is the dipole moment vector,  $F(r, r_Q) = a^2(r^2 - r_Q^2 - r) + r^2 \sqrt{F(r, r_Q)} = r^{-1}(a^2 + a^{-1}(\sigma_1 - \sigma_2)r + 2\sigma_2 + 2r - a + 2r + a^{-1}(\sigma_1 - \sigma_2)r r_Q)$ , with  $\sigma = (\sigma_1 - \sigma_2)$ ,  $a = |a|$ , and  $r = |r|$ .

An important feature of the sphere model is that the result is independent of the conductivities and thicknesses of the layers; it is sufficient to know the center of symmetry. The calculation of the electric potential is more complicated: The results can be expressed only as a series expansion of Legendre polynomials, and full conductivity data are required (45). Furthermore, radial currents do not produce any magnetic field outside a spherically symmetric conductor. Thus MEG is, to a great extent, selectively sensitive to tangential sources, and EEG data are required to recover all components of the current distribution.

The obvious advantage of a simple forward model is that a fast analytical solution is available. It has also been shown (46) that a sphere model fitted to the local curvature of the skull's inner surface (4) provides accurate enough estimates for many practical purposes. However, when the source areas are located deep within the brain or in the frontal lobes, it is necessary to use more accurate approaches.

In the first MCG localization studies the body was approximated as a homogeneous semi-infinite space, which can be regarded as a generalization of a spherical model with the radius extended to infinity (47,48).

However, later computer studies have shown that the semi-infinite approximation is oversimplified, and a more accurate description of the thorax shape is needed in the inverse studies (21,49). A slightly more accurate description of the thorax geometry can be obtained by using cylindrical or spheroidal models. However, the analytical expressions for arbitrary dipolar sources become substantially more complex than in the spherical case (50), and only a few studies to apply spheroids have been reported.

**Numerical Approaches.** When a realistic geometry of the head or the thorax is taken into account, numerical techniques are needed to solve the Maxwell equations. When applying the boundary-element method (BEM), electric potential and magnetic field are calculated from the (quasi-static) integral equations [Eqs. (39) and (40)], which can be discretized to linear matrix equations (46,49,51).

In most BEM applications to the bioelectromagnetic forward problem, the surfaces are tessellated with triangular elements, assuming either constant or linear variation for the electric potential on each triangle. However, the accuracy of the magnetic-field computation may suffer if a dipole source is located near a triangulated surface. The accuracy can be improved, for example, by applying Galerkin residual weighting instead of the standard collocation method and by approximating the surfaces with curved elements instead of plane triangles (52).

Realistically shaped geometries of each subject are usually extracted from MRI data. The regions of interest (e.g., the heart,

the lungs, and the thorax, or the brain, the skull, and the scalp need to be segmented from the data first (see section entitled MRI). The volumes or the surfaces are then discretized for numerical calculations. The segmentation and tessellation problems are still tedious and nontrivial (53).

The relatively low conductivity of the skull greatly facilitates the modeling of MEG data. In fact, a highly accurate model for MEG is obtained by considering only one homogeneous compartment bounded by the skull's inner surface (46). With suitable image processing techniques it is possible to isolate this surface from high-contrast MRI data with little or no user intervention.

The boundary-element model is more complex for MEG, because three compartments need to be considered: the scalp, the skull, and the brain. While the surface of the head can be easily extracted from the MRI data, it is difficult to construct a reliable algorithm to automatically isolate the scalp-skull boundary. In addition, special techniques are required to circumvent the numerical problems introduced by the high conductivity contrast due to the low-conductivity skull.

It is also possible to employ the finite-element method (FEM) or the finite-difference method (FDM) in the solution of the forward problem. The solution is then based on the discretization of Eq. (7). In this case, any three-dimensional inductivity distribution and even anisotropic conductivity can be incorporated (54). However, the solution is more time-consuming than with the BEM, and therefore the FEM or FDM has not been used in routine source modeling algorithms which require repeated calculation of the magnetic field from different source distributions.

## SOURCE MODELING

### The Inverse Problem

The goal of the bioelectric (EEG, MCG) and biomagnetic (MEG, MCG) inverse problems is to estimate the primary source current density underlying the electromagnetic signals measured outside or on the surface of the body. Unfortunately, the primary current distribution cannot be recovered uniquely, even if the electric potential and the magnetic field were known precisely everywhere at the surface and outside the body (56). However, it is often possible to use additional anatomical and physiological information to constrain the problem and facilitate the solution. One can also replace the actual current sources by equivalent generators that are characterized by a few parameters. The values of the parameters can then be uniquely determined from the measured data by a least-squares fit. The solution of the forward problem is a prerequisite for dealing with the inverse problem requiring repeated solution of the forward problem.

**The Current Dipole Model.** The simplest, physiologically sound model for the neural or myocardial current distribution comprises one or several point sources, current dipoles. In the simplest case the field distribution, measured at one time instant, is modeled by that produced by one current dipole. The best-fitting *equivalent current dipole* (ECD) can be found by using standard least-squares optimization methods such as the Levenberg-Marquardt algorithm (56).

In the time-varying dipole model, introduced by Scherg and von Cramon (57,58), an epoch of data is modeled with a set of dipoles whose orientations and locations are fixed but whose

amplitudes vary with time. Each dipole corresponds to a small patch of cerebral cortex or other structures activated simultaneously or in a sequence. The precise details of the current distribution within each patch cannot be revealed by the measurements, which are performed at a distance in excess of 3 cm from the sources.

As a result of the modeling, one obtains the locations of the sources and the orientation of the dipole component tangential to the inner surface of the overlying skull. In addition, traces of the evolution of the source strengths are obtained. Again, the optimal source parameters are found by matching the measured data collected over a period of time with those predicted by the model using the least-squares criterion.

From a mathematical point of view, finding the best-fitting parameters for the time-varying multipole model is a challenging task. Because the measured fields depend nonlinearly on the dipole position, the standard least-squares minimization routines may not yield the globally optimal estimates. Therefore, global optimization algorithms (59) and special fitting strategies (60), taking into account the physiological characteristics of particular experiments, have been suggested. For each candidate set of dipole positions and orientations it is, however, straightforward to calculate the optimal source amplitude waveforms using linear least-squares optimization methods (61).

In cardiac studies, an ECD is applicable for approximating the location and strength of the net primary current density confined in a small volume of tissue. Myocardial depolarization initiated at a single site spreads at a velocity of about 0.4 mm/ms to 0.8 m/ms, and the ECD can be thought to be moving along the "center of mass" of the excitation. In practice, localization based on a single ECD is meaningful only during the first 10 ms to 20 ms of excitation.

Because both nonlinear fits for spatial coordinates and linear fits for dipole moment parameters need to be searched at every time instant, the use of even two ECDs becomes very complicated in cardiac studies. Alternatively, cardiac excitation can be modeled with a set of spatially fixed stationary or rotating dipoles, but attempts to define the time courses of the dipole magnitudes usually result in physiologically unacceptable results.

**The Current Multipole Expansion.** It is often convenient to present the electric potential and the magnetic field as multipole expansions. In the current multipole expansion the field due to the primary current,  $\mathbf{B}_p$  in Eq. (10), is expressed as a Taylor series (62). Thus, more complex source current configurations can be described as higher-order multipole moments, such as quadrupole moments.

Different source models can be built by combining dipole and quadrupole moments. A current dipole is actually the lowest-order term in a general current multipole expansion (62); higher-order terms, such as quadrupoles and octapoles, can be used to account for more complex primary current configurations (48).

**Distributed Source Models.** Another approach often taken in source modeling is to relax the assumptions on the sources and use various estimation techniques to yield a distributed image of the sources. These methods include, for example, the minimum norm estimates (63), magnetic-field tomography

(MFT) (64), and low-resolution electromagnetic tomography (LORETA) (65).

The source images can provide reasonable estimates of complex source configurations without having to resort to complicated multipole fitting strategies. However, one must keep in mind that even if the actual source is pointlike, its image is typically blurred, extending a few centimeters in each linear dimension. Therefore, the size of the "blobs" in the source images does not directly relate to the actual dimensions of the source but rather reflects an intrinsic limitation of the imaging method.

The basic concept relevant to all distributed source estimation methods is the lead field. The signal  $b_k$  detected by the  $k$ th sensor in the sensor array is a linear functional of the primary current distribution  $J^p$  and can be expressed as

$$b_k = \int_G L_k(r) \cdot J^p(r) dr \quad (12)$$

where the integration extends over the source region  $G$ , which can be a curve, a surface, or a volume. The functions  $L_k$  are often called lead fields, which can be readily obtained by solving the forward problem for dipole sources.

The minimum-norm estimate (66) is the current distribution that has the smallest norm and is compatible with the measured data. Here, the norm is defined by

$$\|J\|^2 = \int_G |J(r)|^2 dr \quad (13)$$

The minimum-norm estimate  $J^m$  can be expressed as a weighted sum of the lead fields,  $J^m = \sum_{k=1}^N c_k L_k$ . The weighting coefficients are found by fitting the data, computed from the minimum-norm estimate with those actually measured. Since the lead fields in a large array are almost linearly dependent, regularization techniques are needed to produce stable estimates.

Another type of a distributed source model was developed for reconstructing the sequence of ventricular depolarization by van Oosterom et al. (66). Their model is based on a uniform double layer of constant strength. Lead fields for MCG and ECG sensors are evaluated at each node on the endo- and epicardial surfaces of the heart. These transfer functions are weighted by the Heaviside time-step function to define the onset of excitation at each surface node. Physiological constraints and regularization are then applied to limit the number of solutions.

### Regularization and Constraints

The bioelectromagnetic forward problem can be written as  $b = Lx - e$ , where vector  $x$  represents the unknown (linear and nonlinear) source parameters, vector  $b$  consists of the measured MEG/EEG or MCG/ECG signals, vector  $e$  contains the contribution of measurement noise, and matrix  $L$  is effectively the transfer (lead field) function between the sources and the measurement sensors. Even small contributions of the noise  $e$  make the solution  $x$  very ill-posed. Therefore, regularization techniques are needed to stabilize the solution (67).

In bioelectromagnetic studies dealing with source distributions, the most frequently applied techniques include the truncated-eigenvalue singular value decomposition (68) and

the  $L$ -curve method (67). Another new approach is based on Wiener filtering and orthogonalized lead fields (68). In addition, spatial weighting can be applied to improve the solutions (21,69). Further improvements are achieved by applying more than one constraint at the same time (70).

One can also make explicitly the additional assumption that the activated areas have a small spatial extent. For example, the MFT algorithm obtains the solution as a result of an iteration in which the probability weighting is based on the previous current estimate (64). According to the authors, this procedure produces more focal images than the traditional minimum-norm solutions. Another possibility is to use a MUSIC-type probability weighting (61) combined with cortical constraints to focus the image (71).

An approach that incorporates the desire to produce focal source images is to use the  $L^1$  norm, that is, the sum of the absolute values of the current over the source space, as the criterion to select the best current distribution among those compatible with the measurement (72-74). In contrast to the traditional  $L^2$ -norm cost function (see Eq. (13)), the  $L^1$ -norm criterion yields estimates focused to a few small areas within the source space.

The most powerful way to constrain the bioelectromagnetic inverse problem is to apply anatomical and functional *a priori* information. For example, accurate reconstruction of the cortex surface or myocardial tissue from MRI data limits the spatial extent and orientation of the sources (75). Solutions can also be made more robust by requiring temporal smoothness. Invasively recorded signals such as intraoperative potential recordings from the cortex or from the heart can also be very valuable in developing proper physiological and temporal constraints for distributed sources.

### The Relation between Bioelectric and Biomagnetic Signals

Both bioelectric and biomagnetic fields are generated by the same activity. As a consequence, there must be a correlation both in the temporal waveforms and in the spatial maps of the measured signals. Therefore, it is evident that bioelectric and biomagnetic measurements reveal partly redundant information. However, neither one can be used to uniquely reproduce the other; there are current configurations that produce either electric or magnetic field, but not both. A practical example of a magnetically silent source is a radial dipole in a spherically symmetric conductor. On the other hand, solenoidal currents do not produce any electric potential. This may become important, for example, in cardiac exercise studies (76). Therefore, a combination of magnetic and electric recordings seems appealing to obtain more complete information about the current distributions. Still, few attempts to combine electric and magnetic data have been reported (21,56,77,78).

In the previous considerations it was assumed that the volume conductor is homogeneous or piecewise homogeneous. However, many biological tissues are organized directionally, and the electrical conductivity depends on the direction of the fibers. For example, the conductivity in myocardial fibers is about three times higher in the main fiber direction than across the fibers. Colli-Franzone et al. (79) showed that a classical uniform dipole layer, as representing the myocardial wavefront, should be revised to take into account the anisotropic nature of the tissue. They were able to explain experimentally measured

potential distributions with an oblique dipole layer, where the dipoles may also have tangential components in addition to the normal component.

Wikawa (80) studied isolated animal preparations and employed microSQUIDS and microelectrodes to measure magnetic and electric fields during and after applying a current stimulus. According to their results, the magnetic field is more sensitive to the underlying anisotropy than the electric potential. With such combined electric and magnetic recordings, it is at least in principle possible to determine the intra- and extracellular conductivity values.

The anisotropic properties of the heart are especially evident near the ventricular apex, where the spiral arrangements of the myocardial fibers can be observed on the epicardial surface. It has been argued that this kind of vortex geometry leads to electrically silent components in magnetic field. However, van Oosterom et al. (68) arrived at the conclusion that the anisotropy does not play a significant role in the ECG or MCG during the normal ventricular depolarization. On the other hand, the findings of Brackmeier et al. (76) in pharmacological MCG stress testing indicate that the anisotropy may cause larger repolarization changes in multichannel MCG signals than in the simultaneously recorded ECG maps.

The tissue is directionally oriented also in the brain. For example, the conductivity of the white matter in the direction of the fibers may be 10 times higher than the conductivity across the fibers. In the cerebral cortex, the corresponding factor is about two. In general, the anisotropy influences the body surface potentials and magnetic fields. However, in the sphere model a difference between the radial and the two tangential conductivities does not affect the magnetic field, while the influence on the electric potential is still substantial.

## INTEGRATION WITH OTHER IMAGING MODALITIES

First high-field MRI devices provide precise anatomical data. Besides reconstruction of accurately shaped volume-conductor models, anatomic MRI data on the heart and the brain are necessary to combine the inverse solutions with the anatomy in a clinically useful presentation. Examples of source displays of both MRI slices and three-dimensional surface reconstructions are shown in the section entitled "Applications."

Segmentation of the structures of interest from image data is presently the most time-consuming part in constructing individualized boundary element models. In the medical imaging field, accurate extraction of anatomic structures from image data sequences is still an open problem. In practice, manual extraction of the objects of interest—for example, from MRI slices—is often considered the most reliable technique.

Recently, automated region-based and boundary-based segmentation and triangulation methods have been developed, for example, for extracting the lungs, heart, and thorax, or the brain and skull. In region-based methods, some features based on the intensity of the images are used to merge voxels. The boundary-based methods, in turn, rely on an intensity gradient detector. Both methods have limitations, but the utilization of prior geometrical knowledge, such as triangulated surfaces generated from data of other subjects, provides useful additional information. For example, a deformable pyramid model can then provide automatic segmentation and triangulation of the anatomic objects (58,81).

MRI is still fairly expensive, especially for large patient populations. Thus, methods are being developed to use other imaging methods for reconstructing individualized triangulated surfaces. In cardiac studies, two orthogonal thorax X-ray projections, or ultrasound images of the heart combined with three-dimensional (3D) digitization of the thorax surface, can be utilized to acquire patient-specific geometry models.

In principle, CT images could be used instead of MRI to construct a boundary element model of the head. The skull is particularly easy to isolate from these data. However, the classification of soft tissues is often easier from MR images and the radiation load imposed by a CT scan is generally considered too high for healthy subjects.

To present the MEC/MCG inverse solutions accurately on the individual anatomy, special care needs to be taken with regard to combining the different coordinate frames. Prior to biomagnetic recordings, one has to fix some marker points, for example, with a 3-D digitization system. During MRI or X-ray imaging, specific markers clearly visible and identifiable, such as vitamin pills or tubes filled with MgCl solution, are attached on the reference points. Three or more markers are usually required to achieve sufficient accuracy in the data fusion.

Functional MRI perfusion studies of ischemic or infarcted heart are particularly valuable in developing physiological constraints and in validating the MCG/ECG localization results of ischemia or arrhythmogenic tissue. In brain studies, new possibilities are opened by combining the millimeter-level spatial resolution of functional MRI and the millisecond-scale temporal resolution of MEG and EEG. Weighting of minimum-norm solutions by functional MRI voxel information has been applied, for example, in visual stimulation studies (82).

However, it must be taken into account that fMRI and biomagnetic measurements are not always detecting common activity. Very clear changes of electric and magnetic signals can be easily missed by fMRI if they occur rarely or are very transient thus producing relatively small average changes in the metabolic level. Furthermore, all experimental setups cannot be easily used in both biomagnetic and fMRI studies. It may thus be necessary to often compare the final results of the analysis of each modality rather than aiming at a combination during the source reconstruction.

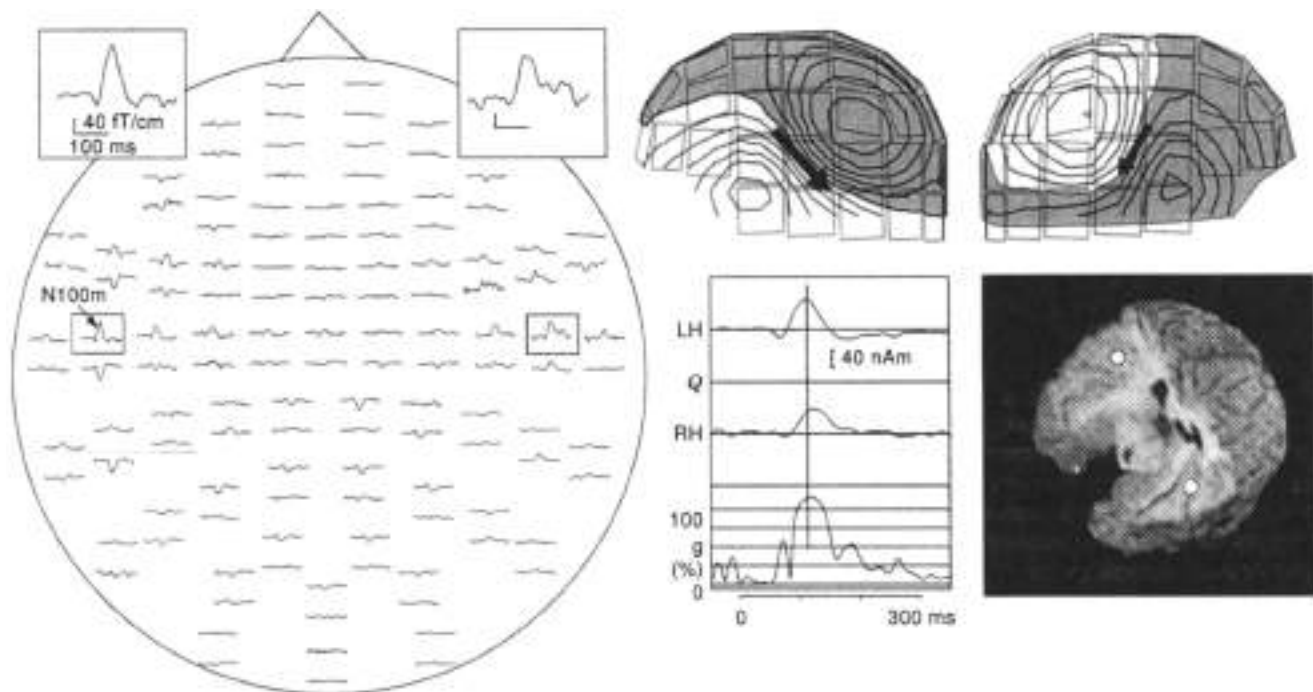
It is also possible to utilize positron emission tomography (PET) studies in combination with the electromagnetic methods. However, PET imposes a radiation load on the subject, and therefore the possibilities to perform multiple studies on a given subject are limited. Furthermore, PET is available only in a few centers, whereas MRI systems capable of functional imaging are generally available in modern hospitals.

## APPLICATIONS

### Brain Studies

**Auditory Evoked Fields.** Fig. 3 shows the results of a typical auditory evoked-response study performed with a whole-head MEG instrument (83). The responses were elicited by 50 ms tones delivered every 4 s to the subject's right ear. The data were averaged over about 100 repetitions with the stimulus onset as a trigger.

The signals were modeled with two current dipoles in a spherically symmetric conductor. The optimal locations,



**Figure 8.** Left: Auditory evoked magnetic fields recorded with a 122-channel magnetometer (30–100 mV, 1 kHz tones presented to the subject's right ear once every 4 s). The head is viewed from above, and the helmet surface has been projected onto a plane; the nose points up. Right, above: The pattern of the field component normal to the helmet surface,  $B_n$ , shows the peak of the response. White indicates magnetic flux into and gray out of the head. The locations of the sensor wires are indicated with squares. The positions and orientations of the two current dipoles modeling the data are projected to the helmet surface. Middle: Time dependence of the dipole strengths, indicating the time behavior of the active area in the left (LH) and right (RH) hemispheres.  $Q$  denotes the dipole moment, goodness-of-fit ( $\rho^2$ ) indicates how well the model agrees with the measurement. Right, below: The locations of the dipoles, projected on an MRI surface rendering, viewed from above. To show the supratemporal surface, frontal lobes have been removed from the images. (Modified from Ref. 83.)

orientations, and time courses of the dipoles were determined with a least-squares search. Fig. 3 shows the averaged data, the distribution of the magnetic field component normal to the measurement surface at the peak signal value, the time courses of the source amplitudes, and the locations of the sources superimposed on a 3-D surface rendering computed from the subject's MRI data. The locations of the sources agree nicely with the known site of the auditory cortex on the supratemporal plane. Furthermore, the time courses of the source amplitudes show that the source in the left hemisphere, opposite to the stimulus, is stronger and peaks about 20 ms earlier than the source on the right.

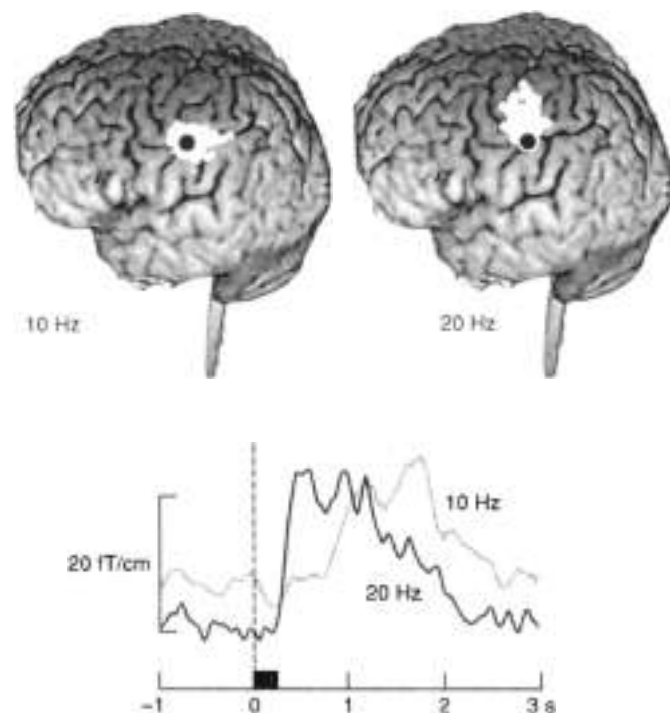
**Characterization of Cortical Rhythms and their Reactivity.** Since the advent of EEG, the rhythmic oscillations of various cortical areas have been described and also utilized in clinical diagnosis, but their functional significance has remained unclear. With the whole-scalp neuromagnetometers, studies of cortical rhythms have become feasible. Because these rhythms do not repeat themselves, it is mandatory to record them simultaneously over the whole scalp.

The neuromagnetic brain rhythms in healthy adults have been recently characterized in (84) and their reactivity has

been quantified during different situations. An efficient way to reveal task-related changes in the level of different frequency components is to filter the signal to the frequency passbands of interest, rectify it, and finally average the rectified signal with respect to the event of interest, like the onset of a voluntary movement (85).

Such an analysis has unraveled new features—for example, of the well-known mu rhythm, which is seen in the EEG records over the somatomotor cortices of an immobile subject. The comb shape of the mu rhythm already indicates the coexistence of two or three frequency components, strongest around 10 Hz and 20 Hz. The sources of the magnetic mu rhythm components cluster over the hand somatomotor cortex, with slightly more anterior dominance for the 20 Hz than for the 10 Hz cluster (see Fig. 4) (85). This difference suggests that the 20 Hz rhythm receives a major contribution from the precentral motor cortex, whereas the 10 Hz component seems mainly postcentral (somatosensory) in origin.

Further support for the functional segregation of these rhythms comes from their different reactivity to movements (15). The level of the 10 Hz rhythm starts to dampen 2 s before a voluntary movement and then returns back within 1 s after the movement. Suppression of the 20 Hz rhythm starts later



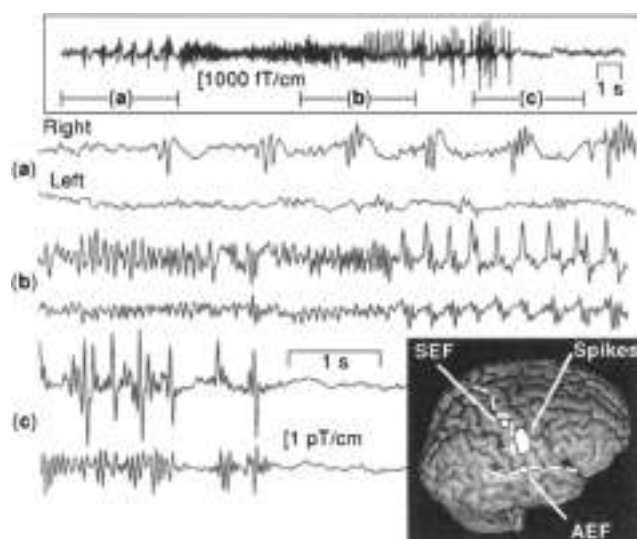
**Figure 4.** Reactivity of spontaneous activity over the somatomotor hand region in association with voluntary right index finger movements. The time dependencies of the 10 Hz and 20 Hz activities are indicated by the two traces showing the temporal spectral evolution (64) of the signal recorded over the left somatomotor hand area. The locations of sources corresponding to 10 Hz and 20 Hz activity are indicated on the 3-D surface rendering of the subject's magnetic resonance images. The site of the source for electrical stimulation of the median nerve at the wrist is shown by the black dot.

and is relatively smaller, and the "rebound" after the movement is earlier and stronger than in the 10 Hz band.

**Locating Epileptic Foci.** Many patients with drug-resistant epilepsy suffer from seizures triggered by a small defective brain area. In preoperative evaluation of these patients, it is important to know whether their epileptic discharges are focal and how many brain areas are involved, what is the relative timing between the foci, and how close they are to functionally irremovable locations such as the motor and speech areas. MEG recordings have been able to answer some of these questions (86,87). The patients cannot be studied with MEG during major seizures, owing to movement artifacts, but in many cases the foci can be identified from interictal discharges occurring during the periods between the seizures.

As an example of a recording during an actual seizure (Fig. 5) depicts MEG signals from a patient who suffered from convulsions in the left side of his face (88). He was able to trigger the seizure by touching the left-side lower gum with his tongue.

The recordings show clear epileptic spikes which appear only in the right hemisphere at first, but later start to emerge in the corresponding areas of the left hemisphere as well. After the 14 s seizure, the epileptic discharges ended abruptly (Figure 5 also depicts locations of the spike ECDs, superimposed on the patient's MRI surface rendering. The sources are clustered along the anterior side of the central sulcus, extending 1 cm to 3 cm lateral to the SI hand area, as determined by somatosensory median-nerve evoked responses. The sources of the epilep-



**Figure 5.** Epileptic discharges after voluntary triggering (88). The trace on the top illustrates MEG activity from the right hemisphere during the whole 14 s seizure. In the middle and lower parts of the figure, selected periods (a, b, c) are expanded and signals generated by the corresponding area in the left hemisphere are shown for comparison. Lower right corner: Locations of ECDs for central and interictal spikes (white cluster), and for auditory (AEF) and somatosensory (SEF) evoked fields, superimposed on the patient's MRI surface rendering. The course of the Rolandic and Sylvian fissures are indicated by the white dashed lines. (Adapted from Ref. 88.)

tic spikes thus agree with the face representation area in the precentral primary motor cortex and are in accord with the patient's clinical symptoms. Spikes generated by the focus in the left hemisphere lagged behind the right-sided spikes by about 20 ms and probably reflected transfer of the discharges through corpus callosum from the primary to the secondary focus. Identification of secondary epileptogenesis is important for presurgical evaluation of patients because the secondary foci may with time become independent, and removal of the primary focus would then no longer be efficient in preventing the seizures.

#### Cardiac Studies

**Ventricular Preexcitation.** Ventricular preexcitation associated with the Wolff-Parkinson White (WPW) syndrome is caused by an accessory pathway between the atria and the ventricles, which may lead to supraventricular tachycardias and life-threatening arrhythmias refractory to drug therapy. Intervention therapy, such as catheter ablation, is then needed, but a necessary condition for successful elimination of the premature conduction is the reliable localization of the accessory pathway.

Catheter ablation techniques have significantly decreased the need for cardiac surgery, but simultaneously increased the need for accurate noninvasive localization techniques. Noninvasively obtained prior knowledge of the site of the accessory pathway can improve the result and shorten the time needed in invasive catheter mapping, and thus diminish patient discomfort and surgical risk. In addition, shortening the time needed in invasive catheterization also reduces radiation exposure due to fluoroscopy monitoring of catheter positions.

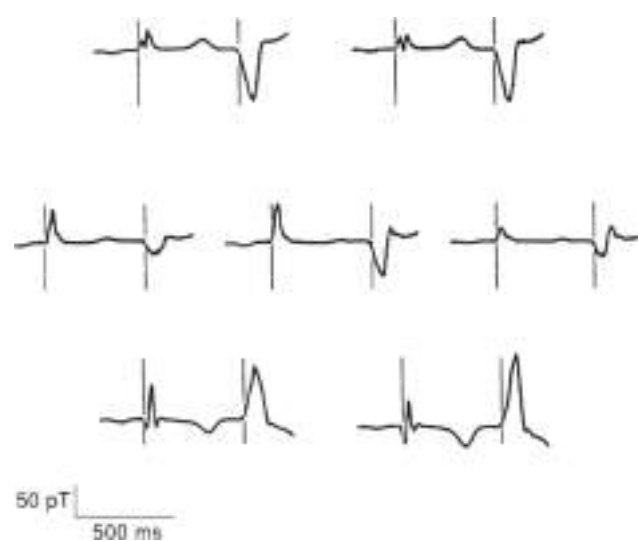
Several MCG studies have been reported on localizing the ventricular preexcitation site in patients with the WPW syndrome (17,20,89,90). The reported accuracy of MCG localizations ranges from 5 mm to 25 mm, which is sufficient to be useful in preablative or presurgical consideration of the patients.

**Ventricular Tachycardia.** Generally, malignant ventricular tachycardia (VT) is much more difficult to locate for ablation treatment than the ventricular preexcitation. It is estimated that the lesion produced by the application of RF currents is about 6 mm in diameter and about 3 mm in depth. Currently, clinical practice for precise localization introduces several catheters through arteries and veins into the ventricles for invasive recordings of cardiac activation sequences. This procedure can be very time-consuming, and noninvasively obtained

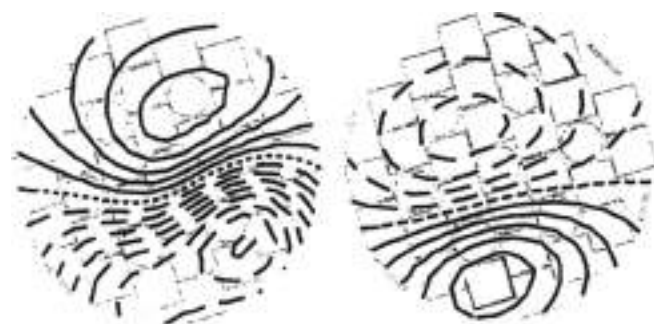
information could shorten the procedure from several hours even to less than 1 hour.

VT patients include postmyocardial infarction patients, patients with different cardiomyopathies, and patients with nonsustained VT. MCG studies reported so far have attempted to locate the origin of ventricular extrasystoles or arrhythmias that have occurred spontaneously during the MCG recording (17,10,20). The results have been compared to the results of successful catheter ablations, presented over X-ray and magnetic resonance images when available. In such comparisons the average MCG locations were found to be within 2 cm from the invasively determined sites.

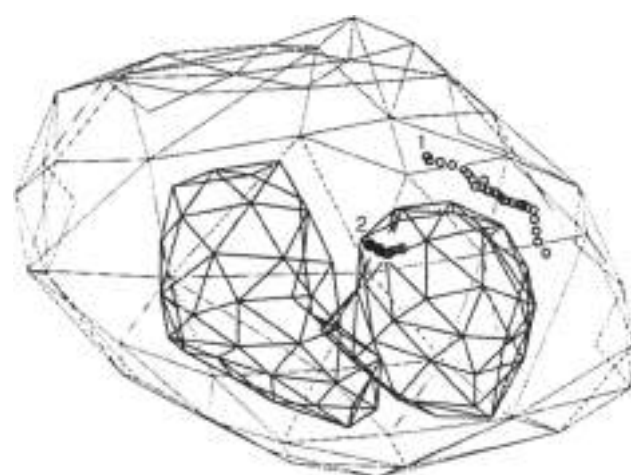
Examples of such studies are displayed in Fig. 6. These results were obtained with an individualized boundary-element torso model; an example is displayed in Fig. 7(a).



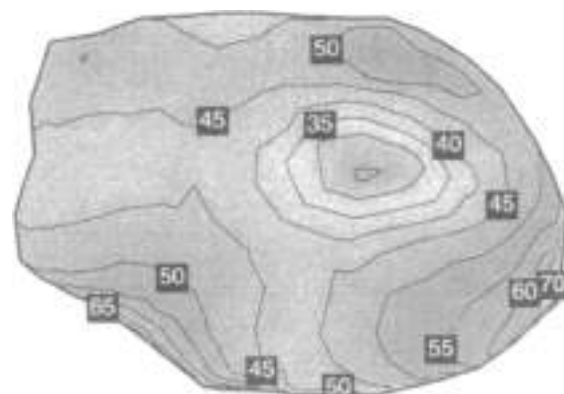
(a)



(b)



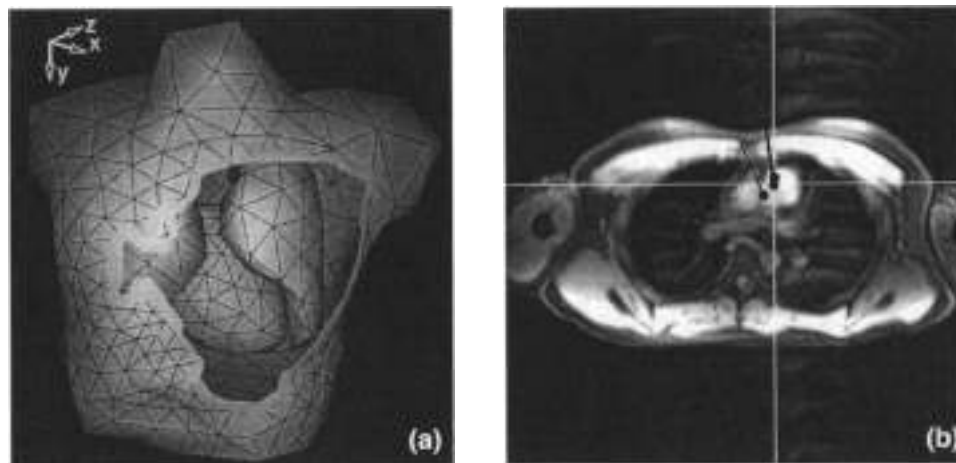
(c)



(d)

**Figure 4.** (a) MCG curves recorded from a patient suffering from ventricular tachycardia (VT). The seven axial gradiometers show (1) a normal sinus-rhythm beat and (2) an arrhythmogenic ventricular extrasystole (VES). (b) Isocontours of the magnetic field component perpendicular to the sensor array (see Fig. 2). The field values were interpolated from the measured data with the minimum norm estimation (36). Solid and dashed lines here indicate, respectively, magnetic flux towards or out of the chest. The step between adjacent contours is 1 pT. (c) MCG localization results obtained with a single moving EIT. (d) Ventricular activation sequence reconstructed from the VES by the method reported in Ref. 32.





**Figure 7.** (a) An example of a boundary element mesh model constructed from MRI data. The surfaces of the body, the lungs, and the heart are tessellated into triangulated networks. The total number of triangles here is about 1500. (b) An example of MCG localization of tachycardia. The patient was suffering from continuous atrial tachycardia with the heart rate of over 140 beats per minute. ECD localization was performed from 67-channel MCG data at the onset of the P-wave, and the 3-D locations were superimposed on the MRI data. Catheter ablation performed later at the location proposed by the MCG result terminated all arrhythmias.

**Cardiac Evoked Fields.** Artificial dipole sources inserted in the heart with catheters (e.g., during routine electrophysiological studies) have been tested to verify the MCG localization accuracy (17,20,32). For example, Fenici et al. (22) studied five patients in whom a nonmagnetic pacing catheter (16) was used to stimulate the heart during MCG recordings. In general, the MCG localization results at the peak of 2 ms catheter stimuli and at the onset of paced myocardial depolarization were within 5 mm from each other. Because these are two physically different sources, the study provided further support for the good localization accuracy of the MCG method. In another recent study (91), simultaneous MCG and ECG mapping recordings were performed during pacing in 10 patients. The localizations were compared to catheter positions documented on fluoroscopic X-ray images. MCG results were, on the average, within 5 mm from the documented catheter position, while the ECG showed somewhat worse accuracy.

**Reconstruction of Distributed Sources.** For comparison of MCG and ECG mapping results, simultaneously recorded MCG and ECG data were applied in reconstructing ventricular depolarization isochrones on the endo- and epicardial surfaces of the heart of a healthy normal subject (92). The results showed almost identical isochrones from both magnetic and electric data. An example of the MCG isochrone reconstructions is shown in Fig. 6(d).

The minimum-norm estimates (MNE) have been applied in estimating the primary current distributions underlying measured MCG signals. An intrinsic problem associated with MNE is that it has a poor depth resolution of the sources without proper regularization and physiological constraints. Various regularization and depth weighting methods have shown promising results (21,68).

Recently, depth-weighted MNE reconstruction has been applied in MCG data recorded in patients with chronic myocardial ischemia (18,93). Clinical validation for the results was provided by SPECT imaging. In general, the site of smallest current density, i.e., the missing depolarization component, was in good agreement with the SPECT result.

## OTHER APPLICATIONS

Until now, magnetic source imaging studies have been focused on the brain or the heart. However, other applications are being developed as well.

Studies of compound action fields (CAF) from peripheral nerves require a very high sensitivity, because the signal amplitudes are below 10 fT (94-97). In addition, signal averaging of hundreds of stimulated sequences may be required to find the CAF waveforms. Analysis of the waveforms demonstrates the quadrupolar nature of neural activity, provided that the observations are performed at a distance from the depolarized segment (96). Multipole analysis has been applied to model the depolarization process; dipole terms reveal the location and intensity of the source, while octupole terms are related to its longitudinal extension along the nerve fibers. In addition to studying propagation of the nerve impulses, multichannel measurements can reveal abnormalities such as proximal conduction blocks in the spinal nervous system (97).

High sensitivity and specific signal processing are also needed to detect low-frequency (0.05-0.15 Hz) magnetic signals from gastrointestinal system (98,99). Distinguishing between gastric and small bowel signals may provide a new tool to study abnormalities in the gastrointestinal system (98). Multichannel recordings allow feasible and continuous monitoring of magnetically marked capsules within the gastrointestinal tract with a temporal resolution in the order of milliseconds and a spatial resolution within a range of millimeters (99).

## DISCUSSION

### Modeling

Despite the inevitable analogies in source analysis, very useful information has been obtained by using relatively simple models. For example, the localization of functional landmarks in the brain using the current dipole model and a spherically symmetric conductor in the forward calculations has already

been developed to the extent of being a reliable clinical tool (16). The results of these studies have also been often verified in direct intraoperative recordings.

The focal source analysis methods are sometimes criticized for being too extreme simplifications of the actual current distributions, which renders them rather useless in the study of complicated functions performed by the human brain. This intuitively appealing opinion is not well backed up by experimental data. Rather, recent fMRI data may be taken to indicate that the significant changes in metabolic activity associated even with complicated cognitive tasks might well be relatively focal.

Reconstruction methods to deal with source distributions are under development, but there are still difficulties in interpretation of the results obtained from measured data. Implementation of available physiological information and constraints is probably needed to obtain a reasonable correlation with actual physiological events in the source regions. If the assumptions of the source model are not compatible with the characteristics of the actual electrophysiological sources, misleading estimates may ensue. As discussed in the section entitled "Distributed Source Models," the distributed source model may produce a distributed estimate even for a focal source. Only very recently have there been attempts to reliably estimate the actual extent of the current source using Bayesian parameter estimation (100) in conjunction with reasonable physiological and anatomical constraints.

Invasively recorded cardiac signals, such as potentials measured during electrophysiological studies on epi- and endocardial surfaces, provide the golden standard for validation of the MCG/ECG inverse solutions. Even though patient populations studied by MCG before or during invasive catheterization are still relatively small, the localization studies of various cardiac arrhythmias have shown encouraging results. Multichannel systems and accurate combination of the results with cardiac anatomy have improved the accuracy to the order of 5 mm to 10 mm, which is sufficient to aid in planning the curative therapy of arrhythmia patients. Further validation for the MCG localization accuracy has been obtained by locating artificial dipole sources, such as pacing catheters inserted into the heart during electrophysiological studies.

#### Future Trends

The arrhythmogenic substrate is not manifested in all normal sinus rhythm recordings, and interventions may be needed during MCG to stimulate controlled arrhythmias to locate them. Thus, MCG should be available in a catheterization laboratory, but the demand of magnetical shielding and liquid helium is, in practice, limiting the use of MCG mapping in guiding invasive arrhythmias localization. For this purpose, compact-size higher-order magnetometer arrays operated without external shielding would be required.

MCG mapping under exercise is a promising tool for noninvasive characterization and localization of myocardial ischemia (78). Improved source modeling and localization methods are under test, especially in patients with coronary artery disease (93).

Despite over 20 years of MCG and MEG research, common standards of measurement techniques, data processing, and presentation are still lacking. Suggestions for such standards are emerging, but it is clear that there is and will be large differences between sensors and their arrangement in multichannel

magnetometers. Fortunately, tools such as MNE (63,68) have been developed to interpolate signal morphologies and isocurrent maps that are directly comparable to studies performed in other centers.

The field of magnetic source imaging may expand during the next few years with the implementation of low-noise high- $T_c$  SQUID arrays that can be operated at the temperature of liquid nitrogen (101–103). The higher noise level of the high- $T_c$  SQUIDs is, particularly in MEG studies, partly compensated by the smaller distance between the sensors and the body. At present, however, the low- $T_c$  SQUIDs are easier to produce and thus cheaper than the high- $T_c$  ones.

The future of MEG, with commercial whole-head instruments now available, looks promising. The capability to monitor activity of several cortical regions simultaneously in real time provides a unique window to study the neural basis of human cognitive functions. Important information can be obtained both from evoked responses and from spontaneous on-going activity.

Effective signal processing and source modeling software is going to be increasingly important to extract all available functional data from the electromagnetic signals emerging from the brain and the heart. The widely discussed issue of whether the electric or magnetic technique is superior to the other is not of primary importance. Rather, one should apply a suitable alliance of different types of imaging methods, taking into account the characteristics and aims of the experiment being performed to yield optimal information about the functions of the biological system.

#### BIBLIOGRAPHY

1. G. Barle and R. McFee. Detection of the magnetic field of the heart. *Am Heart J*, **66**: 95–96, 1963.
2. D. Cohen. Magnetoencephalography: Evidence of magnetic fields produced by alpha-rhythm currents. *Science*, **141**: 784–786, 1968.
3. D. Cohen, E. A. Edelback, and J. E. Zimmerman. Magneto-cardiograms taken inside a shielded room with a superconducting point-contact magnetometer. *Appl. Phys. Lett.*, **16**: 276–280, 1970.
4. R. Hari and R. Ilmoniemi. Cerebral magnetic fields. *CRC Crit. Rev. Biomed. Eng.*, **14**: 93–126, 1986.
5. M. Hämäläinen et al. Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain. *Rev. Mod. Phys.*, **65**: 413–497, 1993.
6. J. Nenonen. Solving the inverse problem in magneto-cardiography. *IEEE Eng. Med. Biol.*, **13**: 487–496, 1994.
7. K. Näätänen, R. Ilmoniemi, and K. Alho. Magnetoencephalography in studies of human cognitive brain function. *Trends Neurosci.*, **17**: 389–395, 1994.
8. W. Andree and H. Nowak (eds.). *Magnetics in Medicine*. Berlin: Wiley VCH, 1992.
9. R. Hari. Magnetoencephalography as a Tool of Clinical Neurophysiology. In E. Niedermeyer and F. Lopes da Silva (eds.), *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*, 3rd ed., Baltimore, MD: Williams and Wilkins, 1993, pp. 1035–1061.
10. R. Salmelin et al. Dynamics of brain activation during picture naming. *Nature*, **368**: 463–465, 1994.
11. R. Salmelin et al. Impaired visual word processing in dyslexia revealed with magnetoencephalography. *Ann. Neurol.*, **40**: 157–162, 1996.

12. R. Näätänen et al., Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature*, **385**: 432-434, 1997.
13. R. Hinri and R. Salmelin, Human cortical oscillations: A neuro-magnetic view through the skull. *Trends Neurosci.*, **20**: 44-49, 1997.
14. C. Teuche et al., Characterizing the local oscillatory content of spontaneous cortical activity during mental imagery. *Cogn. Brain Res.*, **2**: 243-249, 1995.
15. R. Salmelin et al., Functional segregation of movement-related rhythmic activity in the human brain. *NeuroImage*, **2**: 237-243, 1995.
16. J. D. Lewine and W. W. Orrison, Magnetoencephalography. in W. G. Bradley and G. M. Bydder (eds.), *Advanced MR Imaging Techniques*, London: Martin Dunitz, 1997, pp. 333-354.
17. R. Fenici and G. Melillo, Magneto-cardiography: Ventricular arrhythmias. *Eur Heart J*, **14** (Suppl. E): 52-60, 1993.
18. M. Mäkipää et al., New trends in clinical magnetoencephalography. in C. Aze et al. (eds.), *Advances in Biomagnetism Research: Biomag98*. New York: Springer-Verlag, 1998 (in press).
19. M. Oeff and M. Burghoff, Magneto-cardiographic localization of the origin of ventricular ectopic beats. *PACE*, **17**: 617-622, 1994.
20. W. Muehge et al., Evaluation of the non-invasive localization of cardiac arrhythmias by multichannel magneto-cardiography. *Int J Cardiac Imaging*, **12**: 47-59, 1996.
21. G. Strack, R. Lamotte, and M. Gardner, Magneto-cardiographic and Electro-cardiographic Mapping Studies. in H. Weinstock (ed.), *SQUID Sensors: Fundamentals, Fabrication and Applications*, Amsterdam: Kluwer Academic Publishers, 1996, NATO ASI Ser., pp. 413-444.
22. R. Fenici et al., Non-fluoroscopic localization of an amagnetic stimulation catheter by multichannel magneto-cardiography. *PACE*, 1998, in press.
23. J. Muusinen, Magneto-cardiography as identifier of patients prone to malignant arrhythmias, in C. Baumgartner et al. (eds.), *Biomagnetism: Fundamental Research and Clinical Applications*, New York: Springer-Verlag, 1996, pp. 606-611.
24. U. V. Lounasmaa, *Experimental Principles and Methods Below 1K*. London: Academic Press, 1974.
25. T. Ryhänen et al., SQUID magnetometers for low-frequency applications. *J Low Temp Phys.*, **70**: 287-306, 1999.
26. B. D. Josephson, Possible new effects in superconductive tunnelling. *Phys Lett.*, **1**: 251-253, 1962.
27. J. Clarke, A superconducting galvanometer employing Josephson tunnelling. *Philos Mag.*, **13**: 11A, 1966.
28. J. Clarke, W. M. Gouhan, and M. B. Ketchen, Tunnel junctions of SQUID fabrication, operation, and performance. *J Low Temp Phys.*, **26**: 99-144, 1979.
29. C. D. Tesche et al., Practical dc SQUIDs with extremely low  $1/f$  noise. *IEEE Trans Magn.*, **MAG-21**: 1032-1035, 1985.
30. D. Cohen, Magnetic measurement and display of current generators in the brain. Part I: The 2-D detector. *Dig. 12th Int. Conf. Med. Biol. Eng. Jerusalem*, p. 15, 1979 (Petah Tikva, Israel: Beilinson Medical Center).
31. S. N. Erne and G. L. Ruzani, Performance of Higher Order Planar Gradiometers for Biomagnetic Source Localization. in H. D. Halblöhm and H. Lubbig (eds.), *SQUID 88 Superconducting Quantum Interference Devices and their Applications*, Berlin: de Gruyter, 1988, pp. 951-961.
32. P. Carelli and H. Leoni, Localization of biological sources with arrays of superconducting gradiometers. *J Appl. Phys.*, **60**: 645-650, 1986.
33. J. R. T. Kuutila et al., A 122-channel whole-cortex SQUID system for measuring the brain's magnetic fields. *IEEE Trans. Magn.*, **29**: 3315-3320, 1993.
34. D. Cohen, Low-field room built at high-field magnet lab. *Phys. Today*, **23**: 56-57, 1970.
35. S. N. Erne et al., The berlin magnetically shielded room (BMSR): Section B—performance. in S. N. Erne, H. D. Halblöhm, and H. Lubbig (eds.), *Biomagnetism*. Berlin: de Gruyter, 1981, pp. 79-87.
36. V. O. Kelhä et al., Design, construction, and performance of a large volume magnetic shield. *IEEE Trans. Magn.*, **MAG-18**: 260-270, 1982.
37. J. Kuutila et al., Design considerations for multichannel SQUID magnetometers. in H. D. Halblöhm and H. Lubbig (eds.), *SQUID'86 Superconducting Quantum Interference Devices and their Applications*, Berlin: de Gruyter, 1986, pp. 939-944.
38. S. N. Erne et al., The positioning problem in biomagnetic measurements: A solution for arrays of superconducting sensors. *IEEE Trans. Magn.*, **MAG-23**: 1319-1322, 1997.
39. R. Plonsey, *Biostatic Phenomena*. New York: McGraw-Hill, 1969.
40. J. P. Wikswo, Jr., J. P. Barach, and J. A. Freeman, Magnetic field of a nerve impulse: First measurements. *Science*, **200**: 53-55, 1980.
41. J. P. Wikswo, Jr., Cellular magnetic fields: Fundamental and applied measurements on nerve axons, peripheral nerve bundles, and skeletal muscle. *J. Clin. Neurophysiol.*, **8**: 170-188, 1991.
42. J. Sarvas, Basic mathematical and electro-magnetic concepts of the biomagnetic inverse problem. *Phys. Med. Biol.*, **32**: 11-22, 1987.
43. A. C. L. Bernard, I. M. Duck, and M. S. Lynn, The application of electromagnetic theory to electrocardiology. I. Derivation of the integral equations. *Radiophys. J.*, **7**: 443-466, 1967.
44. D. B. Geselowitz, On the magnetic field generated outside an inhomogeneous volume conductor by internal current sources. *IEEE Trans. Magn.*, **MAG-8**: 346-347, 1973.
45. Z. Zhang, A fast method to compute surface potentials generated by dipoles within multilayer anisotropic spheres. *Phys. Med. Biol.*, **40**: 335-349, 1995.
46. M. S. Hämäläinen and J. Sarvas, Realistic conductivity geometry model of the human head for interpretation of neuro-magnetic data. *IEEE Trans. Biomed. Eng.*, **38**: 165-171, 1989.
47. S. N. Erne et al., Modeling of the His-Purkinje Heart Conduction System. in H. Weinberg, G. Strack, and T. Karila (eds.), *Biomagnetism: Applications & Theory*. New York: Pergamon, 1985, pp. 126-131.
48. J. Nöonen et al., Magneto-cardiographic functional localization using current multipole models. *IEEE Trans. Biomed. Eng.*, **38**: 648-657, 1991.
49. J. Nieminen et al., Magneto-cardiographic functional localization using a current dipole in a realistic torso. *IEEE Trans. Biomed. Eng.*, **38**: 656-664, 1991.
50. B. N. Cuffin and D. Cohen, Magnetic fields of a dipole at special volume conductor shapes. *IEEE Trans. Biomed. Eng.*, **BME-24**: 372-381, 1977.
51. B. M. Horroek, Digital model for studies in magneto-cardiography. *IEEE Trans. Magn.*, **MAG-9**: 440-444, 1973.
52. C. Brebbia, J. Telles, and L. Wrobel, *Boundary Element Techniques: Theory and Applications in Engineering*. Berlin: Springer-Verlag, 1984.
53. J. Löfgren et al., A triangulation method of an arbitrary point set selected from medical volume data. *IEEE Trans. Magn.*, **34**: 2224-2233, 1998.
54. H. Buchner et al., Inverse localization of electric dipole current sources in finite element models of the human head. *Electroencephalogr. Clin. Neurophysiol.*, **102**: 267-278, 1997.
55. H. Helmholtz, Ueber einige Gesetze der Vertheilung elektrischer Ströme in körperlichen Leitern, mit Anwendung auf die

- thierisch-elektrischen Versuche. *Ann Phys Chem*, **49**: 211-233, 250-377, 1953.
56. D. W. Marquardt, An algorithm for least-square estimation of nonlinear parameter. *J. Soc. Ind. Appl. Math.*, **11**: 431-441, 1963.
  57. M. Scherg and D. von Cramon, Two bilateral sources of the late erg as identified by a spatiotemporal dipole model. *Electroencephalogr Clin Neurophysiol*, **62**: 232-244, 1985.
  58. M. Scherg, R. Hari, and M. Hamalainen, Frequency Specific Sources of the Auditory N19-P30-P50 Response Detected by a Multiple Source Analysis of Evoked Magnetic Fields and Potentials, in S. J. Williamson et al. (eds.), *Advances in Biomagnetism*. New York: Plenum, 1989, pp. 97-100.
  59. K. Uutela, M. Hamalainen, and R. Salmeila, Global optimization in the localization of neuro-magnetic sources. *IEEE Trans Biomed Eng.*, **45**: 716-723, 1998.
  60. E. Berg and M. Scherg, Sequential Brain Source Imaging: Evaluation of Localization Accuracy in C. Ogura, Y. Coga, and M. Shimokubi (eds.), *Recent Advances in Event-Related Brain Potential Research*. Amsterdam: Elsevier, 1996.
  61. J. C. Mather, P. S. Lewis, and R. Leahy, Multiple dipole modeling and localization from spatiotemporal MEG data. *IEEE Trans Biomed Eng.*, **30**: 541-557, 1992.
  62. T. E. Kaola, On the current multiple presentation of the primary current distributions. *Natur Commun*, **2D**: 660-664, 1963.
  63. M. S. Hamalainen and R. J. Ilmoniemi, Interpreting magnetic fields of the brain: Minimum-norm estimates. *Med Biol Eng Comput*, **32**: 35-42, 1994.
  64. A. A. Javanmard, J. P. R. Bolton, and C. J. S. Clarke, Continuous probabilistic solutions to the biomagnetic inverse problem. *Inverse Problems*, **6**: 523-542, 1990.
  65. R. D. Pascual-Marqui, C. M. Michel, and D. Lehmann, Low resolution electromagnetic tomography: A new method for localizing electrical activity in the brain. *Int. J. Psychophysiol*, **18**: 49-65, 1994.
  66. A. van Oosterom et al., The magnetocardiogram as derived from electrocardiographic data. *Clin Res*, **67**: 1563-1569, 1990.
  67. P. Hansen, Numerical tools for analysis and solution of Fredholm integral equations of the first kind. *Inverse Problems*, **8**: 849-872, 1992.
  68. J. Nummenen et al., Transformation of multichannel magnetocardiographic signals to standard grid form. *IEEE Trans Biomed Eng.*, **42**: 72-77, 1995.
  69. K. Paasilta et al., Comparison of regularization methods when applied to equatorial minimum norm estimates. *Biomed Tech*, **42** (Suppl. 1): 273-276, 1997.
  70. R. MacLeod and D. Brooks, Recent progress in inverse problems in electrocardiography. *IEEE Eng. Med. Biol.*, **17**: 73-81, 1998.
  71. A. M. Dale and M. L. Sereno, Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: A linear approach. *J. Cog. Neurosci*, **5**: 162-176, 1993.
  72. K. Matsuura and Y. Okabe, Selective minimum-norm solution of the biomagnetic inverse problem. *IEEE Trans Biomed Eng.*, **42**: 608-616, 1995.
  73. K. Matsuura and Y. Okabe, A robust reconstruction of sparse biomagnetic sources. *IEEE Trans Biomed Eng.*, **44**: 720-726, 1997.
  74. K. H. Uutela, M. S. Hamalainen, and E. Somersalo, Spatial and temporal visualization of magnetoencephalographic data using minimum-current estimates. *NeuroImage*, **5**: 543-4, 1997.
  75. M. Fuchs et al., Possibilities of functional brain imaging using a combination of MEG and MRT, in C. Pantev (ed.), *Oscillatory Event-Related Brain Dynamics*. New York: Plenum Press, 1994, pp. 436-457.
  76. K. Brackner et al., Magneto-cardiography and head potential mapping: The equalization in normal subjects during pharmacologically induced stress. *J. Cardiovasc. Electrophysiol.*, **8**: 615-626, 1997.
  77. D. Cohen and P. N. Cuffin, A method for combining MEG and PET to determine the source. *Phys Med Biol*, **32**: 85-89, 1987.
  78. M. Fuchs et al., Improving source reconstructions by combining bioelectric and biomagnetic data. *Electroenceph. Clin. Neurophysiol*, **107**: 93-111, 1999.
  79. P. Colli-Franzone et al., Potential fields generated by oblique dipole layers modeling excitation wavefronts in the anisotropic myocardium. Comparison with potential fields elicited by paced dog hearts in a volume conductor. *Circ. Res.*, **51**: 330-346, 1982.
  80. J. P. W. Kow, Jr., Tissue Anisotropy, the Cardiac Bidomain, and the Virtual Cathode Effect, in D. Zipes and J. Jalife (eds.), *Cardiac Electrophysiology: From Cell to Bedside*, 2nd ed., Orlando, FL: 1994. Saunders, pp. 346-361.
  81. P. Ratsman and I. Mugan, Modeling 3d deformable object with the active pyramid. *Int. J. Pattern Recognition & Artif. Intell.*, **11**: 1129-1139, 1997.
  82. G. Simpson et al., Spatiotemporal Mapping of Brain Activity Underlying Visual Attention Through Integrated MEG, EEG, fMRI and MRI in C. Aine et al. (eds.), *Advances in Biomagnetic Research*, **Bioimage96**, Springer Verlag, New York, 1996.
  83. J. P. Mäkelä, Functional differences between auditory cortices of the two hemispheres revealed by whole-head neuro-magnetic recordings. *Hum. Brain Mapp.*, **1**: 43-56, 1993.
  84. R. Salmeila and R. Hari, Characterization of spontaneous MEG rhythms in healthy adults. *Electroencephalogr Clin. Neurophysiol*, **91**: 237-248, 1994.
  85. R. Salmeila and R. Hari, Spatiotemporal characteristics of sensorimotor neuro-magnetic rhythms related to thumb movement. *Neuroscience*, **60**: 527-550, 1994.
  86. W. W. Surberling and D. S. Barré, Neocortical propagation in temporal lobe spike foci on magneto-encephalography and electro-encephalography. *Ann. Neurol.*, **25**: 373-381, 1989.
  87. R. Porsari et al., Magnetoencephalographic localization of epileptic cortex—impact on surgical treatment. *Ann. Neurol.*, **32**: 106-109, 1992.
  88. N. Forsyth et al., Trigeminally triggered epileptic hemifacial convulsions. *NeuroReport*, **6**: 913-920, 1995.
  89. J. Kinnanen et al., Noninvasive magneto-cardiographic localization of ventricular pre-excitation in walf parkinson-white syndrome using a realistic torso model. *Eur Heart J.*, **14**: 168-174, 1993.
  90. M. Mäskjärvi et al., Magneto-cardiography, Supraventricular arrhythmias and pre-excitation syndromes. *Eur Heart J.*, **14** (Suppl. E): 46-52, 1993.
  91. R. Porsari et al., Clinical validation of three-dimensional cardiac magnetic source imaging accuracy with simultaneous magneto-cardiographic mapping, monophasic action potential recordings, and magnetic cardiac pacing. *JNH Int Conf Biomagnetism, Biomaz98*, Abstracts, Sendai, 1998, p. 119.
  92. T. Oostendorp, J. Nenonen, and G. Haukamp, Comparison of inverse solutions obtained from ECG and MCG maps. *Proc. 19th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, pp. CD-rm, 1996.
  93. C. Leyber et al., Non-invasive biomagnetic imaging in coronary artery disease based on individual current density maps of the heart. *Int. J. Cardiol.*, **64**: 83-92, 1998.
  94. L. Tralins et al., Biomagnetic functional localization of a peripheral nerve in man. *Biophys J*, **56**: 1145-1153, 1989.
  95. R. Hari et al., Multichannel Detection of Magnetic Compound Action Fields of Median and Ulnar Nerves. *Electroenceph. Clin. Neurophysiol.*, **72**: 237-250, 1989.

96. I. Hashimoto et al., Visualization of a moving quadrupole with magnetic measurements of peripheral nerve action fields. *Electroenceph. Clin. Neurophysiol.*, **83**: 459-467, 1991.
97. H.-M. Mackert et al., Mapping of tibial nerve evoked magnetic fields over the lower spine. *Electroenceph. Clin. Neurophysiol.*, **104**: 322-327, 1997.
98. W. O. Richards et al., Non-invasive magnetometer measurements of human gastric and small bowel electrical activity. in C. Baumgartner et al. (eds.), *Biomagnetism: Fundamental Research and Clinical Applications*, New York: Springer-Verlag, 1995, pp. 743-747.
99. W. Weischieb et al., High-resolution monitoring of the gastrointestinal transit of a magnetically marked capsule. *J. Pharm. Sci.*, **86**: 1219-1222, 1997.
100. D. M. Schmidt, J. S. George, and C. C. Wood, Bayesian inference applied to the electromagnetic inverse problem. Tech. Report, Los Alamos National Laboratory, Los Alamos, LA-UR 97-4813.
101. D. Drung et al., Integrated  $\text{YBaCuO}_7$  magnetometers for biomagnetic measurements. *Appl. Phys. Lett.*, **68**: 1421-1423, 1996.
102. M. Burghoff et al., Diagnostic application of high-temperature SQUIDs. *J. Clin. Eng.*, **21**: 62-66, 1996.
103. J. M. ter Brake et al., A seven-channel high-T<sub>c</sub> SQUID-based heart scanner. *Meas. Sci. Technol.*, **8**: 327-331, 1997.

#### Reading List

- H. Weinberg, U. Strunk, and T. Katila, Biomagnetism, in J. G. Webster (ed.), *Encyclopedia of Medical Devices and Instrumentation*, Vol. 1, New York: Wiley, 1998, pp. 303-322.
- S. J. Williamson et al., *Advances in Biomagnetism*, New York: Plenum, 1989.
- M. Hoke et al., *Biomagnetism, Clinical Aspects*, Amsterdam: Elsevier, 1992.
- C. Baumgartner et al., *Biomagnetism: Fundamental Research and Clinical Applications*, New York: Springer-Verlag, 1995.
- C. Aine et al., *Advances in Biomagnetism Research: Biomag96*, New York: Springer-Verlag, 1999.

M. S. HÄMMLÄINEN  
DuPont Superconductivity  
J. T. NENONEN  
Helsinki University of Technology

## MAGNETS FOR MAGNETIC RESONANCE ANALYSIS AND IMAGING

Nuclear magnetic resonance (NMR) was discovered in 1946 by Purcell (1) and Bloch (2). Classically, it is the precession of the spins of nuclei with magnetic moment, subjected to a transverse radio frequency (RF) field in the presence of a longitudinal magnetic field. Nuclear species of biological interest having nonzero magnetic moment are listed in Table 1 together with their Larmor precession frequency-field dependence.

Experimentally, NMR is performed as follows (3,4): nuclei are immersed in a static field  $B_0$ , which results in the development of a net polarization along the field direction, occurring with a time constant  $\tau_1$ . Transitions among their spin states are excited by a high-frequency field  $B_1$ , oriented perpendicular to the static field. In a rotating frame of reference rotating at the Larmor precession frequency, the magnetization vector of the spin is tilted from a longitudinal direction (along  $B_0$ ) toward the transverse plane. The angle of precession depends on

Table 1. The Larmor Precession Constant for Various Nucleides

| Nucleide   | Atomic Number | NMR Frequency (MHz/Tesla) |
|------------|---------------|---------------------------|
| Hydrogen   | 1             | 42.5758                   |
| Deuterium  | 2             | 6.5357                    |
| Carbon     | 13            | 10.705                    |
| Oxygen     | 17            | 5.773                     |
| Sulfur     | 23            | 11.362                    |
| Phosphorus | 31            | 17.256                    |

the strength and duration of the applied RF field and is given by

$$\gamma = \gamma B_1 \tau \quad (1)$$

where  $\tau$  is the pulse width and  $\gamma$  is the gyromagnetic ratio.

Following the pulse, the magnetization decays transversely with a time constant  $\tau_2$ , the spin-spin relaxation time. The polarization develops (or decays) along the field with a time constant  $\tau_1$ , the spin-lattice relaxation time.

NMR is the preeminent method for the identification of chemical species in weak solution. It also has useful applications in solid materials. The most exacting specifications for an NMR magnet are imposed by high-resolution NMR. The resonant frequency of a nucleus depends not only on  $B_0$  but also, to a small extent, on the shielding provided by the electronic structure of the chemical compound. This effect is the chemical shift and is distinctive for each chemical species. Thus the resonant frequency of the  $^1\text{H}$  nucleus in water is different from that in benzene ( $\text{C}_6\text{H}_6$ ) or in the methyl or methylene group in alcohol ( $\text{CH}_2\text{CH}_2\text{OH}$ ). These small differences in frequency are typically a few parts per million and provide a means to identify the components of a complex molecule.

Early NMR spectrometers used continuous wave (CW) methods in which the frequency of the  $B_1$  field would be changed slowly and the absorption of a tank circuit enclosing the sample would be recorded as a spectrum of power absorption versus frequency. At the resonant frequency a sharp increase in absorption would be observed. The width of the peak depended, among other things, on the magnification  $Q$  of the tank circuit.

In modern NMR, a pulse of RF of sufficient strength and duration is applied to the sample so that all the spins are excited. The pulse is then switched off, and the signals emitted at various frequencies by the sample during relaxation of the spins are monitored. A Fourier analysis of the signal then transforms the time-dependent spectrum into a frequency-dependent signal, thus revealing the resonance peaks associated with the chemical shifts (3,4).

The uniformity of the static field  $B_0$  is the key to high-resolution NMR and to sharp images in magnetic resonance imaging (MRI). A uniform field allows large numbers of nuclei to precess at exactly the same frequency, thus generating a strong signal of narrow bandwidth. The underlying theory and practice of high-homogeneity superconducting magnets is described in this section. MRI magnets differ from those used for NMR analysis in that spatial distribution of either signal strength or  $\tau_1$  or  $\tau_2$  relaxation times are measured over a volume far greater than that of a sample for chemical analysis. In MRI, the predominant nuclear species examined is hydrogen.

in water. Density of  $\tau_1$  or  $\tau_2$  is measured on planes throughout a body and reconstructed as two-dimensional maps.

The field strengths of NMR magnets are higher than those of MRI magnets. From the discovery of NMR, in 1946, to 1967, magnetic fields were limited to 2 T that could be generated by electromagnets. A 5 T superconducting magnet was introduced in 1967, and slow improvements led to the 20 T magnetic field available in NMR magnets today. The driving forces for the increase in field strength are the chemical shift (separation of the nuclear species which is linearly proportional to field strength), signal strength (which is proportional to the square of the field strength), and signal-to-noise ratio (which is proportional to the 3/2 power of the field strength). Although various experimental techniques have been applied to improve signal-to-noise ratio, including fast-scan pulse sequences, signal averaging, cooled conventional receiver coils and superconducting receiver coils, increased field strength is still desirable for increased chemical shift. Magnets for high-resolution NMR are now almost exclusively superconducting, and it is only that type that is described here.

The optimal field strength for MRI is determined by a number of factors, including reduction in imaging time, reduction in chemical shift artifacts, and reduction in cost, and by limits to the exposure of a patient to electromagnetic radiation, as set by regulations. Even though most MRI magnets have field strengths up to 1.5 T, a few experimental magnets have been built or designed for functional MRI studies with field strengths up to 8 T. In order to achieve the desired combination of field strength, working volume, and stability, superconducting magnets represent the principal type employed. However, both water-cooled resistive magnets, iron-cored electromagnets, and rare earth permanent magnet MRI systems have been used or are in use for special applications. The superconducting magnet is here considered exclusively.

## DESIGN PRINCIPLES

### NMR Magnets

The analysis of weak solutions imposes several requirements on the magnet. In order to obtain usable signal-to-noise ratios, a large volume of sample must be used. This immediately demands good field uniformity so that variations in background field strength do not give rise to different frequencies, which would, of course, mask the small chemical shifts being sought. High field strength is desired as detailed above. Even if these requirements are met, the dilution of the sample may be such that repeated pulses are required. The final signal-to-noise that can be obtained from a number of pulses is proportional to  $\sqrt{N}$ . A run may take many hours or even days to accomplish. During that time not only must the spatial homogeneity of the background field be excellent, but the magnitude of the field also must be constant, or at least must change only very slightly. (The reason that any change is permitted is that a frequency lock can be used to adjust the frequency of the RF to match a slow and slight change in the background field.)

To summarize, the NMR magnet should have high field strength and great uniformity, and the field must be stable.

### MRI Magnets

The essential principles for MRI magnets are identical to those for NMR magnets, but the volumes of homogeneity are much

greater, whereas the homogeneity is somewhat lower. Despite field strengths lower than those of NMR magnets, the stored magnetic energies of MRI magnets are greater by reason of the large bore, which must be sufficient to house correction coils, pulsed gradient coils, and a patient. The MRI system differs significantly from that of the NMR system by including means to superimpose linear field gradients on the background homogeneous field. These pulsed gradients define thin planes in which the field is known but different from that elsewhere. Thus the frequency of nuclear magnetic resonance is spatially encoded so that the signals generated by the relaxing nuclei have frequencies which define their position. As in NMR for chemical analysis, the MRI signals may interrogate either the density of nuclei or the  $\tau_1$  or  $\tau_2$  relaxation times.

## THEORETICAL DESIGN

Almost all superconducting NMR and MRI magnets are solenoids. The reason for that is the relative simplicity and ease of manufacture and design of solenoids, compared with, for instance, extended dipoles. Although the generation of the RF field could be simpler with a transverse background field, the difficulty of manufacture of a high-background field magnet would far outweigh any advantage in the RF coil. The construction of a high-homogeneity solenoid proceeds in three parts: a winding array is designed, based solely on the analysis of the axial variation of the field of a solenoid, the magnet is wound and the spatial variation of its actual field is measured; and the unwanted errors in the field arising from manufacturing imperfections are removed by shimming.

The center field of a solenoid is given by

$$B_0 = \mu_0 J_0 \ln \left[ \frac{a + (a^2 - \rho^2)^{1/2}}{b} \right] (1 + (1 + \beta^2)^{-1/2}), \quad (2)$$

where  $J$  is the overall winding current density,  $a_0$  is the inner radius,  $a$  is the ratio of outer to inner radii, and  $\beta$  is the ratio of length to inner diameter ( $l/d_0$ ). Because SI units are used throughout,  $\mu_0$  is the permeability of free space,  $4\pi \times 10^{-7}$  H/m,  $a_0$  is in meters, and  $B_0$  is in Tesla.

The field strength decreases at points on the axis away from the center of the solenoid. The axial variation of field strength on the  $z$  axis is expressible as a Taylor's series

$$B(z) = B_0 + (d^2 B/dz^2)z^2/2 + (d^4 B/dz^4)z^4/4! - (d^6 B/dz^6)z^6/6! + \dots \quad (3)$$

Only even terms appear because the center of the solenoid coincides with the origin.

Figure 1 illustrates the geometry of a thin solenoid, of radius  $a_0$ , and extending a length  $z_0$  to the right of the origin. For such a thin solenoid, the derivatives of the field at the origin are as follows:

$$\begin{aligned} B_0 &= \frac{1}{2} \mu_0 c z_0 (\pi a_0^2 + z_0^2)^{-1/2} \\ d^2 B/dz^2 &= \frac{1}{2} \mu_0 c \pi a_0^2 (\pi a_0^2 + z_0^2)^{-3/2} \\ d^4 B/dz^4 &= -\frac{1}{2} \mu_0 c \pi a_0^2 (\pi a_0^2 + z_0^2)^{-5/2} \\ d^6 B/dz^6 &= \frac{1}{2} \mu_0 c \pi a_0^2 (3\pi a_0^2 - 12z_0^2) \pi a_0^2 + z_0^2)^{-7/2} \end{aligned} \quad (4)$$

where  $c$  is the sheet current density in amp-turns per meter and  $a_0$  and  $z_0$  are as illustrated in Fig. 1.

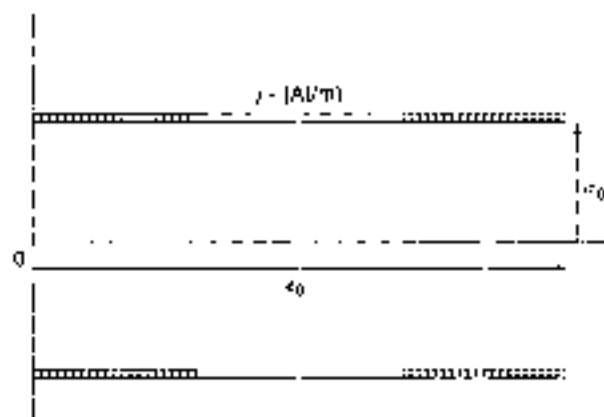


Figure 1. Geometry of a thin solenoid showing the coordinate system used to define current geometries.

The field of a solenoid symmetric about the center plane has even symmetry and no odd derivatives. So, by evaluating the even derivatives of the field at the center, the axial variation of field generated by a solenoid can be calculated to an accuracy determined by the number of derivatives used and by the distance from the center. The derivatives can be treated as coefficients of a Cartesian harmonic series so that

$$B_z = B_0 + b_2 z^2 + b_4 z^4 + b_6 z^6 + \dots$$

in which

$$\begin{aligned} b_2 &= \mu_0 j_0 (3z_0 a_0^2 / (a_0^2 + z_0^2)^{3/2}) / 4 \\ b_4 &= \mu_0 j_0 (45a_0^2 z_0^2 - 60z_0^3 / (a_0^2 + z_0^2)^{5/2}) / 48 \\ b_6 &= -\mu_0 j_0 (15a_0^4 z_0 - 20a_0^2 z_0^3 + 8z_0^5 / (a_0^2 + z_0^2)^{7/2}) / 1440 \end{aligned} \quad (5)$$

For coils of odd symmetry, such as shim coils described later, the corresponding harmonics are

$$\begin{aligned} b_1 &= \mu_0 j_0 (a_0^2 + z_0^2)^{-3/2} / 2 \\ b_3 &= \mu_0 j_0 (3a_0^2 z_0 - 12z_0^3 / (a_0^2 + z_0^2)^{3/2}) / 12 \end{aligned} \quad (6)$$

Notice that the magnitude of any harmonic coefficient is mediated by the denominator of the expressions that each include the term  $(a_0^2 + z_0^2)^{-n/2}$ , where  $n$  is the order of the harmonic. Thus, the generation of high-order harmonics requires coils with large values of current (ampere-turns) or small radii. This is significant in the construction of shim coils, as is noted later.

Associated with an axial variation of field is a radial variation, arising from radial terms in the solution of the Laplace scalar potential equation. For instance, even-order axial variations are accompanied by axisymmetric radial variations (6) of the form

$$\begin{aligned} B_z(z, x, y) &= b_2(z^2 - \frac{1}{2}(x^2 + y^2)) \\ B_4(z, x, y) &= b_4(z^4 - 3(x^2 + y^2)z^2 + \frac{1}{2}(x^2 + y^2)^2) \end{aligned} \quad (7)$$

These equations show that if  $b_2$  or  $b_4$  are zero there will be no axisymmetric radial variation of field.



Figure 2. Principle of harmonic compensation. Coaxial solenoids generating field harmonics of opposite sign.

Figure 2 illustrates a set of nested solenoids. Solenoid 1 gives rise to nonzero values of the harmonic coefficients  $b_2, b_4, b_6, \dots$ . If dimensioned correctly, solenoid 2 by contrast can produce equal values for some or all these coefficients but with opposite polarity. Then at least  $b_2$  and  $b_4$  will have net zero values, and the first uncompensated harmonic to appear in the expression for axial field variation will be the sixth order. A minimum, but not necessarily sufficient, condition is that as many degrees of freedom are needed in the parameters of the coils as there are coefficients to be zeroed.

This method can be extended to as many orders as desired. In most high-resolution NMR magnets, the required uniformity of the field at the center is achieved by nulling all orders up to and including the sixth. That is, the solenoid is of eighth order. In the design of the solenoids, no odd order appears, of course. The first residual harmonic will have a very small value close to the center, although at greater axial distances, the field will begin to vary rapidly. Thus, the design of a high-homogeneity solenoid requires only the calculation of the field or the field harmonics on axis, and those harmonics may be easily calculated using only Cartesian coordinates.

## MANUFACTURING ERRORS

The theoretical design of a high-homogeneity magnet can be simple because only axial terms in the  $z$  field need to be considered. However, the manufacturing process introduces errors in conductor placement which generate both even- and odd-order axial and, most significantly, radial field gradients. Further, the materials of the coil forms, the nonisotropic contraction of the forms and windings during cool-down to helium temperature and the effects of the large forces between the windings may also introduce inhomogeneity. Typically, the homogeneity of an as-wound set of NMR solenoids is not better than  $10^{-5}$  over a 5 mm diameter spherical volume (dsv) at the center. For high-resolution NMR, an effective homogeneity of  $10^{-8}$  over at least 5 mm dsv is required. The improvement of the raw homogeneity to this level is achieved by three steps, superconducting shim coils, room temperature shim coils and, in NMR magnets only, sample spinning. Additionally, in cases of poor raw homogeneity, ferromagnetic shims may be used occasionally in NMR magnets and routinely in MRI magnets to compensate for large errors or significant high-order harmonics. The presence of radial field gradients necessitates a more comprehensive field analysis than is convenient with Cartesian coordinates.

## LEGENDRE FUNCTIONS

The expression of the harmonics of the field in terms of Cartesian coordinates provides a simple insight into the source of the harmonics. However, as the order of the harmonic increases, the complexity of the Cartesian expressions renders manipulation very cumbersome, and an alternative method is needed. The Laplace equation for the magnetic field in free space is conveniently solved in spherical coordinates. These solutions are spherical harmonics, and they are valid only in the spherical region around the center of the solenoid, extending as far as, but not including, the nearest current element. Figure 3 illustrates the coordinate system for spherical harmonics. The convention followed here is that dimensions and angles without subscripts refer to a field point, and with subscripts they refer to a current source.

The axisymmetric  $z$  field generated by a coaxial circular current loop can be expressed in the form of a Legendre polynomial, thus,

$$B_z = \sum_{n=0}^{\infty} g_n r^n P_n(\cos\theta) \quad (8)$$

where  $r$  and  $\theta$  define the azimuth of the field point in spherical coordinates, and  $a$  is  $\cos\theta$ .  $P_n(a)$  is the zonal Legendre polynomial of order  $n$  and  $g_n$  is a generation function given by

$$g_n = \rho_0^n P_{n-1}(\cos\theta_0) \sin(\theta_0) / (2r_0)^{n+1} \quad (9)$$

where  $\theta_0$  and  $\rho_0$  define the position of the current loop in spherical coordinates. In this text, it is the convention that  $n = 0$  represents a uniform field. The field strength given by Eqs. (8) and (9) is constant with azimuth at constant radius  $r$ .

Equations (8) and (9) are equivalent in spherical coordinates to those of Eqs. (4), (5), and (7) on the  $z$  axis but additionally predict the  $z$  field off axis. In the design of the main coils Eqs. (8) and (9) offer no more information than Eq. (5). However, in the calculation of the off-axis  $z$  fields, they provide important additional information that can be used in the optimization of coil design when fringing fields must be considered.

The harmonic components of the  $z$  field can also be expressed in the form of associated Legendre functions of order  $n$ ,  $m$  (1).

These functions define the variation of the local  $z$  field strength at points around the center of the magnet and include variation of the field with azimuth  $\phi$ . Thus,

$$B_z(r, \theta) = r^n [n + (n - 1) P_n(a) / a] \\ \times [C_{n,m} \cos(m\phi) + S_{n,m} \sin(m\phi)] \quad (10)$$

where  $C_{n,m}$  and  $S_{n,m}$  are the harmonic field constants in tesla per meter <sup>$n$</sup> ,  $P_n(a)$  is the associated Legendre function of order  $n$  and degree  $m$ , and  $a$  is  $\cos\theta$ . The order  $n$  is zonal, describing the axial variation of  $z$  field. The degree  $m$  is tesseral, describing the variation of the  $z$  field in what would be the  $x - y$  plane in Cartesian coordinates.  $\phi$  is the azimuth to the point at radius  $r$  from an  $x = -z$  plane.  $\theta$  is the elevation of the point from the  $z$  axis. Tables of the values of the Legendre polynomials can be found in standard texts on mathematical functions (8).

In Eq. (10),  $m$  can never be greater than  $n$ . For example, if  $n = m = 0$ ,  $B_z(0,0)$  is a uniform field independent of position. If  $n = 2$  and  $m = 0$ ,  $B_z(2,0)$  is a field whose strength varies as the square of the axial distance (i.e.,  $B_z$  of Eq. (7)). If  $n = 2$  and  $m = 2$ ,  $B_z(2,2)$  is a field that is constant in the axial direction but increases linearly in two of the orthogonal radial directions and decreases linearly in the other two. Figure 4 shows a map of the contours of constant field strength of a  $B_z(2,2)$  field harmonic for which  $S_{2,2} = 0$ . The  $B_z(2,2)$  field has zero magnitude at the origin and along the  $x$  and  $y$  coordinate axes. Of course, the direction of the zero values of the  $B_z(2,2)$  harmonic will not generally lie in the Cartesian  $x$  and  $y$  planes. Depending on the relative values of  $C_{n,m}$  and  $S_{n,m}$  in Eq. (10), the zero harmonic planes will lie at an angle other than  $\phi = 0$  or  $\pi r/2$ . The constant field contours of  $B_z(2,2)$  extend to infinity along the  $z$  axis and represent, arbitrarily in this figure, values for  $B_z(2,2)$  of  $10^{-4}$ ,  $10^{-2}$ , and  $10^{-1}$ , for example. Within the indicated cylinder centered on the  $z$  axis, the value of the harmonic is everywhere less than  $10^{-6}$ . For higher values of  $m$ , there are more planes of zero value. Thus,  $B_z(4,4)$  has eight planes of zero value,  $B_z(8,8)$  has 16, and so forth. A harmonic  $B_z(4,2)$  defines

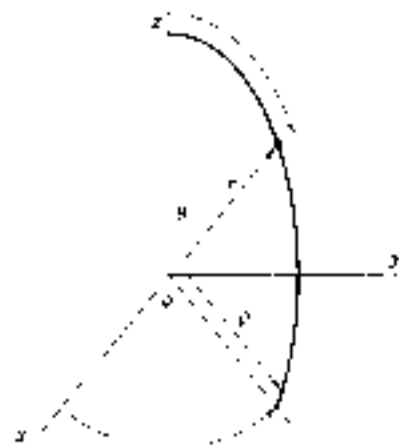


Figure 3. The system of spherical coordinates specifying field points and current sources.

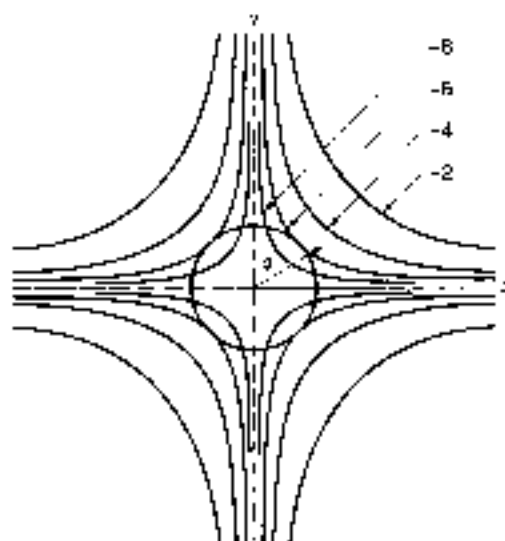


Figure 4. Surfaces of constant magnitude of a  $B_z(2,2)$  harmonic field, showing that the tesseral harmonic is zero when the azimuth  $\phi$  is a multiple of  $\pi/2$ .



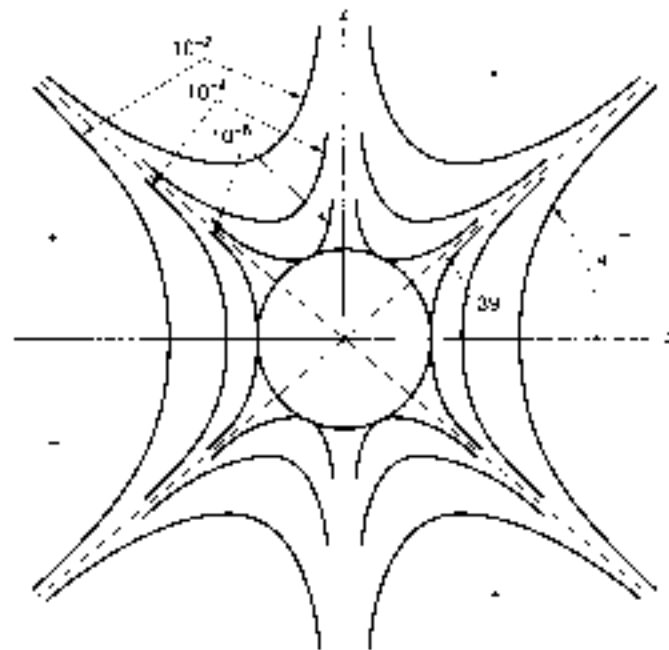


Figure 5. Surfaces of constant magnitude of a  $B_{3,0}$  harmonic field, showing that the zonal harmonic is zero when the elevation  $\mu$  is  $39^\circ$  or  $30^\circ$ .

a field in which the second-degree azimuthal variation itself varies in second order with axial distance.

The zonal harmonics  $B_1(2,0)$ ,  $B_2(3,0)$ ,  $B_3(4,0)$  have conical surfaces on which the value of the field is zero. Thus, for instance,  $B_2(3,0)$  has contours of zero value such as are shown in Fig. 5 to lie at  $\theta = 39^\circ$  and  $90^\circ$ .

The four hyperbolas are actually surfaces of rotation about the  $z$  axis, and each represents a constant value for  $B_2(3,0)$  of, say,  $10^{-3}$  (the uniform field,  $B_0(0,0)$  at the origin having unity value). Within the indicated ellipsoid, centered on the origin, the value of the  $B_2(3,0)$  harmonic is therefore everywhere less than  $10^{-3}$ .

The contours of the zero values of the spherical harmonics are analogous to combinations of Figs. 4 and Fig. 5. The zero values now lie on straight lines radiating from the origin. The surfaces of constant value look like the spines of sea urchins. As for the zonal harmonics, ellipsoidal surfaces roughly describe boundaries within which the magnitudes of the spherical harmonics do not exceed a given value. These error surface diagrams are often used in the design of an MRI magnet to identify the maximum calculated field error within a central volume caused by the highest uncompensated harmonic.

Thus, in general, the deviations from the ideal uniform solenoidal field can be expressed as the sum of a large number of harmonics each described by the associated Legendre function of order  $n$  and degree  $m$ . Although the Cartesian expressions of Eq. (14) can be used for the design of a coil system to generate a uniform field, the associated Legendre functions must generally be used for the analysis of the measured field and the design of shim coils or of ferromagnetic shims to compensate for harmonics with nonzero values of  $m$  (9).

#### Optimization Methods

With the recent rapid increase in the speed and size of computers, an alternative technique for the design of uniform field

magnets has been developed. Not only is a uniform field of specific magnitude required but that should be combined with other criteria. For instance it could be accompanied by the smallest magnet, that is, the minimum of conductor, or by a specified small fringing field. To achieve these ideal solutions, an optimization technique is now generally used. The field strength of a set of coils is computed at points along the axis, and, if fringing field is a consideration, at points outside the immediate vicinity of the system. The starting point may be a coil set determined by a harmonic analysis as described earlier. Now however, mathematical programming methods are employed to minimize the volume of the windings satisfying the requirement that the field should not vary by more than the target homogeneity for each of the chosen points. Again, for purposes of homogeneity, only field on axis is considered because the radial variation of axisymmetric components of field is zero if the axial component is zero. The field strengths at points outside the magnet will be minimized by inclusion of a set of coils of much larger diameter than the main coils but carrying current of reverse polarity.

All design techniques, but particularly that of optimization, are complicated by the highly nonlinear relationship between the harmonic components generated by a coil and the characteristics of the coil. Thus the reversal in sign of the harmonic components occurs rapidly as the dimensions or position of a coil are changed. In the example of an NMR magnet shown in Fig. 6, the value of the second harmonic changes by 4 ppm for an increase in the diameter of the wire in the small coil "1" of only  $\phi 1$  mm. The optimization of the ampere-turns, shape and position of a coil thus affects the various harmonics in highly nonlinear and often conflicting ways.

Design optimization involves the computation of an objective function which contains all the elements that have to be minimized, subject to a set of constraints (10). For example, it may be required to minimize some combination of winding volume or magnet length subject to constraints on the field error at a number of points within the bore and on the fringing field at some point outside the magnet. The objective function would then be of the form

$$\sum_{i=1}^N \rho V_i \cdot L \quad (11)$$

where  $V_i$  is the volume of a coil,  $N$  is the number of coils, and  $L$  is the length of the magnet. The factor  $\rho$  weights the relative

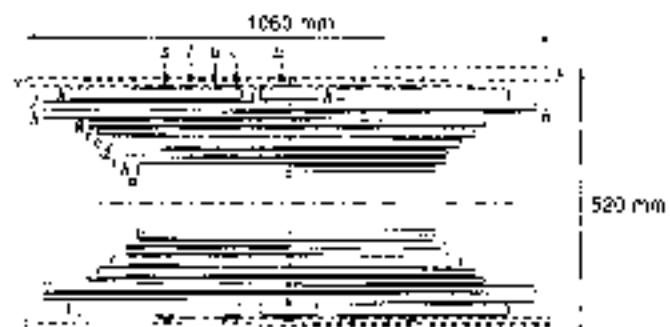


Figure 6. Coil profiles of an actual 8th order compensated NMR magnet. The graded sections a through j produce axial harmonics of which orders 2, 4, and 6 are compensated by sections k and l. Layers i and v are shim coils.

importance of the volume and of the length. This objective function is then minimized subject to the following constraints:

$$\left[ \sum_{c=1}^N B_{c,p} - B_0 \right]^2 < \Delta B^2 \quad (12)$$

$$\sum_{c=1}^N B_{c,p} < H_f \quad (13)$$

In Eq. (12)  $B_{c,p}$  is the field at point  $p$  due to coil  $c$ . The equation represents the constraint on uniformity of field. It could also be expressed in terms of harmonic terms; for example, each even term up to  $P_{17,2}$  being less than  $10^{-6} B_0$ , the center field. (The inclusion of the squared terms in Eq. (12) allows for either positive or negative error field components.) Equation (13) expresses the condition that the fringing field should be less than, say, 1 mT (10 gauss) at a point outside the magnet system. The 10 gauss criterion frequently represents the maximum field to which the public may be exposed in accessible areas around an MRI system.

The minimization of the objective function is performed by a mathematical programming algorithm, whereas the solution of the constraining Eqs. (12) and (13) will require a nonlinear technique (such as Newton-Raphson), in order to deal with the extremely nonlinear variation of the harmonics as they change with coil geometry (11).

## SHIELDING

The minimization of the external fringing field is becoming increasingly important for the siting of MRI systems, so the active shielding of MRI magnets with center fields up to 2 T is now almost universal. (Active shielding of MRI magnets with center fields above 2 T is uneconomical and is not generally attempted.) Active shielding is generally achieved by the inclusion in the coil array of two reverse polarity coils at diameters typically twice that of the main coils. Because of the large dipole moment of an MRI magnet, the unshielded fringing field will extend several meters from the boundary of the cryostat. Consequently, active shielding is applied to many MRI magnets with central fields of over 0.5 T (12). The effect of the shielding on the harmonics of the center field must, of course, be included in the design of the compensation coils.

## SHIMMING

The harmonic errors in the field of an as-built magnet divide into purely axial variations (axysymmetric zonal harmonics, which are accompanied by radial variations dependent on the elevation  $\theta$  from the  $z$  axis, but independent of  $\phi$ ) and radial variations (sectoral harmonics, which depend on  $\phi$ , where  $\phi$  is the angle of azimuth in the  $x - y$  plane).

In order to compensate for the presence of various unwanted harmonic errors in the center field of the as-built coils, additional coils capable of generating the opposite harmonics are applied to the magnet. For each set of  $n$  and  $m$  in the associated Legendre functions, a current array can be designed in the form of a set of arcs of varying azimuthal extent and symmetry and with various positions and extents along the  $z$  axis. The magnitude of the harmonic field that an array generates can be

controlled by the current. This is the principle of variable harmonic correction for both MRI and NMR magnets. (Correction by means of ferromagnetic shims is not variable.)

The shimming of the unwanted harmonics is a process in two independent parts. First, there is the design of as many sets of coils as are needed to generate the compensating harmonics. Second is the measurement of the actual field errors to determine the magnitudes of the various harmonic components and the application of currents to the previously designed coils to provide the compensation. In fact, because superconducting shims must be built into the magnet prior to installation in the cryostat and cooldown, the range of harmonic errors in the field of the as-built magnet must be largely anticipated. Typically it might be assumed that the level of harmonic error decreases by a factor of three for each unit increase in  $n$  or  $m$ . Therefore, as a rough guide it has been found that compensation of up to  $B(3,0)$  for the zonal harmonics and up to  $B(2,3)$  for the sectoral harmonics is satisfactory in most cases for the superconducting shims of small bore NMR magnets. There will also be a set of room temperature shims in a high resolution NMR system. These will compensate for errors typically up to  $B(6,0)$  and  $B(4,5)$  in many cases. Typically there may be up to 28, but exceptionally up to 45 independent shims in all. They will be constructed according to a different principle from the superconducting shims. The shimming of MRI magnets is accomplished by current shims, typically up to  $n = 3$  and  $m = 2$ , and by ferromagnetic shims.

## Superconducting Axial Shims

These will be simple circular coils combined in groups so as to generate a single harmonic only (13). Thus, a coil to generate  $B(3,0)$  must generate no  $B(1,0)$  nor  $B(5,0)$ . Because the superconducting shim coils need to generate only a small fraction of the field due to the main coil, they generally need only comprise one to three layers of conductor. For that reason the harmonic sensitivities can be calculated directly from Eqs. (4) and (5). A set of axial shims providing correction of  $B(n,0)$  harmonics for  $n = 1$  through 3 are shown in Fig. 7. Note that, for a fixed linear current density, only the angles defining the start and end of each coil are needed, together, of course, with the current polarities, either side of the center plane of the magnet, odd for  $n = 1, 3, 5, \dots$  and even for  $2, 4, 6, \dots$ .

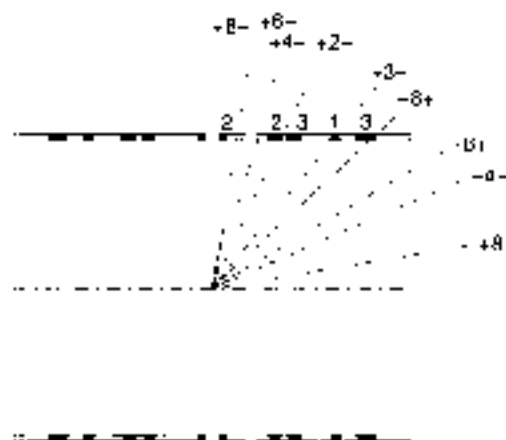


Figure 7. A set of axial shim coils for harmonic correction up to  $B(3,0)$ . These coils generate small harmonics of 4th order and higher.

The set of coils illustrated in Fig. 7 generate negligible harmonics above the third order,  $B(3,0)$ . The individual coils of each harmonic group are connected in series in sets, there being in each set enough coils to generate the required axial harmonic but excluding, as far as is practical, those harmonics that are unwanted. Thus, in the figure, coils labeled 2 generate second-order  $B(2,0)$  but no fourth order. However, they do generate higher orders. The first unwanted order is  $B(6,0)$  but that is small enough that it may be neglected. So also with all higher orders because the denominator in the expressions of Eqs. 14) and (5) strongly controls the magnitude of the harmonic. Also illustrated in the figure is the effect on harmonic generation of the angular position of a circular current loop. Each of the dashed lines lies at the zero position of an axial harmonic. Thus, at an angle of  $70.1^\circ$  from the  $z$  axis, the  $B(4,0)$  harmonic of a single loop is zero. Two loops carrying currents of the same polarity and suitable magnitude may be located on either side of the  $70.1^\circ$  line to generate no fourth-order harmonic yet generate a significant second order harmonic. Similarly, a coil for the generation of only a first order axial harmonic is located on the line for zero third order. The zero first order harmonic line is at  $90^\circ$ , the plane of symmetry. In order therefore to generate a third order with no first, two coils must be used, with opposing polarities. The coils are all mirrored about the plane of symmetry, but the current symmetries are odd for the odd harmonics and even for the even harmonics. The loops may be extended axially as multiturn coils while retaining the property of generating no axial harmonic of a chosen order, if the start and end angles subtended by the coils at the origin are suitably chosen.

The principles described earlier can be applied both in the design of shim coils and in the selection of main coil sets. A further observation from the zero harmonic lines of Fig. 7 is that the higher harmonics reverse sign at angles close to the plane of symmetry of the system. This implies that, to produce single, high-order harmonics, coil positions close to the plane of symmetry must be chosen because the other coil locations where the sign of the harmonic reverses are too far from the plane of symmetry to be usable; the coils lying a long way from the plane of symmetry generate weak high-order harmonics.

### Superconducting Radial Shims

The radial shims are more complex than those for purely axial harmonics because the finite value of  $m$  requires a  $2m$  field symmetry in the azimuthal distribution of current arcs, the polarity of current always reversing between juxtaposed arcs in one  $z$  plane (6,9). For instance,  $m = 2$  requires four arcs, as shown in Fig. 8. However, as for  $m = 0$ , the set of current arcs shown in Fig. 8 will generate  $B(n,m)$ , where  $n$  is 2, 4, 6, etc., or 1, 3, 5, etc., depending on even or odd current symmetry about the  $z = 0$  plane. So, the positioning of the arcs along the  $z$  axis is again crucial to the elimination of at least one unwanted order,  $n$ . Fortunately, the azimuthal symmetry generates unique values of the fundamental radial harmonic  $m$ . (Eight equal arcs cannot generate an  $m = 2$  harmonic.) However, depending on the length of the arc, higher radial harmonics may be generated. For the shim coil configuration of Fig. 8, the first unwanted radial harmonic is  $m = 6$ . The higher tesseral harmonics are much smaller than the fundamental because of the presence in the expression for the field of a term  $(r/r_0)^2$ . Generally, the arc length is chosen to eliminate the first higher-degree radial harmonic. As an example, if the arc length of each shim coil shown

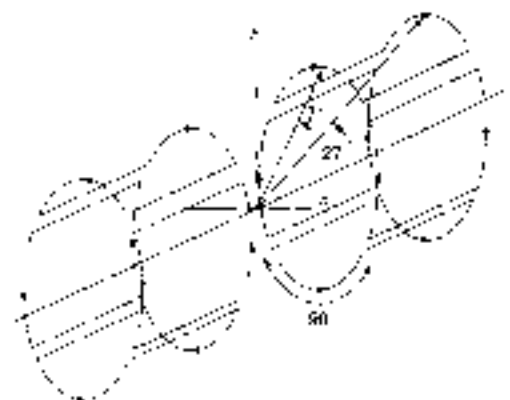


Figure 8. Schematic of a set of radial shim coils for correction of a  $B(2,2)$  harmonic showing the positioning necessary to eliminate  $B(4,2)$  and  $B(4,4)$ .

in Fig. 8 is  $90^\circ$  the  $B(6,6)$  harmonic disappears. The  $B(10,10)$  harmonic is negligible.

The superconducting shims are almost invariably placed around the outside of the main windings. Although the large radius reduces the effective strength of the harmonics they generate, the shim windings cannot usually be placed nearer to the center of the coil because of the value of winding space near the inner parts of the coil and because of the low critical current density of wires in that region due to the high field. A comprehensive treatment of shim coil design may be found in Refs. 6 and 9. Those references also include details of superconducting coil construction. It should be noted, however, that some expressions in Ref. 6 contain errors.

### Ferromagnetic Shims

Ferromagnetic shimming is occasionally used in high field, small bore NMR magnets, but its principal use is in MRI magnets. It is in that application that it will be described. The principle invoked in this kind of shimming is different from that of shim coils. The shims now take the form of discrete pieces of ferromagnetic material placed in the bore of the magnet. Each piece of steel is subjected to an axial magnetizing field at its position sufficient to saturate it. It then generates a field at a point in space that is a function of the mass of the shim and its saturation magnetization  $H_s$ , with little dependence on its shape. For ease of example, a solid cylinder of steel will be assumed. The axis of the cylinder is in line with the field, as shown in Fig. 9. (In Fig. 9 the axis labeled  $z$  is that of the shim, not that of the MRI magnet itself. In fact, the shim will usually be placed at the inside surface of the bore of the MRI magnet.)

The field  $H_s$  caused by the ferrimagnetic shim, contains both axial and radial components. The axial component  $H_z$  is the correcting field required, and it adds arithmetically to the field of the magnet. The radial component adds vectorially to the field and produces negligible change in the magnitude of the axial field. Therefore, only the axial component of the shim field must be calculated. If the saturation flux density of the shim is  $B_s$ , the axial shim field is given by

$$B_z = BV[2 - \tan^2 \alpha] / (\tan^2 \alpha + 1)^{3/2} (1/4\pi z^3) \quad (14)$$

where  $V$  is the volume of the shim and  $z$  and  $\alpha$  are as shown in Fig. 9.

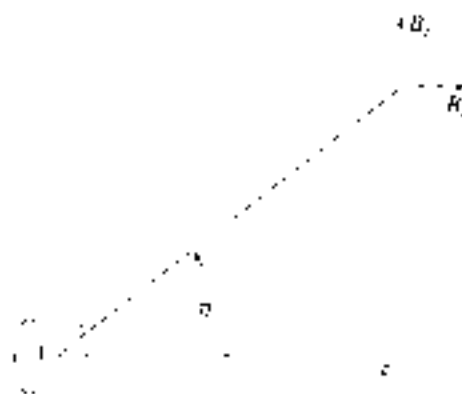


Figure 9. Field vectors generated by a ferrimagnetic shim in the bore of an NMR magnet.  $B_2$  adds arithmetically to the main field,  $B_1$  adds vectorially and so has negligible influence on the field.

The practical application of ferromagnetic shims involves the measurement of the error fields at a number of points, and the computation of an influence matrix of the shim fields at the same points. The required volumes (or masses) of the shims are then determined by the inversion of a  $U$ ,  $W$  matrix, where  $U$  is the number of field points and  $W$  is the number of shims. In an MRI magnet, the shims are steel washers (or equivalent) bolted to rails on the inside of the room temperature bore of the cryostat. In the occasional ferromagnetic shimming of an NMR magnet, the shims are coupons of a magnetic foil pasted over the surface of a nonmagnetic tube inserted into the room temperature bore or, if the cryogenic arrangements allow, onto the thermal shield or helium bore tube. As in the design of the magnet, linear programming can be used to optimize the mass and positions of the ferromagnetic shims (e.g., to minimize the mass of material).

#### Resistive Electrical Shims

The field of an NMR magnet for high resolution spectroscopy must be shimmed to at least  $10^{-7}$  over volumes as large as a 10 mm diameter cylinder of 20 mm length. If, as is usually the case, substantial inhomogeneity arises from high-order harmonics ( $n$  and  $m$  greater than 1), superconducting shims are of barely sufficient strength. This arises because of the large radius at which they are located, as in NMR magnets (e.g., in the regions  $x$  and  $y$  of Fig. 6). In general, the magnitude of a harmonic component of field generated by a current element is given by

$$B_n \propto r^{n-1}/r_0^{n+1} \quad (15)$$

where  $r$  is the radius vector of the field point, and  $r_0$  is the radius vector of the source. Thus, the effectiveness of a remote source is small for large  $n$ .

In order to generate useful harmonic corrections in NMR magnets for large  $n$  and  $m$ , electrical shims are located in the warm bore of the cryostat. Although in older systems these electrical shims took the form of coils tailored to specific harmonics, modern systems use matrix shims. Essentially, the matrix shim set consists of a large number of small saddle coils mounted on the surface of a cylinder. The fields generated by unit current in each of these coils form an influence matrix, similar to that

of a set of steel shims. The influence matrix may be either the fields produced at a set of points within the magnet bore, or it may be the set of spherical harmonics produced by appropriate sets of the coils.

#### FIELD MEASUREMENT

The accurate measurement of the spatial distribution of field in the  $\mu$ s-wound magnet is essential to shimming to high homogeneity. Sometimes, measurement of the field is possible at very low field strengths with tiny currents flowing in the windings at room temperature. That may allow mechanical adjustment of the positions of the main compensation coils ( $k - k$  in Fig. 6) to reduce the  $B_{1,0}$ ,  $B_{2,0}$  and  $B_{1,1}$  harmonics. Major field measurement is made with the magnet at design field strength and in persistent mode. The methods of measurement in NMR and MRI magnets are generally different.

In NMR magnets, because of the small bore, the field is measured by a small NMR probe on the surface of a cylindrical region about 8 mm diameter and over a length of up to 10 mm. The measurements are made at typically 20 azimuthal intervals. From these field measurements, the predominant harmonics can be deduced, using a least-squares fit, and shimmed by means of the superconducting shim coils, both axial and radial. With subsequent measurements, as the harmonic content becomes smaller, the higher harmonics become evident and in turn can be shimmed. The field measurements are usually reduced to harmonic values because the shim sets are designed to generate specific harmonics. The correcting current required in any particular shim set is then immediately determined. Measurement and shimming is always an iterative process, generally requiring several iterations to achieve homogeneities of better than  $10^{-6}$  over 5 mm dia.

Field measurement in an MRI magnet is usually performed differently because much more space is available and because knowledge of the magnitudes of the harmonics in associated Legendre polynomial form is an advantage in the shimming process. In this case, the measuring points will lie on the surface of a sphere. Typically, the diameter of this sphere may be 500 mm. The field is measured at intervals of  $\phi$ , often  $30^\circ$ , around each of the circles of intersection with this spherical surface of several  $z = \text{const}$  planes, called Gauss planes. From these measurements and by the property of orthogonality of the associated Legendre functions, the values of the constants  $C_{nm}$  and  $S_{nm}$  in Eq. (10) can be deduced by the following methodology.

The double integral

$$\int_{-1}^{+1} \int_0^{2\pi} P_n(m)u(\cos\theta\cos\phi)P_l(j)u(\cos\theta\sin\phi)du d\phi$$

is nonzero only if  $i = n$  and  $j = m$ . Then, for  $m > 0$ , its value is

$$2\pi(n+m)!/2^n(n-m)!$$

So, if both sides of Eq. (10) are multiplied by  $P_l(j)u(\cos\theta\sin\phi)$  or  $P_l(j)u(\sin\theta\sin\phi)$  and double integrated, the right-hand side will

be nonzero only if  $l = n$  and  $j = m$ . Then

$$\int_0^{2\pi} \int_0^{\pi} B_z(r, \theta; P_l, \mu_l) |\cos j\phi| d\theta d\phi \\ = C_{l,m} r^n (n + m + 1) 2\pi (n + m)! (2n + 1)! (n - m)!$$

and

$$\int_0^{2\pi} \int_0^{\pi} B_z(r, \theta; P_l, \mu_l) |\sin j\phi| d\theta d\phi \\ = S_{l,m} r^n (n - m + 1) 2\pi (n + m)! (2n - 1)! (n - m)! \quad (16)$$

Equation (16) are realized in practice by the measurement of the field at each of 60 points (for example) and the multiplication of each value by the spherical harmonics,  $P_l(\mu) \cos j\phi$  and  $P_l(\mu) \sin j\phi$  at that point. The integration is numerical. The method usually employed is Gaussian quadrature, similar in principle to Simpson's rule for numerical integration in Cartesian coordinates, but in which the  $z = \text{const}$  planes are the roots of the Legendre polynomial and the weights assigned to the values measured on each of these planes are derived from the Lagrangian. Tables of the roots and weights are found in standard texts on numerical analysis (14). For the purposes of example, assume the number of planes  $p = 5$ , and the number of azimuthal points per plane  $q = 12$ , for a total of 60 points on the surface. The planes are at  $z/r = 0$ ,  $z/r = \pm 0.5386$ , and  $z/r = \pm 0.9662$ ,  $r$  is the radius of the spherical surface. The corresponding weights are 0.5689, 0.4786, and 0.2769. The measurements are made on the circles of intersection and two numerical integrations of Eq. (16) performed, one for the  $\cos(j\phi)$  terms and the other for the  $\sin(j\phi)$  terms. Then the values of  $C_{l,m}$  and  $S_{l,m}$  are obtained from

$$C_{l,m} = \left[ \sum_p \sum_q w_p B_{l,m}(r_p, \theta_p) P_{l,m}(u_p) \cos m\phi_p \right] \\ \times [(2n + 1)(n - m)! / (2\pi r^n (n + m + 1))] \quad (17)$$

$$S_{l,m} = \left[ \sum_p \sum_q w_p B_{l,m}(r_p, \theta_p) P_{l,m}(u_p) \sin m\phi_p \right] \\ \times [(2n + 1)(n - m)! / (2\pi r^n (n - m + 1))] \quad (18)$$

where  $B_{l,m}(r_p, \theta_p)$  is the field at the point  $p, q$ , the subscripts  $p$  and  $q$  denote each of the 60 points, and  $w_p$  is the Gaussian weighting for the plane  $q$ .

## NMR MAGNET DESIGN AND CONSTRUCTION

Practical issues peculiar to the design and construction of NMR magnets include the following:

The wire diameter must be such that layers of windings near the inner radius of the solenoids do not generate discrete field fluctuations of a size comparable to the desired homogeneity. For example, if the wire diameter is very large, say, greater than 3 mm, the field will develop a fine structure away from the  $z$  axis. If the winding lay of a large diameter wire is helical in each layer, a helical

structure may arise in the amplitude of the field with consequent problems in the correction of the resulting high harmonics.

Nonmagnetic coil forms must be used because the presence of discrete regions of ferro- or strong paramagnetism will generate large harmonics of high order (large  $n$  and possibly also  $m$ ), which would be very difficult to shim.

The index of the wire must be high. All high-resolution NMR requires high field stability, with a decay not exceeding about  $10^{-4}$  per hour. To achieve that, the magnets operate in persistent mode. A superconducting switch is closed across the winding after energization so that the current flows without loss in a resistanceless circuit. The superconducting switch consists of a small coil (usually noninductive) of a superconducting wire equipped with a resistance heater. When the heater is energized, the temperature of the coil is raised above the critical value, and the coil becomes resistive. The charging voltage applied to the magnet then causes only a small current to flow in the switch. When the magnet has been charged, the switch heater is turned off, the coil cools, and the magnet current can then flow through the switch without loss. If a magnet does not need frequent resetting, its rate of field decay must be small. The joints between wire lengths and the switch and the magnet must be superconducting, and the wire must be without resistance. Although the joints can indeed be made so that their critical currents exceed the operating current, the effective resistance of the wire, owing to its index, may be high enough that decay in persistent mode exceeds acceptable levels for NMR. The resistance of the wire, manifest as a low value of the index, arises from variation in the critical current along the length of the wire. If a short region exists where the superconducting filaments are thin or have low pinning strength, a fraction of the current transfers between superconducting filaments through the copper (or bronze) matrix, giving rise to the resistance. The voltage associated with this resistance appears in critical current measurements on samples of the wire.

Figure 10 shows the typical trace of voltage gradient along a superconducting wire in a fixed field as a function of current. The defined critical current is that at which a voltage gradient of, typically, 0.1  $\mu\text{V}/\text{cm}$  is measured. As the current is increased beyond the critical value, that voltage gradient increases. An

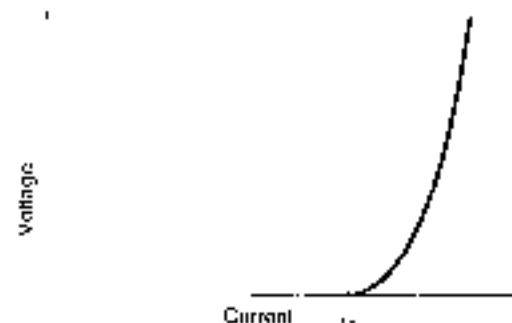


Figure 10. Voltage gradient along a composite superconductor as a function of steady current, showing the effect of index.

approximation to the gradient near to the critical current is

$$v \sim i(i_c)^N \quad (19)$$

where  $i/i_c$  is the ratio of actual current to critical current and  $N$  is the index of the wire. The higher the value of  $N$ , the sharper is the superconducting to normal transition. Clearly, for values of  $i$  below the critical value, the voltage gradient will be small; the larger the index, the smaller the gradient. So, for NMR magnets, an appropriate combination of index and the ratio of  $i/i_c$  must be chosen. The index of most niobium-titanium (NbTi) wires available for MRI and NMR is typically 50. However, for niobium-tin (Nb<sub>3</sub>Sn) wires, the index is lower, typically around 30, and the matrix is the more resistive bronze. So, for high field NMR magnets using Nb<sub>3</sub>Sn inner sections, lower ratios of  $i/i_c$  are necessary. The concept of the index is only an approximation to the behavior of voltage as a function of current. In fact, theory and measurement indicate that the effective index increases as  $i/i_c$  decreases below 1 (15,16). Field decay arising from the index is constant and is distinguished from that caused by flux creep. The latter is a transient effect. It dies away with a logarithmic time dependence after a magnet has been set in persistent mode.

Protection of the magnet from the consequences of quenching must be compatible with the electrical and thermal isolation of the magnet from room temperature systems. Quenching is an spreading irreversible transition from the superconducting to the normal resistive state in the winding. The energy released during quenching in an NMR or MRI magnet must be dissipated as heat in the winding. In order to limit the energy and hence heat dissipated in any part of the winding, the magnet must be electrically divided into sections, each of which is provided with a shunt, often in the form of diodes. This subdivision limits the energy that can be transferred between sections and thereby minimizes the temperature rise and voltage generated within a section during quenching (17).

### NMR Magnet Design

Figure 6 illustrates the winding array of a typical 750 MHz NMR magnet (18). Table 2 specifies the dimensions and winding specifications of the sections.

At a current of 307.86 A, these windings generate 17 616 T at the center; that corresponds to 750 MHz proton resonance frequency. The total inductance is 109.2 H, and the stored energy is 5.17 MJ. The first nonzero harmonic of the design is the 12th. The coils  $x$ ,  $t$ ,  $z$ ,  $r$ , and  $u$  and their mirror images are the axial shim coils located in the annular space  $x$ . The radial shim coils are located in the space  $y$ .

The winding of the Nb<sub>3</sub>Sn sections of high field NMR magnets presents particular problems. The wire is wound in the unreacted state after which it must be heated at about 700°C for up to 200 h to transform the separate niobium and tin components into the superconducting compound. The wire is insulated with S-glass resin, with a softening temperature of about 1000°C. (An alternative is E-glass. Although the E-glass may start to soften during the heat treatment, it is stronger in the prefired state than S-glass and therefore better survives the exigencies of winding.) After the heat treatment the winding is consolidated by impregnation with epoxy resin.

The forms on which the Nb<sub>3</sub>Sn wire is wound must also endure the heat treatment without distortion. Stainless steel is the universal choice for the coil forms although titanium alloys

have been used. The alloy 316 L is generally preferred because of its very small magnetic susceptibility. If the form is assembled with welds, those must be made with nonmagnetic filler, if used. The inner bore of the form must be quite thick if distortion is not to occur. The reason for that lies in the expansion coefficients of the wire and of stainless steel. The unreacted Nb<sub>3</sub>Sn wire consists of bronze, niobium, tin, tantalum, and copper. During reaction, the copper and bronze have negligible strength, and the mechanical properties of the niobium and tantalum dominate. Their coefficients of thermal expansion are smaller than that of stainless steel with the consequence that, as the temperature rises during the heat treatment, the bore of the form will expand faster than the inner diameter of the winding. If the bore tube is thin, it can buckle against the constraint of the winding.

The need for thick bore tubes leads to windings of several wire diameters on one form. The thick bore tube occupies space that could otherwise be used by field-generating winding. In order to minimize the diluting effect of the bore tube, large winding builds are used. However, in the high field regions of Nb<sub>3</sub>Sn windings, the wire diameter must be graded to optimize the cross section of Nb<sub>3</sub>Sn corresponding to the local field. An alternative to the thick stainless steel bore tube is the transfer of the reacted winding to an aluminum form before impregnation with epoxy resin. This has been used occasionally, as in the example of Fig. 8 (18).

### MRI Magnet Design

The design construction of MRI magnets follows the principles involved in the construction of NMR magnets (19). The forces and energies are generally greater. For example, the force tending to center each large end coil of the MRI magnet illustrated in Fig. 11 is 1,339,000 N (150 tons). NbTi conductor is used exclusively in MRI magnets because, to date, center fields of no more than 5 T are used. The NbTi filaments in the copper matrix of composite NbTi wires are twisted to approximate transposition. Because of the low fields in MRI magnets, wires with few NbTi filaments can be used. Those filaments can then be arrayed as a single circular layer within a copper matrix. The filaments are then fully transposed and are magnetically very stable. Mechanical perturbation is nevertheless a problem, and attention has to be paid to the interface between a winding and the coil form against which it presses. Because of the high stored energies, large copper cross sections are needed

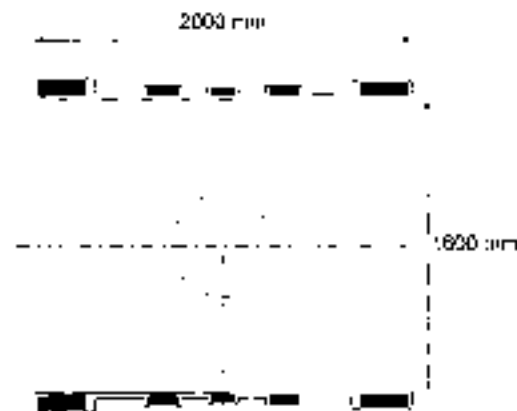


Figure 11. Coil profile of a 1.5 T unshielded MRI magnet illustrating typical coil placement.

**Table 2. Dimensions of the Windings of a 700 MHz NMR Magnet**

| Section Number | Peak Field (T) | Wire Type          | Wire Diameter (mm) | Inner Diameter (mm) | Outer Diameter (mm) | Winding Length (mm) | Number of Turns |
|----------------|----------------|--------------------|--------------------|---------------------|---------------------|---------------------|-----------------|
| a              | 17.62          | Nb <sub>3</sub> Sn | 2.4                | 43                  | 52.2                | 600                 | 1000            |
| b              | 16.97          | Nb <sub>3</sub> Sn | 2.22               | 52.2                | 65.3                | 600                 | 1629            |
| c              | 16.93          | Nb <sub>3</sub> Sn | 1.64               | 64.8                | 79.9                | 650                 | 1770            |
| d              | 14.8           | Nb <sub>3</sub> Sn | 1.63               | 89.4                | 102.5               | 650                 | 1780            |
| e              | 13.83          | Nb <sub>3</sub> Sn | 1.63               | 99.0                | 115.2               | 700                 | 3870            |
| f              | 11.60          | Nb <sub>3</sub> Sn | 1.61               | 121.7               | 134.8               | 760                 | 4194            |
| g              | 9.48           | NbTi               | 1.41               | 141.3               | 149.5               | 800                 | 3398            |
| h              | 8.31           | NbTi               | 1.39               | 158.0               | 171.0               | 1000                | 6949            |
| i              | 5.70           | NbTi               | 1.14               | 176.0               | 188.2               | 1000                | 6892            |
| j              | 2.13           | NbTi               | 1.30               | 194.2               | 212.5               | 88.7                | 1020            |
| k/l            | 0.14           | NbTi               | 1.30               | 194.2               | 212.5               | 377.1               | 4110 (each)     |

At a current of 9700 A these windings generate 11.6 T at the center, that corresponds to 500 MHz proton resonance frequency. The total inductance is 109.2 H and the stored energy is 5.17 MJ. The best current-to-magnetic field ratio of the design is 10.7/0.01.

in the conductor to avoid over heating during quenching. Currents are typically up to 600 A. A common form of conductor is a composite wire embedded in a copper carrier. The latter frequently has a grooved rectangular cross section into which the composite wire is pressed or soldered. Insulation may be cotton or kapton wrap instead of enamel, and wax may be used as an impregnant as an alternative to epoxy.

Most whole-body MRI magnets used in clinical applications have room temperature bores of between 1 and 1.5 m, with fields up to 2 T. An example of the profile of the windings of a whole-body MRI magnet is illustrated by the simple five-coil system of Fig. 11.

The center field is 1.5 T and the dimensions of the windings are listed in Table 3. The compensation is to tenth order (10 ppm over a 500 mm sphere). The current is 394 A for a 1.5 T center field. The inductance is 78 H and the stored energy 6 MJ.

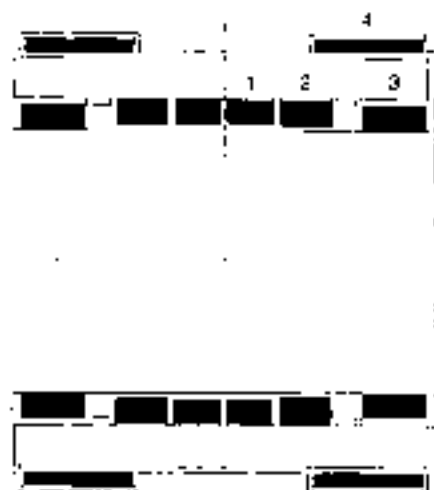
The fringing field of this magnet extends a long way from the cryostat in which the coils are housed. The 1 mT (10 gauss) line is at an axial distance of 11.3 m and at a radial distance of 8.8 m from the center. Access to the space within these limits must be restricted because of the dangers to the wearers of pacemakers, the attraction of ferromagnetic objects, and the distortion of video monitors.

This may be an expensive restriction in a crowded hospital. Therefore, methods of shielding the space from the fringing fields are frequently used. Three methods are generally available, close iron, remote iron, and active shielding. The use of iron close to the coils has been used in a few instances. However, the iron must be at room temperature, to avoid otherwise severe cryogenic penalties. That leads to difficulties in balancing the forces between the coils and the iron in order to minimize the loads that the cryogenic supports must resist.

Remote iron takes the form of sheet, typically several millimeters thick, placed against the walls of the MRI room. This involves rather awkward architectural problems but is used frequently where the restricted space can still extend several meters from the cryostat.

The third form of shielding is by superconducting coils, built around the main coils, operating in series with the main coils as part of the persistent circuit, and in the same cryogenic environment. These shield coils generate a reversed field to cancel, or reduce, the external fringing field. Typical of the resulting magnet is the eight-coil configuration shown schematically in Fig. 12.

Particular aspects of the illustration follow. The outwardly directed body forces in unshielded MRI windings are supported in tension in the conductor. However, in the shielded version, these forces are too large to be supported by the conductor alone, and a shell is applied to the outside of the winding against which the accumulated body forces act. Thus, the body forces on coils 3 and 4 are supported on their outer surfaces by a structural cylinder. Coil 4 provides the compensation of the dipole moment of the three inner windings so that the fringing fields of the magnet are much reduced from those of the unshielded



**Figure 12.** Schematic of an actively shielded MRI magnet showing the large coils needed to generate a main field while the shield coils generate an opposing field.

**Table 3. Example of a 1.5 T Whole-Body MRI Magnet**

| Coil Number | Inner Radius (mm) | Outer Radius (mm) | Left End (mm) | Right End (mm) | Number of Turns |
|-------------|-------------------|-------------------|---------------|----------------|-----------------|
| 1           | 741.9             | 807.1             | -977.0        | 724.9          | 3030            |
| 2           | 742.2             | 785.7             | -891.9        | -244.8         | 1180            |
| 3           | 742.1             | 777.0             | -743.9        | 74.9           | 960             |
| 4           | 742.3             | 785.7             | 244.0         | 191.9          | 1180            |
| 5           | 741.9             | 807.1             | 724.9         | 977.0          | 3030            |

magnet. The reduction in the volume of the restricted space is about 90%. The magnet is much heavier (and more expensive) than the simple unshielded type and the structural design of the cryostat and the suspension system is accordingly stronger. The highest fields generated at the center of shielded whole-body MRI magnets is 2 T. See also Ref. 12.

## CRYOGENICS

As for NMR magnet systems, the economic operation of superconducting MRI magnets demands cryogenic systems with low heat in-leak. The evolution of MRI cryostats has been significant over the past 15 years. They have changed from simple liquid helium, liquid nitrogen shielded reservoirs with relatively high cryogen evaporation rates to single or multi-cryocooled cryostats. In one embodiment, no refrigerant is used in some types of cryocooled MRI magnets; in other examples, a combination of cryocoolers and refrigerants provide a zero evaporation rate. Dismountable current leads are an essential feature of any magnet system with low refrigerant evaporation rate, and have been a standard feature of MRI magnet systems since 1974. An implication of dismountable current leads is the need for the MRI magnet to be self-protecting during quenching, just as an NMR magnet must be.

## PULSED GRADIENT COILS

In addition to the uniform background field, which it is the function of the MRI magnet to generate, pulsed gradient fields must be superimposed on that field in order to create the spatial encoding of the resonant frequencies of the protons (or other species) within the body. Those pulsed gradient fields are generated by three sets of room temperature coils, each set being driven by a powerful ramped current source. The pulsed field gradients are linear ( $dB/dz$ ,  $dB/dx$ ,  $dB/dy$ ), as far as it is possible to design pure first-order gradient coils. Two problems arise in the overall design as a consequence of these pulsed gradient fields. The first is the effect on the superconductor of the periodic incident fields. Although the thermal shields and coil forms lie between the gradient coils and the superconductor, the incident pulsed field would still be significant there. Those small fields would cause a loss within the conductor through the mechanisms of hysteresis and coupling.

The second is the distortion of the gradients arising from currents induced in adjacent structures, such as the thermal shields and helium vessel, also sometimes the room temperature bore tube, if that is metallic. This distortion is minimized by locating sets of shield coils near the room temperature bore tube. These active shield coils are energized in opposition to the main pulsed gradient coils. They serve to confine the return flux of the gradient coils to flow in the space between the main gradient coils and the shield coils. The eddy currents induced in the surrounding structures are thereby minimized. The shield coils reduce the efficiency of the pulsed gradient system, that effect becoming more pronounced as the diameter of the main coils becomes a large fraction of that of the shield coils. At a diameter ratio greater than about 0.55, the efficiency is so reduced that the driving power required for useful gradient fields becomes prohibitively large. Figure 13 illustrates the form of the shielded  $dB/dx$  or  $dB/dy$  pulse coils. The  $dB/dz$  coils are simple solenoids surrounded by shielding solenoids.

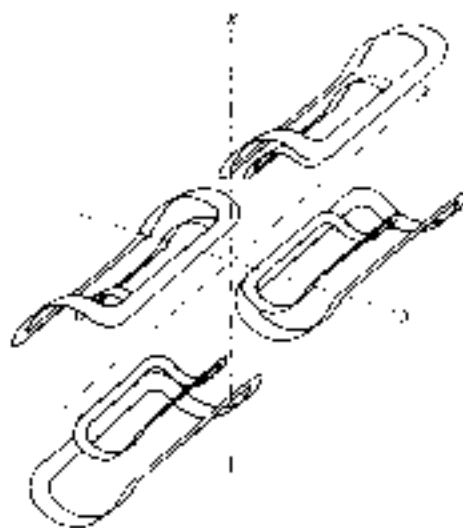


Figure 13. Schematic of actively shielded pulsed gradient coils for the  $dB/dx$  gradient showing the spacing between the main and shield coils.

## BIBLIOGRAPHY

1. E. M. Purcell, H. C. Torrey, and R. V. Pound, Resonance absorption by nuclear magnetic moments in a solid, *Phys. Rev.* **69** 37, 1946.
2. F. Bloch, W. W. Hansen, and M. Packard, Nuclear induction, *Phys. Rev.* **69**, 127, 1946.
3. E. Becker, *High resolution NMR. Theory and Applications*, New York: Academic Press, 1960.
4. R. R. Ernst, G. Bodenhausen, and A. Wokaun, *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, Oxford: Clarendon Press, 1987.
5. D. B. Montgomery, *Solenoid Magnet Design*, New York: Wiley-Interscience, 1969, p. 4.
6. M. D. Saxe and S. K. Kan, High resolution nuclear magnetic spectroscopy in high magnetic fields, *Adv. Electron. Electron. Phys.*, **34** 1-93, 1973.
7. W. R. Smythe, *Static and Dynamic Electricity*, New York: McGraw-Hill, 1950, pp. 147-148.
8. M. Abramowitz and I. A. Stegun (eds.), *Handbook of Mathematical Functions*, Washington, DC: US Dept. of Commerce, Natl. Bur. of Standards, 1968.
9. P. Romeo and D. I. Huft, Magnetic field profiling: Analysis and correcting coil design, *Magn. Resonance Med.*, **1**: 44-65, 1984.
10. M. K. Thompson, H. W. Brown, and V. C. Srivastava, An inverse approach to the design of MRI main magnets, *IEEE Trans. Magn.*, **MAG-30**: 106-112, 1964.
11. W. H. Press et al., *Numerical Recipes. The Art of Scientific Computing*, Cambridge, UK: Cambridge Univ. Press, 1987.
12. P. J. Davies, R. T. Elliott, and D. G. Hawksworth, A 2 Tesla active shield magnet for whole body imaging and spectroscopy, *IEEE Trans. Magn.*, **MAG-27**: 1677-1690, 1991.
13. E. S. Rohrer and W. F. B. Pritchard, A general method of design of axial and radial shim coils for NMR and MRI magnets, *IEEE Trans. Magn.*, **MAG-24**: 523-526, 1968.
14. P. Davis and P. Rubinowitz, Abcissae and weights for Gaussian quadrature of high order, *J. Res. NBS*, **61B**(2545), **AMS**(55): 35-37, 1956.
15. Y. Iwasa, *Case studies in Superconducting Magnets*, New York: Plenum, 1994, pp. 306-307.



16. J. E. C. Williams et al., NMR magnet technology at MIT, *IEEE Trans Magn.* **MAG-28**: 627-630, 1992.
17. P. J. Maddock and G. B. James, Protection and stabilisation of large superconducting coils, *Proc Inst. Electr. Eng.* **116**: 743-746, 1968.
18. A. Zhukovsky et al., 750 MHz NMR magnet development, *IEEE Trans Magn.* **MAG-28**: 644-647, 1992.
19. D. G. Hawksworth, Superconducting magnets systems for MRI, *Int. Symp. New Develop. in Appl. Superconductivity*, Singapore: World Scientific, 1989, pp. 731-744.

JOHN E. C. WILLIAMS  
Massachusetts Institute of  
Technology

## MICROWAVE FERROELECTRIC DEVICES

Control of reactance in microwave circuits, devices, and systems is a common method by which the response of a microwave circuit such as a filter, resonator, or phase shifter can be tuned. Devices based on a class of voltage-dependent nonlinear dielectrics known as ferroelectrics provide an alternative to semiconductor varactor diodes and ferromagnetic components which are the most common devices of this type.

After an overview of the relevant materials issues, this article describes microwave devices that exploit the variation of the ferroelectric's permittivity with applied dc electric field. The main feature of these "tunable" microwave devices is the change of their capacitance, impedance, or phase velocity. We will describe varactors, oscillators, tunable filters, and phase shifting devices.

## FERROELECTRIC MATERIALS

Even though they do not contain iron, the name ferroelectric was selected because they possess a response to an electric field that, although not the dual, is analogous to a ferromagnetic material's response to a magnetic field. Ferroelectrics are a subgroup of nonlinear dielectrics. The complex permittivity of a ferroelectric material is a function of both the temperature and an applied dc electric field.

Ferroelectric materials possess spontaneous polarization below a temperature referred to as the Curie temperature, at which point they undergo a phase transition. Above the Curie temperature, they are in a paraelectric state where spontaneous polarization disappears, but they still retain a nonlinear dielectric constant with applied electric field. In this article, we will use the term ferroelectric to describe these materials even if they are being operated at temperatures where they are in the paraelectric phase. Unlike ferromagnetic materials, ferroelectric materials in either the ferroelectric or paraelectric state are reciprocal; that is, the transmission coefficient through these devices is the same for different directions of propagation. Until recently, the dielectric losses in ferroelectric materials excluded them from being used at microwave frequencies. However, the continued improvement of ferroelectric materials suitable for use at microwave frequencies has resulted in the design of many microwave devices.

Tunability can be defined for a ferroelectric as the fractional change in the dielectric constant with applied dc bias voltage

or

$$\text{Tunability} = \frac{\epsilon_{r,\text{max}} - \epsilon_{r,\text{min}}}{\epsilon_{r,\text{max}}} \quad (1)$$

where  $\epsilon_{r,\text{min}}$  is the dielectric constant when no bias voltage is applied, and  $\epsilon_{r,\text{max}}$  is the dielectric constant when maximum dc bias is applied. The dielectric constant of a ferroelectric decreases as the bias voltage is increased. Although larger tunability is a desirable feature for most microwave applications, a ferroelectric material with larger tunability usually has a relatively larger dielectric loss. Optimizing material tunability and loss to meet the needs of a particular microwave application remains a challenging task.

Ferroelectrics are inherently broadband. That is, they do not have a low frequency limit like ferrites. Switching time for these materials has been measured to be less than a nanosecond (1), which is sufficient for most microwave applications. Also, these materials are radiation-hardened. Ferroelectrics can be manufactured in bulk, thick-film, and thin-film form.

Like other ceramics, bulk ferroelectric ceramics can handle high peak powers. The limit on the average power is determined by the loss tangent ( $\tan \delta$ ) of the ferroelectric. There are many known ferroelectrics, but the most widely used ferroelectric at microwave frequencies is barium strontium titanate,  $\text{Ba}_{1-x}\text{Sr}_x\text{TiO}_3$  (BSTO). BSTO with  $x = 0.5$  is frequently used for microwave applications since the Curie temperature is well below room temperature yet a reasonable tunability is retained. Bulk ceramics of this composition typically possess relative dielectric constants in the order of 1000 and loss tangents of 0.02 at 10 GHz (2,3). BSTO's Curie temperature can be controlled by varying the barium to strontium ratio. A bulk composite material can be engineered by adding nonferroelectric oxides to BSTO to reduce the dielectric constant and the loss tangent (2,3). For room-temperature operation of these composites in the paraelectric phase, tunability,  $\epsilon_r$ , and  $\tan \delta$  decrease with decreasing barium content; they also decrease with increasing oxide content. Tunability increases linearly with an increase in bias voltage. Bulk ceramics can be produced using usual ceramic processing techniques.

Thin films are compatible with integrated circuits, and they need lower bias voltages than does bulk material. Thin films can be manufactured by any of the common thin-film deposition techniques, pulsed-laser deposition, sputtering, metal organic chemical vapor deposition (MOCVD), and so on. In thin films, control of the ferroelectric composition and incorporation of doping is also possible to reduce losses at microwave frequencies (4). Thin films can potentially be less costly and easier to manufacture, but they cannot handle high power levels. Thin films of the ferroelectric strontium titanate,  $\text{SrTiO}_3$  (STO), are used at microwave frequencies because of their compatibility with the high-temperature superconductor (HTS) yttrium barium cuprate (YBCO).

Between bulk and thin-film ferroelectrics lies the realm of thick-film ferroelectrics, which can be produced via tape casting.

Throughout this article, it will be assumed that the ferroelectric is homogeneous and that it is linear with respect to a small, time-varying electric field. The dielectric strength of these ferroelectrics is relatively high. Large dc electric fields (in the range of a few MV/m) can be applied to STO and BSTO before dielectric breakdown occurs. Assuming that a particular

ferroelectric composition meets the tunability,  $\epsilon_r$ , and  $\tan \delta$  requirements of the application, the next two sections describe the various issues that need to be addressed before designing a microwave ferroelectric device.

## MICROWAVE DEVICE CONSIDERATIONS

Most microwave devices can be categorized according to their physical size in relationship to the wavelength at their upper frequency of operation. Those devices that are very small compared to their operational wavelength are called "electrically small" and can be modeled with discrete circuit components. The term "lumped-element" model is often employed. For devices whose dimensions are larger, it is usually necessary to take into account the frequency-dependent effects. Such devices are described by distributed networks of common circuit elements such as inductors and capacitors. The most common distributed device is the transmission line which is modeled by a ladder network of series inductors and shunt capacitors. As a lumped element, a capacitor which uses ferroelectric material yields a tunable circuit reactance.

Transmission lines can usually be described by their phase velocity and characteristic impedance, which are given by

$$v_c = \frac{1}{\sqrt{LC}} \quad (2)$$

and

$$Z_c = \sqrt{\frac{L}{C}} \quad (3)$$

respectively, where  $L$  is the inductance per unit length and  $C$  is the capacitance per unit length. By introducing a material whose dielectric constant,  $\epsilon_r$ , is controlled (or tuned) by a direct current (dc) bias voltage, the phase velocity and characteristic impedance can be varied by changing capacitance.

An issue to consider when choosing a microwave device topology is the power-handling requirements. For high-power applications, the number of device topologies that are appropriate is limited. Further discussion is offered in the next section.

In any nonlinear material, device, or system, another important practical consideration is the strength of signals generated at other than the desired frequency. Since most systems are bandwidth-limited, the most troublesome condition arises when two desired signals,  $f_1$  and  $f_2$ , both within the passband produce signal at frequencies  $2f_1 - f_2$  and  $2f_2 - f_1$ , also within the passband. A plot of the signal strengths of  $f_1$ ,  $f_2$ ,  $2f_1 - f_2$ , and  $2f_2 - f_1$  is often used to determine the third-order intercept point (IP3), which is an important figure of merit.

Since ferroelectrics have a high dielectric constant, the circuits that employ these materials tend to have very low impedance. Therefore, impedance matching is also another major issue to be addressed when using ferroelectrics. A consequence of a voltage-dependent capacitance being utilized to tune the phase velocity of a transmission line as given by Eq. (2) is that the characteristic impedance of the transmission line is also tuned per Eq. (3). This further complicates the impedance matching problem.

## MICROWAVE GUIDING STRUCTURES

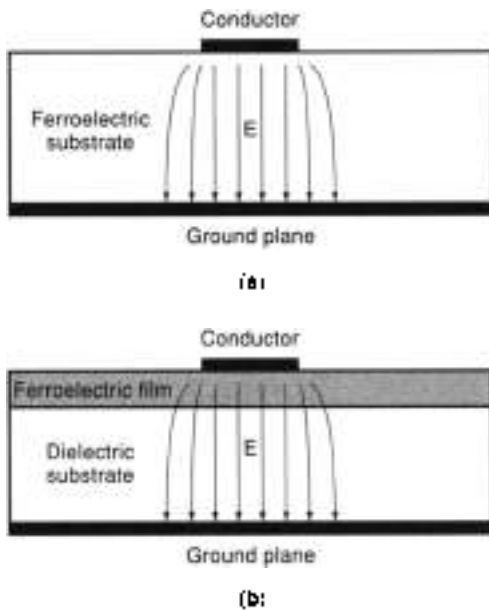
At microwave frequencies, ferroelectrics can be introduced into many different types of rectilinear structures that are used to guide electromagnetic waves. These guiding structures include parallel plate and rectangular waveguides, which can be loaded (or filled) with ferroelectric material. There are also many planar structures that use the ferroelectric material as a tunable substrate, like microstrip, slotline, coplanar strip, and coplanar waveguide. Each structure has a different set of advantages and disadvantages. For applications where ferroelectrics are used to provide bias-dependent propagation properties, it is convenient to divide the guiding structures into two categories: (1) geometries that can handle high microwave power but require large bias voltages (parallel plate and rectangular waveguide) and (2) those which are compatible with small microwave power levels and require only modest bias voltages (planar structures). Note that the ferroelectric permittivity is a function of the electric field. In a planar structure, the bias voltage is applied across a thinner ferroelectric, and so smaller bias voltage will produce the same variation of the permittivity as a larger bias voltage (which creates a similar electric field intensity) would produce in a parallel plate or rectangular waveguide.

Planar waveguiding structures contain the metallization defining the waveguiding structure delineated on one or more planes. Often this metallization layer is on the top surface of a dielectric substrate. Hence, these geometries are compatible with photolithographic processing. Since the metallization delineating the waveguide is at the interface of two regions (usually dielectric), the guided wave is propagating such that a portion of the field is in each region. The choice of the correct planar transmission line is determined by many factors including (1) orientation of the bias field and microwave field with the ferroelectric region, (2) thickness of the ferroelectric material, and (3) compatibility with other circuit elements. Planar structures are compatible with ferroelectric thin films and with semiconductors for microwave monolithic integrated circuits. However, planar structures require a dc block to isolate the radio frequency (RF) from the high dc voltage that is used to tune the ferroelectric permittivity (5).

### Microstrip

The most common planar transmission line is microstrip. As shown in Fig. 1, both the bias and the dominant mode microwave electric field are oriented primarily normal to the interface. It should be noted that although microstrip is the most widely used planar transmission line, when using thin-film ferroelectrics deposited on a substrate with the delineated metal layer on top as shown in the Fig. 1(b), the high dielectric constant of the ferroelectric results in decreased tuning efficiency. This can be seen by considering the capacitance per unit length to be a series combination of the ferroelectric capacitor and the substrate capacitor. The capacitance contribution from the thin-film ferroelectric is much smaller than that from the substrate, and the tunability of the phase velocity and the characteristic impedance of the dominant mode are reduced accordingly.

Tunable filters employing parallel-coupled microstrip resonators do not suffer from this inefficiency since the coupled-line mode (6) possesses a significant electric field component parallel to the surface. This coupled-line mode is similar to



**Figure 1.** Microstrip planar transmission line cross section with (a) a homogeneous bulk substrate and (b) a thin film on a bulk substrate. The electric field between the metal strip on the top surface and the metal ground plane is primarily normal to the surface.

the coplanar strip transmission line (discussed below) with a ground plane. Practical design equations for microstrip on layered dielectric substrates are based on a quasi-static analysis (7).

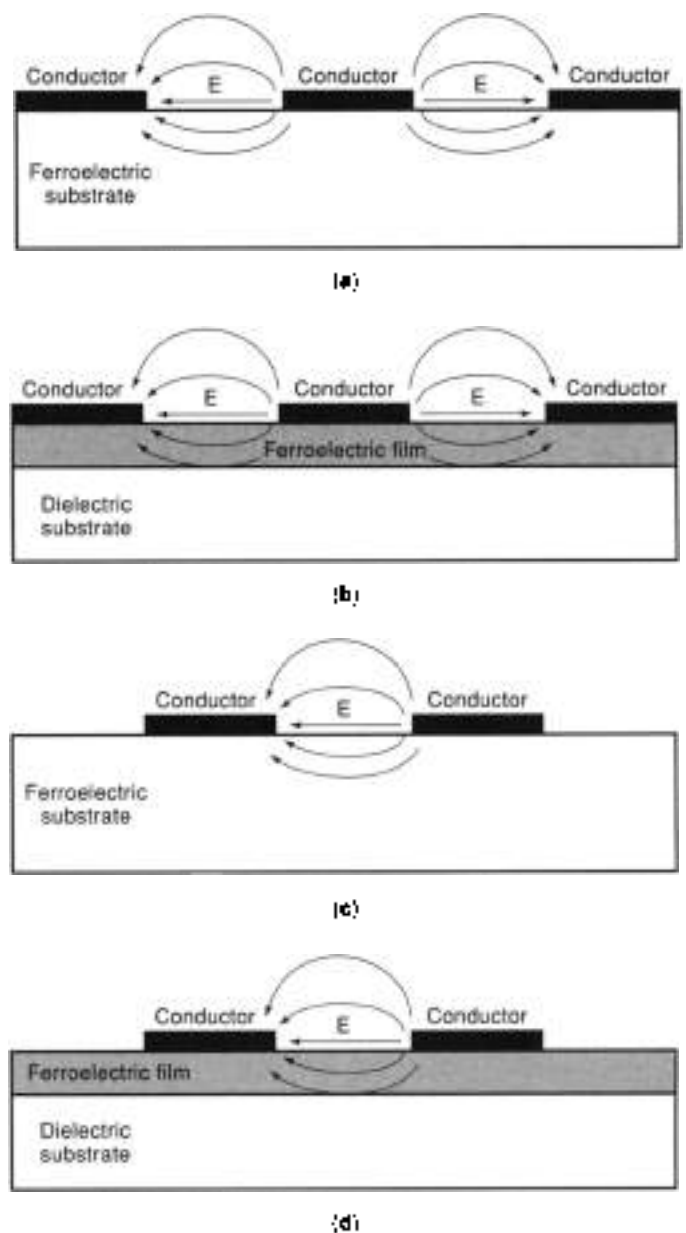
**Coplanar Waveguide, Coplanar Strip, and Slotline**

Other planar transmission lines such as coplanar waveguide, coplanar strip, and slotline have two or more conductors on the patterned surface. Hence the electric fields of both the dominant microwave mode and the bias are tangential to the substrate surface as shown in Fig. 2. The dominant mode of these planar transmission lines can be efficiently tuned with a bias field whether a bulk ferroelectric substrate or a thin-film ferroelectric on bulk dielectric substrate is employed. In the later case, good tunability is retained since the thin-film capacitance and substrate capacitance are in parallel. Design equations for a coplanar waveguide on multilayered dielectric substrates are available (7). Similar analyses using a partial-capacitance-conformal-mapping approach can be applied to other geometries to account for the ferroelectric thin film.

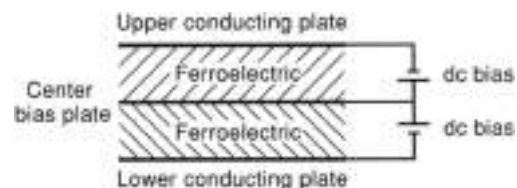
For the coplanar waveguide, the narrower the gaps between the center conductor and the ground planes, the higher the electric field intensity (and tunability) for a given bias voltage. Although coplanar waveguide is one of the simplest transmission lines, the microwave current density is sharply peaked at the edges of the strips causing large conductor losses. The problem is enhanced by the high dielectric constant and small thickness of the ferroelectric (8).

**Parallel Plate and Rectangular Waveguide**

A parallel-plate waveguide is a two-conductor guiding structure that supports transverse electromagnetic (TEM) waves. Thus, the electric and magnetic field are orthogonal to each other and to the direction of propagation. Figure 3 shows how this type of waveguide can be loaded with a ferroelectric medium to provide



**Figure 2.** Coplanar waveguide (CPW) transmission line cross section with (a) a homogeneous bulk substrate and (b) a thin film on a bulk substrate and coplanar strip (CPS) transmission line cross section with (c) a homogeneous bulk substrate and (d) a thin film on a bulk substrate. The electric field between the metal strips on the top surface is primarily parallel to the surface.



**Figure 3.** Parallel plate waveguide phase shifter (propagation into or out of paper).

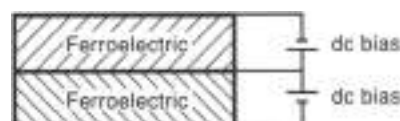


Figure 4. Rectangular waveguide phase shifter (propagation into or out of taper).

a variable phase velocity, which is given by

$$v_z = \frac{1}{\sqrt{\mu \epsilon}} \quad (4)$$

where  $\mu$  and  $\epsilon$  are the permeability and permittivity of the ferroelectric. Both the dc and the RF electric field are vertical. The ferroelectric is bifurcated with an electrode that is used to apply the dc bias with respect to the grounded waveguide walls.

Rectangular waveguides are popular in the microwave region. They are single-conductor guiding structures that confine the electromagnetic wave in the interior of the waveguide. Typically, the waveguide is operated in the dominant  $TE_{10}$  mode. Figure 4 shows how a ferroelectric can be used in this type of waveguide to provide a variable phase velocity, which is given by

$$v_p = \frac{1}{\sqrt{\mu \epsilon \left[ 1 - \left( \frac{\lambda_0}{2a} \right)^2 \right]}} \quad (5)$$

where  $\lambda_0$  is the cutoff wavelength. Again, both dc and RF electric field are vertical, and the ferroelectric is bifurcated by an electrode. Unlike the parallel plate waveguide (which has no sidewalls), rectangular waveguide has sidewalls. Therefore, a slot needs to be cut into a sidewall to connect the electrode to a dc power supply. The area of the slot opening must be small to prevent the microwave energy from leaking out of the slot.

## APPLICATIONS

In this section, we will describe several applications of ferroelectrics at microwave frequencies. These include varactors, voltage-controlled oscillators (VCOs), tunable filters, and phase shifting devices.

### Varactor

Varactors are variable-resistance circuit elements. They are used in switching or modulation of a microwave signal, for the generation of harmonics in an applied microwave signal, and in the mixing of two microwave signals of different frequency. As a discrete tunable capacitor, ferroelectric-based capacitors are applicable in a number of microwave circuits including VCOs, tunable filters, and oscillators. Parallel-plate configurations have not been successfully implemented (due to high required processing temperatures) in producing a high-quality ferroelectric thin film on a low-surface-resistance metal. Interdigitated capacitors (IDTs), where the metal electrodes are deposited on top of the ferroelectric (either thin-film or thick-film) or bulk substrate, have proven to be a more practical option. A typical interdigitated capacitor is shown in Fig. 5. Although strongly dependent on the ferroelectric material involved, Figs. 6 and 7 show the level

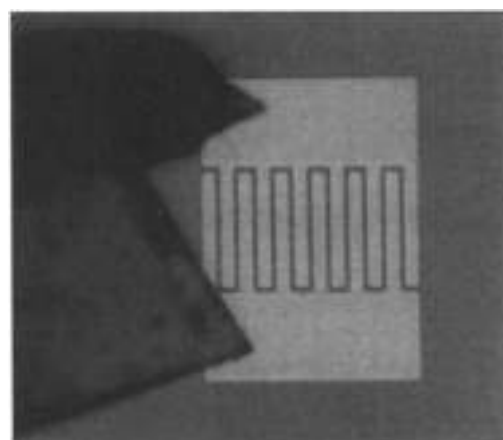


Figure 5. Photograph of a typical interdigitated capacitor on a thin-film ferroelectric covered substrate. The gap between fingers on the metal electrodes is 6  $\mu\text{m}$ . A microwave probe is shown contacting the device from the left.

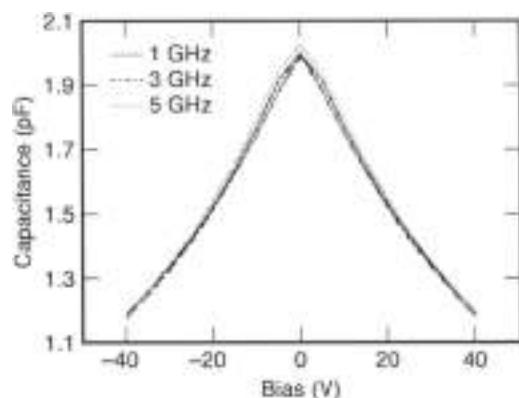


Figure 6. Capacitance versus bias voltage of a  $\text{Sr}_{0.5}\text{Ba}_{0.5}\text{TiO}_3$  thin-film interdigitated capacitor on an  $\text{MgO}$  substrate for frequency values of 1, 3, and 5 GHz. The data represents bias swept from  $-40$  V to  $40$  V and back to  $-40$  V.

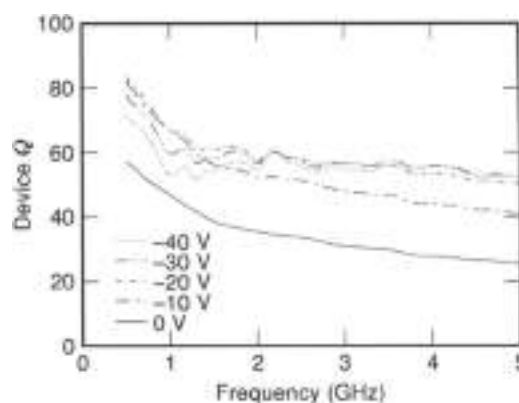


Figure 7. Interdigitated capacitor quality factor as a function of frequency for a  $\text{Sr}_{0.5}\text{Ba}_{0.5}\text{TiO}_3$  thin-film on an  $\text{MgO}$  substrate with bias voltages of 0 V,  $-10$  V,  $-20$  V,  $-30$  V, and  $-40$  V.

of performance available from thin-film ferroelectric interdigitated capacitor technology. Losses in ferroelectric interdigitated capacitors arise from the losses in the ferroelectric material as denoted by the dielectric loss tangent and from resistive losses in the metal electrodes. Using a low-surface-resistance metal such as silver or tin (some cases) superconductors minimize the electrode loss component, and in most cases the unloaded quality factor of the ferroelectric varactor is given by

$$Q_u = \frac{1}{\tan \delta} \quad (6)$$

Design and modeling of layered interdigitated lumped element capacitors is based on a conformal mapping approach (10).

### Voltage-Controlled Oscillator

Resonant circuits are used in oscillators and tunable filters. An oscillator provides a sinusoidal signal, and it is used as a source of microwave energy. The oscillator output should be clean (noiseless) and stable (frequency and power should not change with time). A high- $Q$  resonator is used in an oscillator circuit to obtain good frequency stability and low noise. The tunable capacitance of a ferroelectric-based capacitor is particularly applicable to a class of devices called a voltage-controlled oscillator (VCO). The bias-controlled change in reactance varies the oscillation frequency of an active element such as a transistor. Although there are many different oscillator topologies the designer can choose from, to first order the oscillation frequency can be varied in proportion to the square root of the bias-dependent capacitance. In principle, there is no difference in the design of a VCO using a ferroelectric capacitor and a semiconductor varactor diode (11). A VCO employing a ferroelectric tunable ring resonator has demonstrated 3% frequency tunability at 17 GHz (12). Phase noise is one of the primary limitations of any VCO application. Although several mechanisms contribute to phase noise, in many cases the  $Q$  factor of the tunable element is the limiting factor. As can be seen from Leeson's formula, we have (11)

$$A(f_n) = \frac{1}{2} \left[ 1 + \frac{1}{f_n^2} \left( \frac{f}{2Q_L} \right)^2 \right] \frac{FkT}{P_{av}} \left( 1 - \frac{f}{f_n} \right) \quad (\text{dBc/Hz}) \quad (7)$$

where  $f_n$  is the offset frequency,  $f$  is the oscillation frequency,  $P_{av}$  is the power level,  $F$  is the noise figure,  $f_c$  is the 1- $f$  noise corner frequency, and  $Q_L$  is the loaded quality factor which is related to the unloaded device quality factor  $Q_u$  by

$$Q_L = \frac{1}{\frac{1}{Q_u} + \frac{1}{Q_{EXT}}} \quad (8)$$

where  $Q_{EXT}$  is the external quality factor. As can be seen in Leeson's formula, the quality factor of the variable capacitor, which is, to first order, the reciprocal of the dielectric loss tangent, has a major impact on the phase noise of the VCO.

### Tunable Filter

A filter is any device or circuit which exhibits frequency selectivity; that is, the amplitude and phase of the output signal

are functions of frequency. A simple example is a bandpass filter where, ideally, all frequencies in a certain range are passed without change to the signal, whereas at any other frequency no signal appears at the output. There are many different ways to realize a microwave filter (13). Many of these rely on resonators which are coupled together in a carefully controlled fashion to realize the desired filter transfer function. Tunability of the filter transfer function can be achieved with ferroelectrics (14,15). It has been demonstrated that the center frequency of a microwave filter can be tuned by approximately 10% using ferroelectrics. In practice, there are many filter topologies that lend themselves to tuning with ferroelectrics. Conceptually the simplest to envision is tuning the center frequency of a bandpass filter composed of coupled half-wavelength resonators by varying the phase velocity of the resonant elements and hence their resonant frequency. From Eq. (2) it can be seen that the phase velocity is inversely proportional to the square root of the capacitance. Since many filter topologies rely on capacitive coupling of resonators, utilizing tunable coupling between resonators allows the design of tunable bandwidth filters.

### Phased Array Antenna

Phased array antennas can steer transmitted and received signals without mechanically rotating the antenna. Each radiating element of a phased array is normally connected to a phase shifter and a driver, which determines the phase of the signal at each element to form a beam at the desired angle. The most commonly used phase shifters are ferrite and diode phase shifters. Ferrite phase shifters are preferred at microwave frequencies, but they are expensive. The cost of a phased array mainly depends on the cost of phase shifters and drivers, and thus lower-cost phase shifting devices need to be developed to make the phased array antenna affordable for more applications. In this section, three different applications of ferroelectrics to phased array antennas will be described.

**Ferroelectric Lens Antenna.** The cost of a phased array depends mainly on the cost of phase shifters and drivers. A typical array may have several thousand elements as well as several thousand phase shifters and drivers; hence, it is very expensive. Therefore, reducing the cost and complexity of the phase shifters, drivers, and controls is an important consideration in the design of phased arrays. The ferroelectric lens phased array uniquely incorporates bulk phase shifting (2,3,16,17); the array does not contain individual phase shifters but rather uses ferroelectric material. This will reduce the number of phase shifters from  $nm$  to  $m$  to  $n$ , where  $n$  is the number of columns and  $m$  is the number of rows in a phased array. The number of phase shifter drivers and phase shifter controls is also significantly reduced by using row-column beam steering. The ferroelectric lens has the advantages of small lens thickness, high power-handling capability, and simple beam-steering controls, and it uses very low power to control the phase shift. Thus, it leads to low-cost phased arrays. However, it should be noted that the use of row-column steering may limit the level of side lobes that can be achieved.

**Description of Ferroelectric Lens and Its Operation.** The ferroelectric lens is shown in Fig. 5; each column of the lens is a set of conducting parallel plates that are loaded with bulk ferroelectric material. The material is bifurcated by a center conducting plate that is used to apply the dc bias voltage to

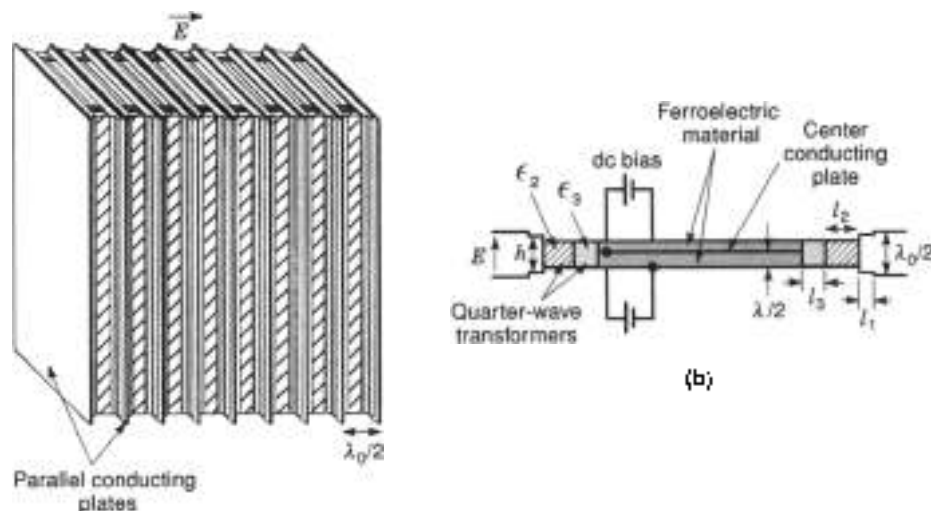


Figure 8. Ferroelectric lens.

(a)

the ferroelectric. The separation between the parallel plates at the input and output end is  $\lambda_0/2$ , where  $\lambda_0$  is the free space wavelength. Since only the TEM mode is desired, the separation between the parallel conducting plates is reduced to avoid higher-order mode propagation in the dielectric loaded section of the waveguide. Specifically, the separation between the center bias plate and either conducting plate is less than  $\lambda/2$ , where  $\lambda$  is the wavelength in the ferroelectric. Quarter-wave dielectric impedance transformers are used to match the empty waveguide to the ferroelectric loaded waveguide.

For scanning applications, a phase-shifting device must provide 360° differential phase shift. The amount of the ferroelectric material needed (in the direction of propagation) to obtain 360° differential phase shift is (16)

$$t = \frac{\lambda_0}{\sqrt{\epsilon_{r,max}} + \sqrt{\epsilon_{r,min}}} - \frac{\lambda_0}{\sqrt{\epsilon_{r,max}} |1 - \sqrt{\epsilon_{r,max}}|} \text{ tunability} \quad (10)$$

where  $\epsilon_{r,max}$  is the dielectric constant when no bias voltage is applied, and  $\epsilon_{r,min}$  is the dielectric constant when maximum dc bias is applied. Tunability is the fractional change in the dielectric constant as defined earlier. Thus, the thickness of the ferroelectric material needed is a function of the dielectric constant and the tunability of the ferroelectric, and the wavelength. Also, it can be shown that in order to obtain 360° phase shift, the dielectric loss through the ferroelectric is (16)

$$|dB| = \frac{27.3 \tan \delta}{1 - \sqrt{\epsilon_{r,max}}} \text{ tunability} \quad (10)$$

It may be noted that the lens loss is independent of the ferroelectric permittivity and depends only on its loss tangent and tunability.

In general, the ferroelectrics with higher dielectric constant offer higher tunability, which is desired to reduce the lens thickness. However, matching the lens to free space is easier for smaller  $\epsilon_r$ . Therefore, a compromise is needed between reducing the lens thickness (to reduce overall lens size) and achieving reasonable impedance match to reduce reflections from the lens

surface. For a typical value of  $\epsilon_r \sim 100$ , it is possible to obtain a tunability of 20%, which results in a reasonable lens thickness of  $\sim \lambda_0$  (e.g., 3 cm at 10 GHz). From Eq. (10), it can be seen that the  $\tan \delta$  must be less than 0.005 to limit the lens loss to less than 1 dB. The existing ferroelectric materials are a bit more lossy ( $\tan \delta = 0.008$  at 10 GHz).

**Phased Array Configurations Using Ferroelectric Lens for Two-Dimensional Scanning.** The ferroelectric lens offers electronic scanning in one plane. The lens proposed here can be fed by a non-scanning planar array, like a slotted waveguide array. A combination of slotted waveguide array with phase shifters and the lens proposed here can be used as a phased array that can scan in two planes. A space feed can be used with the combination of two lenses proposed here (with a polarization rotator in between) to scan the beam in two planes. The details of these alternatives are discussed elsewhere (16,17).

**Theoretical and Experimental Results.** For the theoretical analysis of the ferroelectric lens, an individual section between two conducting parallel plates of the lens can be considered as one column of a phased array. The column can be analyzed as a two-dimensional (2-D) parallel-plate waveguide with electric field of the TEM mode normal to the plates as shown in Fig. 6. A matching network was designed using mode matching technique assuming that the dielectric constant of the ferroelectric varies from 120 to 80 (33% tunability) over a frequency range of 8 GHz to 12 GHz (40% bandwidth) and that  $\epsilon_{r1} = 0.8$  in. and  $\lambda = 0.1$  in. The computed results are shown in Fig. 9. The matching network parameters are  $t_1 = 0.2956$  in.,  $t_2 = 0.1860$  in.,  $t_3 = 0.0505$  in.,  $h = 0.2345$  in.,  $\epsilon_2 = 2.54$ , and  $\epsilon_3 = 35$  (see Fig. 8).

Experiments were performed with the ferroelectric composition  $\text{Ba}_{0.7}\text{Sr}_{0.3}\text{TiO}_3$  with 60% oxide. This material offered a good compromise among  $\epsilon_r$ ,  $\tan \delta$ , and tunability. At 10 GHz, for this composition,  $\epsilon_r = 100$  and  $\tan \delta = 0.0079$ . The ferroelectrics were 1 in. long (in the direction of propagation), 0.05 in. high and 5 in. ( $\sim 4\lambda$  at 10 GHz) wide. Figure 10 shows the measured transmission and reflection coefficients at zero bias. The reflection coefficient is sufficiently small over a wide frequency band as the theory had predicted in Fig. 9. Figure 10 also shows that the loss increases with frequency; this is due to two reasons. First,  $\tan \delta$  increases with frequency, which is

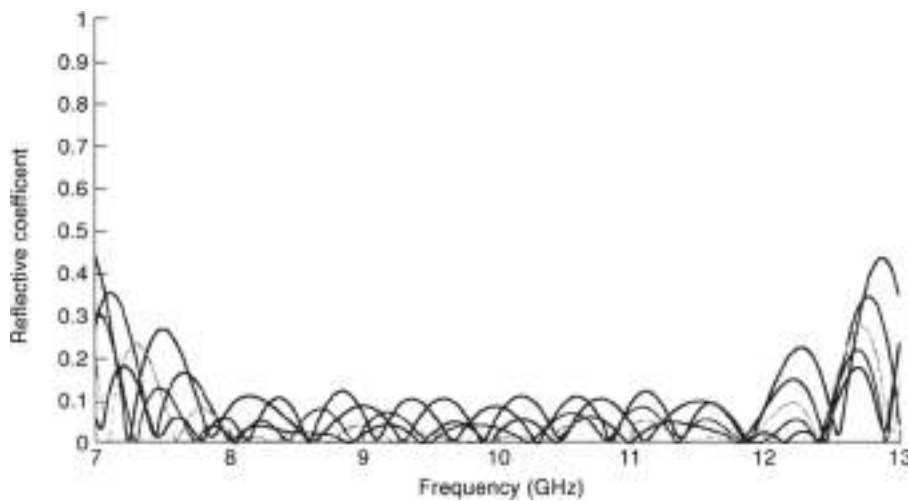


Figure 9. Theoretical reflection coefficient for  $\epsilon_r = 80$  to 120

expected for ceramics, second, the electrical length (in terms of wavelengths) of the ferroelectric in the direction of propagation increases with frequency since the physical length is kept constant (1 in.).

Figure 11 shows the measured phase shift as a function of the bias voltage for various frequencies. As expected, the phase shift increases linearly with frequency because the electrical length of the material increases with frequency. Since ferroelectrics are good insulators, the dc current requirements are very low. For example, at 10 kV bias voltage, the dc current drawn was 0.05 mA, and thus the dc power dissipated is only 0.5 W. The bias voltage can be reduced by further bifurcating the ferroelectrics using interdigital electrodes).

Figure 12 shows the reflection coefficients as a function of frequency for various bias voltages. The standing wave ratio

(SWR) is less than 2 for frequency range of 8 GHz to 12 GHz as theoretically predicted earlier (see Fig. 9).

**Traveling Wave Antenna.** Another type of phased array antenna that also uses bulk phase shifting is a traveling wave antenna, as shown in Fig. 13. The antenna is a slab of ferroelectric material with conducting strips on the top side of the slab and a ground plane on the bottom. This type of antenna is well-suited for millimeter-wave applications when a low-loss dielectric (not a ferroelectric) is used as a substrate, and frequency variation is used to scan the antenna beam electronically. Instead of changing the frequency, the dielectric constant of the ferroelectric substrate can be changed to electronically scan the antenna beam in the *E* plane (17–19). It can be shown (18) that the radiation angle of the antenna beam is given by

$$\theta = \sin^{-1} \lambda_c \left( \frac{1}{\lambda_0} - \frac{1}{d} \right) \quad (11)$$

where  $\lambda_0$  and  $\lambda_c$  are the free space and guide wavelength, respectively, and  $d$  is the spacing between the conducting strips. The guide wavelength and thus the scan angle changes as the dielectric constant of the substrate changes. The main advantage of this type of antenna is that only a single dc power supply

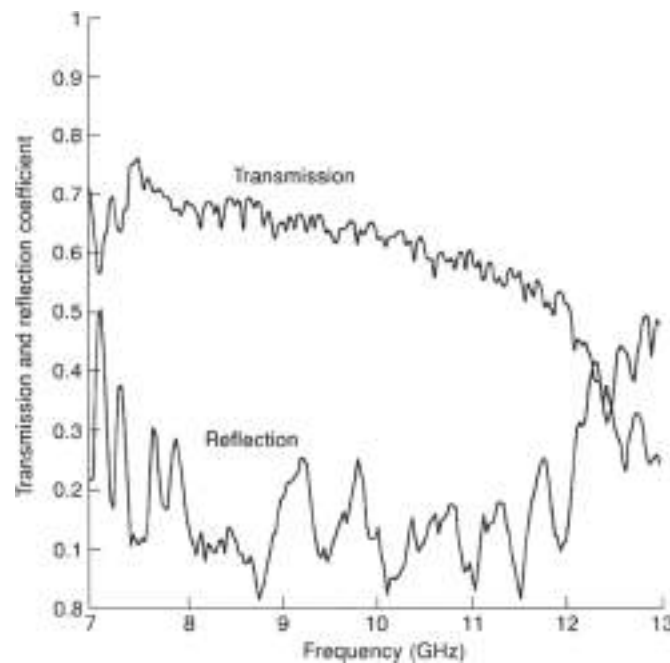


Figure 10. Measured reflection and transmission coefficient at zero bias voltage.

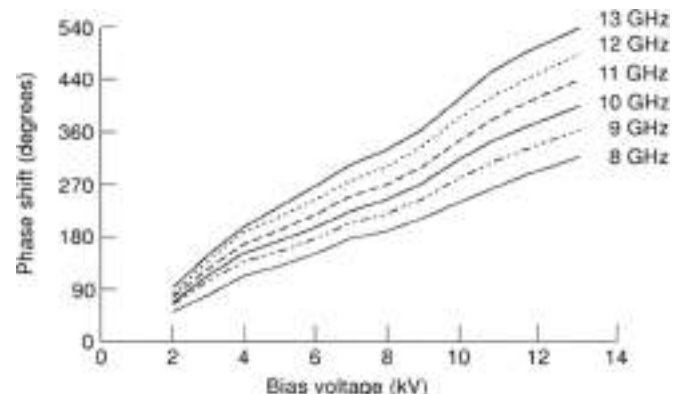


Figure 11. Measured differential phase shift.

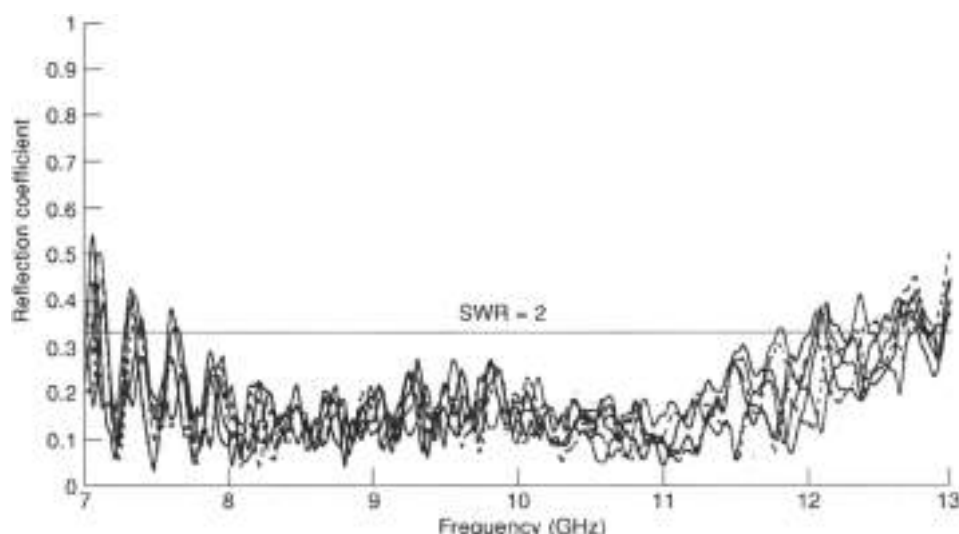


Figure 12. Measured reflection coefficient for various bias voltages (0 to 15.5 kV).

is needed to scan the beam, and it is a very simple structure. However, this antenna has a major disadvantage that makes it quite impractical. Since the physical size of most microwave antennas is at least a few wavelengths (if not a few tens of wavelengths), the loss that the electromagnetic wave would suffer as it travels down the antenna is enormous. Also, the instantaneous bandwidth of this antenna is very small because it is a frequency scan antenna. That is, the beam pointing direction changes as the frequency changes. Like the ferroelectric lens, the traveling wave antenna offers electronic beam scanning in one plane. Electronic scanning in the other plane (azimuth plane in Fig. 13) can be achieved with phase shifters in the linear array fed for this antenna.

**Discrete Phase Shifter.** Ferroelectrics have also been proposed for discrete phase shifter applications at microwave frequencies. There are several advantages of using ferroelectrics over ferrites in phase shifters. First, since ferroelectrics are voltage-driven devices, the dc control power requirements are small. However, unlike latching ferrite phase shifters that only require current pulses, the bias voltage needs to be applied to

the ferroelectrics during the entire transmit and/or receive cycle. Second, ferroelectrics provide reciprocal phase shift. Third, the high dielectric constant of the ferroelectric has the effect of decreasing the overall size of the phase shifter. At the present time, however, the dielectric loss in ferroelectrics is higher than that in ferrites at microwave frequencies.

The basic design equations for a discrete phase shifter are the same as those for the ferroelectric lens. For the same electric field applied in the lens, the discrete phase shifter should provide similar phase shift using the same ferroelectric. Phase shifters have been designed using ferroelectric-loaded rectangular waveguides as well as planar transmission lines, like microstrip and coplanar waveguide, on a ferroelectric substrate (5,20,21).

For the rectangular waveguide, impedance matching techniques similar to the ones used in the ferroelectric lens can be applied. For the microstrip line, several impedance matching techniques have been tried (5) including quarterwave transformers, open circuit stubs and radial stubs. For the coplanar waveguide, usually the lines are tapered to provide a 50  $\Omega$  impedance (21).

### High Temperature Superconductors and Ferroelectrics

The discovery of high temperature superconductors (HTS) has generated many tunable device designs (22) using both thin films and bulk ferroelectrics. One of the main incentives is the promise of the low conductor loss associated with HTS. In addition, both BTO and BSTO are closely lattice matched with the HTS yttrium barium cuprate,  $YBa_2Cu_3O_{7-x}$  (YBCO), meaning that the ferroelectric and YBCO can be epitaxially grown on top of each other to form multilayer thin film structures. This has been done with STO and YBCO. To take advantage of the low microwave losses, operation must be below the critical temperature for YBCO. Most of the research has been done at 77 K, the liquid nitrogen boiling temperature. In BSTO, the barium to strontium ratio can be adjusted so that BSTO can be operated in the paraelectric phase at 77 K, and STO remains paraelectric down to the lowest temperatures.

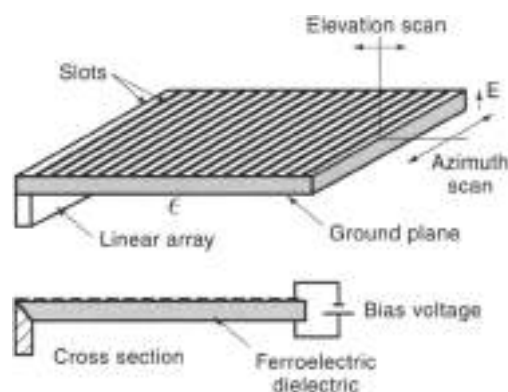


Figure 13. Ferroelectric traveling-wave antenna.



**BIBLIOGRAPHY**

1. P. K. Larsen et al., Nanosecond switching of thin ferroelectric films, *Appl. Phys. Lett.*, **59**(5): 811-813, 1991.
2. J. B. L. Rao, D. P. Patel, and L. C. Sengupta, Phased array antennas based on bulk phase shifting with ferroelectrics, *Integr. Ferroelectr.*, **22**: 307-316, 1996.
3. J. B. L. Rao et al., Ferroelectric materials for phased array applications, *Dig. IEEE Antennas Propag. Soc. Int. Symp.*, **4**: 2264-2267, 1987.
4. J. S. Horvath et al., Structure-property relationships in ferroelectric thin films for frequency agile microwave electronics, *Integr. Ferroelectr.*, **22**: 279-289, 1999.
5. R. W. Balbutt, E. E. Kuscia, and W. C. Dreck, Planar microwave electro optic phase shifters, *Microwave J.*, **35**(6): 63-79, 1992.
6. K. C. Gupta, R. Garg, and J. Bahl, *Microstrip Lines and Slot Lines*, 2nd ed., Boston, MA: Artech House, 1996.
7. J. Svecina, A simple quasi-static determination of basic parameters of multilayer microstrip and coplanar waveguide, *IEEE Microw. Guided Wave Lett.*, **MGWL-2**: 395-397, 1992.
8. S. S. Gevorgian et al., HTS-ferroelectric devices for microwave applications, *IEEE Trans. Appl. Supercond.*, **ASC-7**(2): 2453-2461, 1997.
9. J. M. Pond et al., Microwave properties of ferroelectric thin films, *Integr. Ferroelectr.*, **22**: 317-326, 1998.
10. S. S. Gevorgian et al., CAD models for multilayered substrate interdigital capacitors, *IEEE Trans. Microwave Theory Tech.*, **MTT-44**(6): 896-894, 1996.
11. G. D. Vendelin, A. M. Pavon, and U. L. Rohde, *Microwave Circuit Design*, New York: Wiley, 1990.
12. R. R. Romanofsky, P. W. Van Kester, and P. A. Miranda, A cryogenic GaAs-PHEMT/ferroelectric K-band tunable oscillator, *3rd European Workshop on Low Temperature Electronics*, San Marino, Italy, June 24-26, 1998.
13. G. Matthaei, L. Young, and E. M. T. Jones, *Microwave Filters, Impedance Matching Networks, and Coupling Structures*, Norwood, MA: Artech House, 1980.
14. S. S. Gevorgian et al., Tunable superconducting band-stop filters, *IEEE MTT'S 1998 Int. Microw. Symp. Dig.*, **2**: 1027-1030, 1998.
15. G. Satchanarayanan, P. Van Kester, and P. A. Miranda, A novel K-band tunable microstrip bandpass filter using a thin film HTS-ferroelectric/dielectric multilayer configuration, *IEEE MTT'S 1998 Int. Microw. Symp. Dig.*, **2**: 1011-1014, 1998.
16. J. B. L. Rao, D. P. Patel, and V. Kriehaveky, Voltage controlled ferroelectric lens phased arrays, accepted for publication in *IEEE Trans. Antennas Propag.*
17. J. B. L. Rao, G. V. Trank, and D. P. Patel, Two low-cost phased arrays, *Proc. 1996 IEEE Int. Symp. on Phased Array Systems and Technology*, 119-124, 1997.
18. V. K. Varadan et al., Electronically steerable leaky wave antenna using a tunable ferroelectric material, *Smart Mater. Struct.*, **3**: 470-475, 1994.
19. T. W. Bradley et al., Development of a voltage variable dielectric (VVD) electronic scan antenna, *Proc. Radar 97, IEE Proc.*, **449**: 395-396, 1997.
20. V. K. Varadan et al., Ceramic phase shifters for electronically steerable antenna systems, *Microw. J.*, **35**(1): 116-127, 1992.
21. C. M. Jackson, New phase shifters for smart systems, in V. K. Varadan (ed.), *Smart Structures and Materials 1996, Smart Electronics*, Proc. SPIE **2446**, 219-225, 1995.

22. O. Vendik, I. Miruzenko, and L. Ter-Murtusyan, Superconductors Spur Applications of Ferroelectric Films, *Microwaves and RF*, **33**(7): 67-70, 1994.

D. P. PATEL  
 J. M. POND  
 J. B. L. RAO  
 Naval Research Laboratory

**PELTIER EFFECT**

It has been known for more than 150 years that the transport of electric charge in conductors is accompanied by flow of heat/energy. In conducting materials there are three types of reversible effects that arise when both electric currents and temperature differences are present: the Seebeck effect, the Peltier effect, and the Thomson effect. These phenomena are known as *thermoelectric effects*, and they have been proven to be very fundamental for the advance of our understanding of the properties of metals and semiconductors. Studies of thermoelectric effects provide information about the electronic structure and the interactions between electrons and both phonons and impurities. In addition, they have significant applications to industrial technology such as energy conversion, power production, and refrigeration. Today a great variety of thermoelectric generators and thermoelectric infrared detector coolers exist.

The history of the Peltier effect started in 1834, when a French watchmaker and amateur scientist named Jean Charles Athanase Peltier discovered that when an electric current  $I$  is forced through a junction between two different materials, which are initially at uniform temperature, heat flows from one material to another. The amount of heat that is liberated (junction heats up) or absorbed (junction cools down) per unit of time ( $Q$ ) is proportional to the current. The coefficient of proportionality,  $\Pi$ , is called the Peltier coefficient,  $Q = \Pi I$ , where the sign of  $Q$  (plus for heating and minus for cooling) depends on the direction of the current. The Peltier heating/cooling is a reversible phenomenon in the sense that it depends linearly on the current, and it should be distinguished from the irreversible Joule heating (quadratic dependence) that takes place in any single conductor. Namely, if heat is released when current flows from one conductor to the other conductor, then upon current reversion heat will be absorbed. The simple physical principle that describes all the thermoelectric effects is as follows: As electrons travel through the junction between the two different conductors, on average, they will either lose or gain energy, since their electronic states in these conductors have different energies. For example, if a temperature difference is maintained across the junction, in the absence of external flow of current, electrons will diffuse from the hot conductor to the cold one where they can find states of lower energy. Eventually, more electrons will be accumulated on the cold side of the junction, and this results in the appearance of a voltage difference (Seebeck effect).

Although the Peltier effect is usually the effect that takes place when two different materials are in contact, the Peltier coefficient  $\Pi$  can be defined for each individual conductor and

it is an intrinsic property of the conductors. The same is true for all other thermoelectric coefficients, and generally they are strongly temperature-dependent. However, it is difficult to measure  $\Pi$  for individual conductor experimentally, and usually the Seebeck coefficient can be measured rather readily and accurately. In order to calculate  $\Pi$ , the Kelvin relations can be employed. In 1857 Lord Kelvin, using equilibrium thermodynamics, derived two very simple equations that relate the three thermoelectric coefficients (Kelvin relations):

$$\Pi = TS \quad (1)$$

$$\alpha_1 = T \frac{dS}{dT} \quad (2)$$

where  $S$  is the thermopower or Seebeck coefficient and  $\alpha_1$  is the Thomson coefficient. More rigorously, these relations follow from the theory of irreversible thermodynamics, as it was shown by Onsager almost a century later.

The temperature dependence of  $\Pi$  can be easily estimated for metals, using the free-electron gas model. The Thomson coefficient has the meaning of specific transport heat capacity (1 per unit of electric charge). In metals, only a small fraction of conduction electrons contribute to transport, approximately  $T/T_F$ , where  $T_F$  is the Fermi temperature of the electrons, typically of the order of  $10^4$  K, and we have  $\mu_T \approx k_B T$  ( $T \ll T_F$ ). Kelvin's relations then immediately give us

$$\Pi \approx \frac{k_B T^2}{e T_F} \quad (3)$$

It is worth mentioning the value of the ratio of the Boltzmann constant to the electric charge of the electron,  $k_B/e = 85.4 \mu\text{V}/\text{K}$ , since it helps to understand the order of magnitude of the thermoelectric effects. The above estimate of the Peltier coefficient holds practically in most cases of pure nonmagnetic metals and can be derived using the kinetic theory of electrons in metals. Both the Peltier coefficient and the Seebeck coefficient in superconducting materials are equal to zero for temperatures below their critical temperature.

## PELTIER EFFECT IN METALS AND SEMICONDUCTORS

The thermoelectric coefficients are defined in the context of linear response theory: The temperature gradient or the electric field induce small perturbations from the equilibrium state of the conductor so that the electric charge current and the heat current are linear functions of these gradients. The heat current density  $J_h$  is analogous to the electrical current density  $J_e$ , but instead of measuring electric charge, it measures thermal energy being carried. The formal definitions are:

$$J_e = \int v(\mathbf{k}) f(\mathbf{k}) d\mathbf{k} \quad (4)$$

$$J_h = \int v(\mathbf{k}) \cdot \mu(\mathbf{k}) f(\mathbf{k}) d\mathbf{k} \quad (5)$$

where  $v(\mathbf{k})$  is the velocity of the electrons,  $\mu$  is the temperature-dependent chemical potential, and  $f(\mathbf{k})$  is the distribution function of the electrons. For metals that are in equilibrium,  $f(\mathbf{k})$  is given by the Fermi-Dirac distribution function  $f(\mathbf{k}) = [1 + \exp(\epsilon(\mathbf{k}) - \mu)/k_B T]^{-1}$ . In addition, in the equilibrium state (i.e.,

absence of "external forces" such as electromagnetic fields or temperature differences) both current densities are zero. The dynamics of the electron transport is hidden in the distribution function since it carries information about the scattering of electrons by impurities or lattice vibrations. Usually it is determined by the solution of Boltzmann transport equation with appropriate boundary conditions. Upon linearization with respect to the electric field  $\mathbf{E}$  and temperature gradient  $\nabla T$ , Eqs. (4) and (5) provide the mathematical connection between all the thermoelectric coefficients:

$$J_e = \sigma \mathbf{E} - S \nabla T \quad (6)$$

$$J_h = \alpha \Pi \mathbf{E} - \kappa \nabla T \quad (7)$$

where  $\sigma$  and  $\kappa$  are the electrical conductivity and the thermal conductivity, respectively.

For a degenerate metallic conductor, the Peltier coefficient is given by the celebrated Mott formula:

$$\Pi = \frac{\pi^2 (k_B T)^2}{3e} \left( \frac{d \ln \sigma(E)}{d \ln E} \right)_{E=\mu_T} \quad (8)$$

where  $\sigma(E)$  is the conductivity that would be found in a metal for electrons of average energy  $E$ ,  $\sigma(E) = e^2 \nu(E) n(E) \tau(E)$ , and  $\epsilon_F$  is the Fermi energy. When the scattering of electrons is mainly due to impurities and is isotropic—as in dilute alloys—the mean free path  $l$  is constant. Consequently, the relaxation time is  $\tau(E) = l/v(E) \propto E^{-1/2}$ , since the average electron velocity  $v(E)$  is proportional to  $\sqrt{E}$ . The electron density of states for a free electron gas is  $n(E) \propto \sqrt{E}$ . As a result, Eq. (8), reads

$$\Pi = \frac{\pi^2 (k_B T)^2}{3e} \frac{1}{\epsilon_F} \quad (9)$$

In pure metals and for temperatures above a few degrees kelvin, the relaxation time is determined primarily by electron-phonon scattering. Approximately,  $\tau(E) \propto E^{1/2}$  and

$$\Pi = \frac{\pi^2 (k_B T)^2}{e} \frac{1}{\epsilon_F} \quad (10)$$

The Peltier coefficient for metals is of the order of  $50 \mu\text{V}$  at room temperatures, and for a wide range of temperatures it is quadratic in temperature. Moreover, because of the smallness of the degeneracy factor  $k_B T/\epsilon_F$ , the thermoelectric effects in metals are very weak. This is not true in semiconductors, where this factor is absent.

For a semiconductor with relatively few conduction electrons that are characterized by the Boltzmann distribution, the Peltier coefficient has the following form (11):

$$\Pi = \frac{k_B T}{e} \left( \frac{5}{2} - \frac{2 \ln \tau(E)}{3 \ln E} - \frac{\mu(T)}{k_B T} \right) \quad (11)$$

where

$$\mu = k_B T \ln \left( \frac{1}{2} N_c \left( \frac{2\pi\hbar^2}{m^* k_B T} \right)^{3/2} \right) \quad (12)$$

$m$  is the effective mass of the carriers in the semiconductor and  $N_c$  is the carrier concentration—the number of charge carriers per unit volume—which has a strong temperature dependence. Generally, the relaxation time can be written as

$$\tau(E) = g/T^s \quad (13)$$

where  $g/T^s$  is a function of temperature only and  $s$  is a constant that depends on the type of scattering that is dominant. Usually  $s$  is in the range from  $-2$  to  $2(2)$ . Consequently, the second term in Eq. (11) is equal to  $s$ , whereas the last term is at least of the order of unity. From the above analysis, it is clear that the order of magnitude of the Peltier coefficient is of the order of  $\mu\text{V}$ , which is much higher than that of metals, and varies linearly with the temperature. The simplicity of the above expression for  $\Pi$  is restricted only to extrinsic semiconductors. When the carriers are holes (i.e., in  $p$ -type semiconductors),  $\Pi$  has opposite sign. For intrinsic semiconductors, where both type of carriers are present, the Peltier coefficient takes a more complicated form and its magnitude, generally, is smaller.

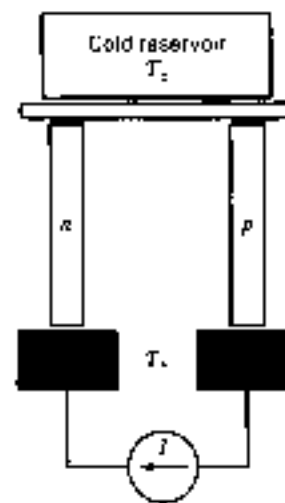
**PELTIER REFRIGERATION**

The most important applications of the Peltier effect lie in the possibility to create efficient cooling solid-state devices that can operate at room temperatures. The basic essentials of the operation of the Peltier cooler (or heater) are straightforward: Two different materials, usually a  $p$ -type and an  $n$ -type semiconductor, are brought in contact. Sometimes  $p$  and  $n$  refer to the positive and negative thermoelements, respectively. Let  $T_c$  be the temperature of the junction which is in thermal equilibrium with a reservoir that we want to cool down. The other two ends remain at a temperature  $T_h > T_c$  and are connected to a battery that produces an electric current  $I$  that passes through the junction. A diagram of this type of solid-state thermoelectric device is presented in Fig. 1. There are three types of processes (one reversible and two irreversible) that transfer thermal energy from the cold to the hot reservoir:

- The rate of Peltier heat absorption  $Q_1 = \Pi_{pn}I$ , which is determined by the Peltier coefficients of the two semiconductors  $\Pi_{pn} = \Pi_n - \Pi_p$ , and the current  $I$ .
- The rate of heat generation  $Q_2 = K\Delta T$ , due to existence of a temperature difference between the two reservoirs  $\Delta T = T_h - T_c$ . It is characterized by the sum of the thermal conductances of the two semiconductors  $K = K_n + K_p$ .
- The rate of Joule heat production, which is delivered to each reservoir—that is,  $Q_3 = I^2/2R$ , where  $R$  is the total resistance of the two semiconductors. The junction resistance is assumed to be negligible compared to the bulk resistances of the two thermoelements.

In order to evaluate the efficiency of the cooling device, one has to calculate the coefficient of performance, which is defined as the ratio of the rate of heat removed from the cold reservoir to the total electrical power supplied by the battery. Upon maximization with respect to the electric current (2), the coefficient of performance for the Peltier cooling unit is

$$\phi = \frac{T_c}{T_c} \frac{\Omega}{\Omega + 1} \quad (14)$$



**Figure 1.** Schematic of the solid state refrigerator. Due to the Peltier effect the current drives heat out of the cold reservoir, which is in thermal equilibrium with the  $n$ - $p$  junction, towards the dark region. The maximum temperature difference that can be achieved is given by Eq. (17). Upon current reversion, the inverse process takes place.

where  $\Omega = \sqrt{1 + z^2 T_h}$  and  $T_h = (T_c + T_h)/2$  is the average temperature of the two reservoirs. The figure of merit  $z$  of the coupled system is a function of the thermopowers, electrical conductivities ( $\sigma$ ), and thermal conductivities ( $\kappa$ ) of the two materials:

$$z = \left( \frac{S_{pn}}{\sqrt{\kappa_n/\sigma_n} + \sqrt{\kappa_p/\sigma_p}} \right)^2 \quad (15)$$

The optimal value of the current is

$$I_{opt} = \frac{S_{pn} \Delta T}{R(\Omega - 1)} \quad (16)$$

where  $S_{pn} = S_n - S_p$  is the difference between the thermopowers of the two components. The maximum value of the coefficient of performance is limited by the Carnot cycle coefficient of performance  $\phi_c = T_c/\Delta T$ . The second factor in Eq. (14) is between zero and unity and describes the reduction of the efficiency due to irreversible processes of thermal conduction and Joule heat generation that occur in the device. The maximum temperature difference that the Peltier refrigerator can achieve is obtained by setting the optimal coefficient of performance,  $\phi$ , equal to zero—that is,  $\Omega = T_h/T_c$ —or

$$\Delta T_{max} = \frac{zT_c^2}{2} \quad (17)$$

To get a sense of the operation values of the parameters, say that the figure of merit is  $z = 2 \times 10^{-4} \text{ K}^{-1}$ , the hot reservoir is kept at approximately room temperature  $T_h = 300 \text{ K}$ , and the cold reservoir is kept at ice temperature, about  $T_c = 270 \text{ K}$ ; then  $\Omega = 1.258$ , the efficiency is  $\phi = 0.60$ , and  $\Delta T_{max} = 73 \text{ K}$ .

A unique property exists in the thermoelectric devices that is based on the reversible nature of the effects. For example, the solid-state refrigerator can act as a heating device (heat

pump by reversing the direction of the current. The coefficient of performance for the heat pump  $\phi_p$  is equal to  $\phi_p = \phi + 1$ , and it is greater than unity. The optimal values for the operation of the heating device are the same as for the cooling device. Also if the battery that provides the current is replaced by a resistance load  $R_l$ , and we exchange the temperatures of the two reservoirs in Fig. 1, then upon heat deposition in the hot junction the device can act as a power generator (current is generated in the load, due to the Seebeck effect). However, depending on the type of operation, the design of the thermoelectric device has to be adjusted so that it provides the optimal performance. For a thermoelectric power generator it can be shown (2) that under optimal conditions of operation, the efficiency, defined as the ratio of the output electric power on the resistance load to the input heat at the hot junction is

$$\phi_p = \frac{\Delta T}{T_c} \frac{\Omega - 1}{\Omega + T_c/T_h} \quad (18)$$

The optimal resistance load is  $R^{opt} = \Omega R$  and the output power is

$$P_{out} = \frac{\Omega}{R} \left( \frac{S_m \Delta T}{\Omega + 1} \right)^2 \quad (19)$$

The above equations indicate that the figure of merit is the most fundamental quantity for thermoelectric refrigeration and provides us with a quality criterion for the selection of the materials to be used for thermoelectric devices. As a matter of convenience, the quantity that is most studied in the literature is the dimensionless figure of merit  $ZT$ , which is defined for an individual conductor as

$$ZT = \frac{\pi^2}{T R K} = \frac{TS^2}{RK} \quad (20)$$

To a good approximation, the total figure of merit can be taken to be the average of the two individual figures of merit. It is evident from the above expressions that the efficiency of the Peltier cooler depends on the temperatures  $T_c$  and  $T_h$  (sometimes called the ambient temperature) and also on the transport properties of the materials through the figure of merit. It is also clear that the figure of merit should be as large as possible. This is actually the central problem in the thermoelectric refrigeration and, although some success has been attained in the past, is still an open problem. In order to obtain an effective Peltier cooler, the Seebeck coefficients of the two components should have opposite signs and should be large. Moreover, we need materials with high electrical conductivity and low thermal conductivity. Furthermore, the strong temperature dependence of the figure of merit, which usually peaks at some temperature and then decreases quickly, must be considered, since it is not easy to find good positive and negative thermoelements that have the same properties within the same temperature window.

A pair of thermoelectric materials usually performs well within some temperature range, as determined by Eq. (17), but might not fit our requirements. One way to resolve this problem is to design a multistage cooling unit or refrigeration cascade. At each stage the thermoelectric materials can achieve some

fraction of the total temperature difference that is of interest. Thus, in order to get a large temperature difference, a variety of materials have to be chosen so that they operate optimally within some temperature window. The cascades have the form of pyramids of pairs of thermoelements that are electrically in series, and the hot junction of the first stage acts as a cold junction for the second stage, and so on. For multilayer refrigerating devices the heat cooling capacity of each layer is required to be larger than that of the previous stage since, at each level of operation, not only the amount of heat absorbed from the previous cold junction, but also the electrical power put into the preceding stage, has to be removed. Therefore, for each extra stage the number of Peltier elements has to be larger than that of the previous one. In a cascade type of arrangement, not only the attainable temperature differences are larger but also the coefficient of efficiency increases significantly. Therefore, cascades are certainly superior to single-stage thermoelectric devices. However, there are some technical limitations related to thermal isolation of each unit, and only two to three stages are used in practice.

Using metals as thermoelectric components is not preferable, in spite of the favorable small values of the ratio  $\pi^2/\beta$ . From the Wiedemann-Franz law we have

$$\pi^2/\beta = LT = 2.45 \times 10^{-8} T^2 (\text{K}) \quad (21)$$

where the Lorenz number is

$$L = \frac{\pi^2}{3} \left( \frac{k_B}{e} \right)^2 \quad (22)$$

This is due to the fact that the Peltier effect in metals is very weak (as was shown in the previous section) at all temperatures, and, as a result, typical values of the figure of merit,  $ZT$ , for metals at room temperatures is approximately 0.05. The performance of a metal-based cooler is approximately 1% and has no practical value. Furthermore, the metals do not meet the requirements for efficient cooling at liquid nitrogen temperatures. For lower temperatures, the Peltier coefficient in metals can be increased substantially by adding certain magnetic impurities (e.g., iron) to pure metals. A rapid increase in the density of states at energies close to the Fermi level occurs, related to the Kondo effect. From the Mott formula, one then can see that this change in the density of states would increase the Peltier coefficient.

Semiconductors have more advantages than metals, since the Peltier coefficient is larger. Also, by varying the concentration of the impurities (doping) a further optimization of thermopower, electrical conductivity, and thermal conductivity is possible. This flexibility of carrier concentration adjustment made the semiconducting materials the most favorite among researchers for the past 50 years. Usually higher doped semiconductors are preferable, but the doping impurities that determine the carrier concentration must be so chosen, to maximize  $S^2/R$ . However, highly doped semiconductors have large phonon thermal conductivity in addition to the electronic one. At room temperatures, typical values of  $ZT$  for semiconductors are near unity. The semiconducting compound  $\text{Bi}_2\text{Te}_3$  and other solid solutions such as bismuth selenide or antimony telluride were extensively studied in the past (3) for thermoelectric refrigeration at room temperatures. They can provide

a maximum temperature difference of the order of 40 K, and the efficiency is of the order of 10%. At cryogenic temperatures, the carriers in the semiconducting materials are frozen, which limits their potential use for refrigeration in this temperature range.

The performance of a Peltier refrigerator is much smaller than the one of conventional vapor-compression-type systems that are used for domestic refrigerators. The attainable cooling power is approximately 10 W, and there is a need for substantial improvement in the figure of merit of the Peltier thermomaterials. The use of reduced dimensionality systems was recently suggested (4) in order to increase the figure of merit of conventional bulk semiconducting thermoelectric materials which for the past 20 years seemed to reach their maximum potential performance. Based on the latest advances in nano-lithography and fabrication technology, it is possible to confine electrons in a very narrow, one-dimensional region and thus create *quantum wires* or *one-dimensional conductors*. In such exotic systems, lattice thermal conductance appears to diminish due to increased phonon scattering from the surface. The results show a very drastic change in the figure of merit at room temperatures: for a bulk three-dimensional bismuth-telluride alloy  $ZT = 0.67$ , whereas for a wire with 5 Å diameter  $ZT = 14$  (4).

For Peltier refrigeration at and below liquid-nitrogen temperatures (77 K), Bi-Sb alloys are promising candidates for the *n*-type material, with individual figure of merit about  $7 \times 10^{-3} \text{ K}^{-1}$  at 80 K. On the other hand, traditional *p*-type thermoelectric materials, like bismuth-antimony-tellurides, have low individual figure of merit and would significantly degrade the potential performance of solid-state refrigerators. That is, the total figure of merit is less than the figure of merit of the negative thermoelement. If we use a superconducting material in which the Seebeck coefficient is zero and the ratio of the thermal to electrical conductivity is significantly smaller than that of the semiconductor, the efficiency will be restored. The use of high- $T_c$  superconductors like YBCO, with high critical current density, as a *p*-type material offers this alternative solution (5). A maximum temperature drop of 7 K with the hot junction at 78 K was reported recently (5).

In spite of their low cooling power, there are certain advantages that make Peltier refrigerators more desirable. For example, they have a smaller size, they use no refrigerant, and they lack moving mechanical parts. Also among their features are low weight, maintenance-free operation, and extreme silence, which make them of major interest for military applications. Another application is related to the property that it is very simple to control the rate of cooling by adjusting the current, while reversal of the current direction transforms the cooling device into a heater. This makes Peltier coolers/heaters very useful units for temperature control system. Adjusting the current within some range of values, heat can either be removed or added to one of the junctions that is in contact with the device whose temperature we want to be stabilized. In general, thermoelectric units can be part of miniature electronic and optoelectronic devices. For example, incorporation of a Peltier cooler/heater was suggested to stabilize the output wavelengths of a scanning laser diode with broad thermal scanning range (6).

As a final comment, it should be mentioned that the marathonian search for more exciting and exotic thermoelectric materials at room temperatures still continues (7).

## PELTIER REFRIGERATION IN THE MILLIKELVIN TEMPERATURE RANGE

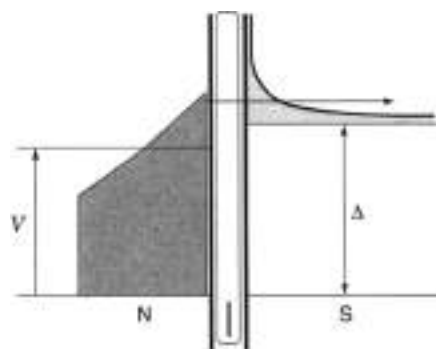
At very low temperatures—below a few hundred millikelvin—the coupling between electron and phonons in metals is very poor, and as a result the heat-energy flow between the phonon system (lattice) and the electron gas can be negligible. Therefore, electrons and phonons can be viewed as almost independent thermodynamic systems with well defined, but different temperatures. The energy transfer between lattice and electrons can be derived quantum mechanically by calculating the rate at which electrons exchange (absorb or emit) energy with phonons. The final result is calculated to be (8)

$$P_{e-l} = \Sigma V^2 T^6 - T_l^6 \quad (23)$$

where  $V$  is the volume of the metal,  $\Sigma$  differs from material to material, but it is a constant and in general depends on the strength of electron-phonon coupling. When the electronic temperature  $T$  is equal to lattice temperature  $T_l$ ,  $P_{e-l}$  vanishes as it should, since then electrons and phonons are in thermal equilibrium, and therefore there is no energy flow from one system to the other. Most notably, the minimum electronic temperature that can be achieved when some small incident Joule heat  $P_{Joule}$  is deposited in the metal, possibly due to some external voltage source, is

$$T_{min} = (P_{Joule}/\Sigma V^2)^{1/6} \quad (24)$$

Traditionally, in order to achieve such low temperatures, the metals are immersed into liquid helium, a rather indirect procedure because the lattice is cooled first. A novel technique was suggested recently (9) that exploits the thermal transport properties of a normal-metal-insulator-superconductor (NIS) tunnel junction, in order to decrease the electronic temperature of the metal (usually called *normal metal* because of the absence of superconductivity which can be suppressed by an applying magnetic field). The principle of the Peltier cooling in hybrid superconducting structures can be understood as follows. At zero temperature, electrons in the normal metal are distributed among the energy levels in a way that all levels with energy below the Fermi energy are occupied and all states with higher energy are empty. Namely, their distribution function is a step function. At very low temperatures the occupational probabilities of the energy states are determined by the Fermi-Dirac distribution function: approximately a step function with a thermal smearing of  $k_B T$  about the Fermi energy. The electrons lying in a strip of energy  $k_B T$  above the Fermi energy carry more energy than the rest. If we manage to extract only those electrons from the normal metal, then the smeared distribution function will be sharpened. Consequently the electronic temperature will be lowered. This effect can be achieved with the help of the adjacent superconductor, which possesses a gap in the excitation spectrum. When the biased voltage of the NIS tunnel junction is close to the superconducting gap  $\Delta$  and the tunnel barrier between the normal metal and the superconductor is very strong, only the hot electrons of the normal metal can tunnel effectively to the superconductor, since for the rest there are no states available in the superconducting electrode. Figure 2 illustrates the highlights of refrigeration in



**Figure 2.** Energy-space diagram for the NIS junction. Biasing the junction at voltages about the superconducting gap, the highly excited electrons from the normal metal tunnel to the superconducting region where unoccupied electronic states are available. The horizontal arrow describes the tunneling process. (Based on Ref. 9.)

an NIS junction. In this space-energy diagram, all energies are measured from the chemical potential of the superconductor. When the junction is biased about the superconducting gap, only electrons from the tail of the distribution of the normal metal can tunnel to the superconductor where they can find available empty states (shown by the arrow).

Mathematically, the heat transport through the NIS junction can be formulated along the lines of tunneling Hamiltonian theory. Under the condition that the tunneling probability through the barrier (insulating region) is small, then the heat current  $P$  out of the normal electrode and the electric current through the junction are

$$P(V) = \frac{1}{e^2 R_T} \int_{-\infty}^{\infty} dE N(E) E (E - eV) (f(E - eV) - f(E)) \quad (25)$$

$$I(V) = \frac{1}{e R_T} \int_{-\infty}^{\infty} dE N(E) (f(E - eV) - f(E)) \quad (26)$$

where  $R_T$  is the resistance of the tunnel junction when both electrodes are in the normal state (9–11). It is assumed that the two electrodes (N and S) have the same temperature.  $f(E)$  denotes the equilibrium distribution function for the electrons:

$$f(E) = \frac{1}{1 + \exp(E/k_B T)} \quad (27)$$

$N(E)$  denotes the density of states in the superconducting region, given by the BCS theory

$$N(E) = \frac{|E|}{\sqrt{E^2 - \Delta^2}} \Theta(E - \Delta) \quad (28)$$

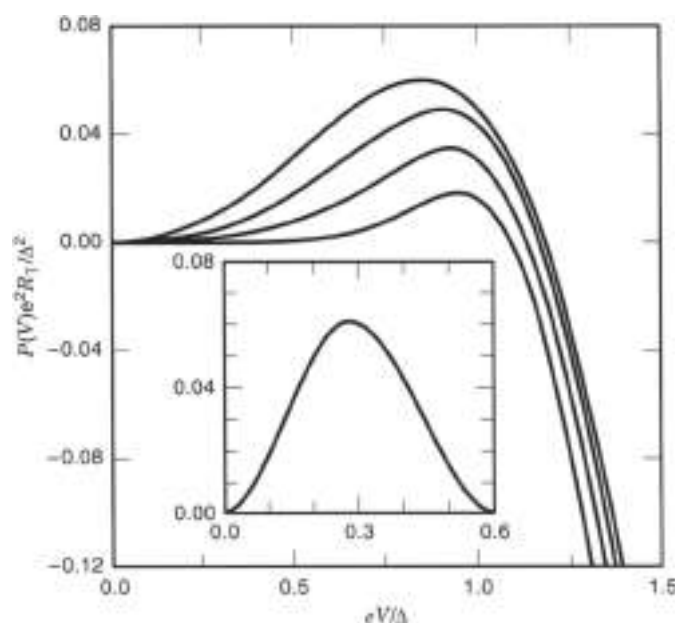
and taken to be zero within the gap region.

Several properties can be deduced from Eq. (25) for the heat current. First of all, it is a symmetric function of the voltage,  $P(-V) = P(V)$ ; that is, the heat flows out of the normal metal regardless of the direction of the electric current. Second, it is inversely proportional to the resistance of the junction. Namely, the cooling power will be increased if the junction area or the

conductivity of the insulating barrier is increased. One then would expect that when transmission probability through the contact between the normal metal and the superconductor increases, the cooling power will be magnified.

However, as was demonstrated (10), at larger transparencies coherent two-electron tunneling (Andreev reflection) begins to dominate the electron transport and suppresses the flow of heat from the N region to the S region. Andreev reflection is a special mechanism of transport in which a quasiparticle from the normal metal with energy below the superconducting gap combines with a Cooper pair from the superconducting electrode to produce a hole in the normal metal that travels in the opposite direction of the incident electron. It can be also visualized as a Cooper pair breaking that creates an electron and a hole that travel coherently in the normal metal. Andreev reflection process dominates the transport in nanocontacts (NS interfaces), and as an imminent result from its existence, electrons with all energies, including those with energies below the superconducting energy gap, can tunnel to the S electrode. Thus, it limits the cooling power of the NIS junction. Moreover, the heat current exhibits a nonmonotonic dependence on the interface transparency: It increases at small transparencies and decreases at larger ones. The interplay between the single-electron tunneling process and the Andreev reflection type of transport, determines the crossover value for the transmission probability which maximizes the cooling power (10). At low temperatures the transition value of the transparency is proportional to the ratio  $(k_B T/\Delta)^2$ .

Figure 3 shows the heat current as a function of the bias voltage for different temperatures calculated numerically using Eq. (25). The optimal value of the heat current is obtained for voltages about the gap  $V \approx \Delta/e$  and has a nonmonotonic behavior with respect to the temperature. When the applied



**Figure 3.** The heat current for a NIS junction as a function of the bias voltage. From top to bottom the curves correspond to temperatures  $T = 0.3, 0.2, 0.15, 0.1$  K. The inset shows the normalized heat current  $P(V)_{opt}/(e^2 R_T^{-1} \Delta^2)$  under optimal bias voltage conditions, as a function of the temperature  $k_B T/\Delta$ . (From Ref. 10.)

voltage is less than the gap, heat is extracted from the normal metal and dissipated in the superconducting region through electron-phonon collisions, but for higher voltages  $P$  becomes negative and the normal metal is heated. The inset of Fig. 3 shows the heat current at the optimal bias voltages as a function of the temperature. The maximum value of the optimal heat current  $P \approx 0.06 \lambda^2 / e^2 R_T$  is reached at  $k_B T \approx 0.1 \lambda$  and decreases at lower temperatures as  $(k_B T / \Delta)^2 \ll 10$ . In particular, for temperatures smaller than the gap,  $k_B T \ll \Delta$ , we can obtain from Eq. (25) the following result for the cooling power of the junction:

$$P(\Delta, \varphi) \approx \frac{\sqrt{7} \sqrt{2} - 1}{4} \frac{(3/2) \lambda^2}{e^2 R_T} \left( \frac{k_B T}{\Delta} \right)^{1/2} \quad (29)$$

The power supplied by the voltage source is  $P_e(V) = VI(V)$ . The electric current at the optimal bias, for low temperatures, can be calculated in a similar way as the heat current. As a result, the supplied power at the optimal voltage is

$$P_e(\Delta, \varphi) \approx \frac{\sqrt{7} \sqrt{2} - 1}{\sqrt{2}} \frac{(3/2) \lambda^2}{e^2 R_T} \left( \frac{k_B T}{\Delta} \right)^{1/2} \quad (30)$$

If we introduce the efficiency at the optimal point of operation of the device,  $\eta = P(\Delta, \varphi) / P_e(\Delta, \varphi)$ , Eqs. (29) and (30) read

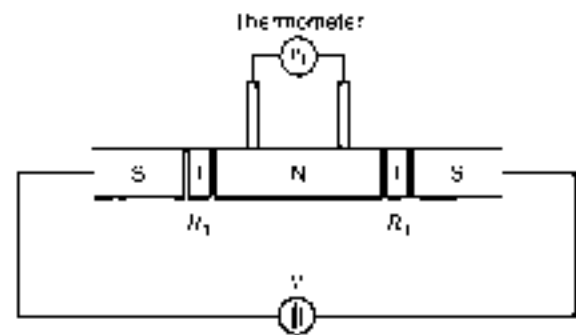
$$\eta = \frac{\sqrt{2} (3/2) k_B T}{4 (3/2) \lambda} \quad (31)$$

For typical values of the gap for conventional low-temperature superconductors  $\Delta \approx 2$  K and  $T = 200$  mK, the estimated efficiency is of the order of 7%.

In the initial experiment (9), a small metallic film of copper with volume  $V = 0.4 \mu\text{m}^3$  and ambient temperature of  $T = 100$  mK was cooled down to 85 mK. The barrier resistance  $R \approx 10$  k $\Omega$  was much larger than the resistance of the metallic island, approximately  $10 \Omega$ . The cooling power was about 7 fW at the ambient temperature. The refrigerating device was consisted of a tunnel junction between Cu (normal metal) and an aluminum superconducting electrode (NIS). The current was driven through the system with another superconducting electrode (Pb) which was in metallic contact (SN) with the normal thin film.

A dramatic improvement of the performance of the NIS microrefrigerator was achieved recently (11). Leiva et al. (11) combined two NIS junctions in series to form a symmetric SINIS hybrid superconducting structure. Because the heat power is a symmetric function of the voltage, when the junction is biased symmetrically at the optimal points, that is,  $V \approx \pm \lambda/e$ , even though the current passes through the normal metal in one direction, the heat flows out of the metallic island through both junctions. Moreover, one of the advantages of this structure is the efficient thermal isolation of the metallic island that we desire to refrigerate. The junction resistances were approximately 1 k $\Omega$ , and the volume of the metallic island was about  $0.05 \mu\text{m}^3$ . Starting from 300 mK, a 75% decrease in the electronic temperature was obtained. The achieved cooling power 1.5 pW was three orders of magnitude larger than before.

A schematic diagram of the SINIS microrefrigerator is depicted in Fig. 4. The two superconducting electrodes on the



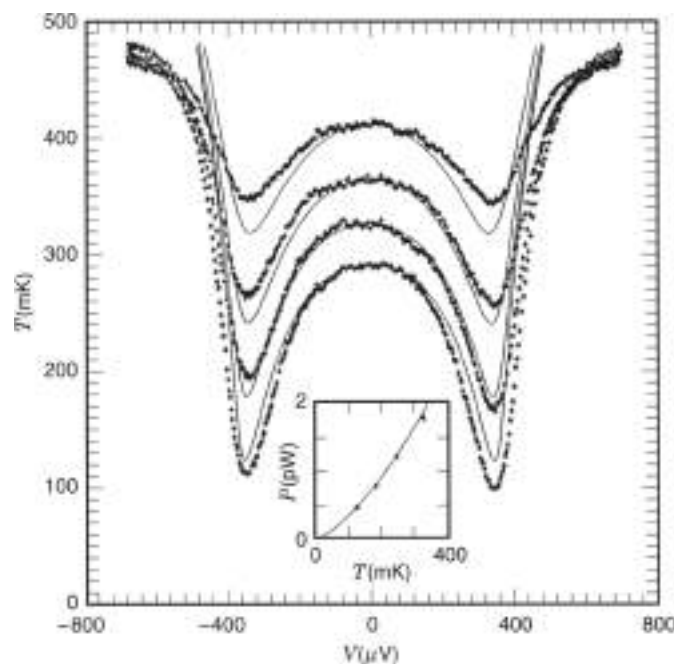
**Figure 4.** A schematic of the SINIS Peltier refrigerator. A metallic island (N) is connected via thin insulating regions (I) to superconducting electrodes (S). The two junctions are characterized by the same resistance,  $R_1$ . The symmetrically biased junctions effectively reduce the electron temperature of the island, which is recorded in the thermometer. (Based on Ref. 11.)

top of the metallic island can be used to measure its temperature. The current-voltage characteristic ( $I - V$ ) is recorded, from which the temperature of the metal can be calculated. It should be mentioned that in the tunneling limit the ( $I - V$ ) characteristics depend only on the electronic temperature. Figure 5 shows the performance of the SINIS Peltier refrigerator at various ambient temperatures (value of  $T$  when  $V = 0$ ). The value of the bias voltage at the temperature dips is approximately  $V \approx 2\Delta/e$ ; each junction is biased symmetrically  $eV/2$  with respect to the metal. Namely, the superconducting gap of the aluminum is  $\lambda = 180 \mu\text{eV}$ . The voltage dependence of the temperature can be calculated from the energy balance equation,

$$P(V) = P_{cool} = \Sigma U(T^e - T_l^e) \quad (32)$$

For  $V = 0$  we have  $T_l = T$ . For the SINIS system, one has to calculate in a self-consistent way the electric current and heat current because there is a voltage drop across the normal metal which shifts the chemical potential of the metallic island. The temperature dependence of the maximum cooling power is plotted in the inset of Fig. 5.

When comparing the NIS Peltier microrefrigerator to conventional millikelvin refrigeration schemes like adiabatic demagnetization or dilution cryostats, there are many advantages and disadvantages. Some of the advantages are: smaller size and weight (can operate even at zero gravity), faster cooling; easier to construct, operate, and maintain; reliability. One of the major drawbacks is the small cooling power: the superconducting electrodes can remove only some picowatts of power, and the cooling is basically efficient only for small volume conductors. The cooling power of a dilution refrigerator at temperatures below 100 mK is in the range of tens to hundreds of microwatts. In addition, the cooling properties of the NIS junctions are related only to the electronic temperature. Attempts to cool down the lattice based on tunneling principles show even smaller cooling power. However, the NIS Peltier microrefrigerators are of importance for devices that dissipate small amounts of power like x-ray or infrared detectors. Furthermore, cryogenic NIS junctions can be used as sensitive bolometers to detect particles and radiation. When some incident radiation is absorbed by the metal, it affects mostly the electronic temperature due to



**Figure 6.** Performance of the SINIS cooling device. By varying the bias voltage the temperature of the metallic island is recorded. Each curve corresponds to different ambient temperature. Solid lines represent the theoretical fit with one fitting parameter  $\tau$ . In the inset the optimal cooling power is plotted as a function of the ambient temperature. Asterisks represent the experimental result, and the solid line is twice the result of Eq. (25). (From Ref. 11.)

weak coupling of lattice with the electrons. The extra energy is deposited in a form of highly energetic thermal excitations that tunnel to the superconductor. Temperature increments can be detected via changes in the  $I$ - $V$  characteristics of the device (12). As a final remark, Peltier microrefrigerator can serve as an alternative solution for the cooling of bolometric detectors (which operate best below 100 mK) and as temperature sensors that are carried by satellites for astronomical observations.

## BIBLIOGRAPHY

1. D. K. C. MacDonald, *Thermoelectricity: An Introduction to the Principles*, New York: Wiley, 1982.
2. R. R. Heikes and R. W. Ure, Jr., *Thermoelectricity: Science and Engineering*, New York: Interscience, 1961.
3. H. J. Goldsmid, *Thermoelectric Refrigeration*, New York: Plenum, 1961.
4. L. D. Hicks and M. S. Dresselhaus, Thermoelectric figure of merit of a one-dimensional conductor, *Phys. Rev. B* **47**: 16631, 1993.
5. H. J. Goldsmid et al., High- $T_c$  superconductors as passive thermoelements, *J. Phys. D* **21**: 344, 1988; T. Nishino and T. Hashimoto, Refrigeration character of new-type Peltier refrigerator using high- $T_c$   $\text{BiPb}_{2-x}\text{Sr}_{1-x}\text{Ca}_2\text{Cu}_3\text{O}_y$  superconductors, *Ion. J. Appl. Phys.*, **33**, 1726, 1994.
6. D. Ja. S. J. Hoering, and J. G. Baker, A scanning temperature control system for laser diodes, *Meas. Sci. Technol.*, **4**: 111, 1993.
7. G. Mahan, R. Sales, and J. Sharp, Thermoelectric materials: New approaches to an old problem, *Phys. Today*, March, **42**, 1997.
8. E. C. Wellstood, C. Urbina, and J. Clarke, Hot-electron effects in metals, *Phys. Rev. B* **49**: 5942, 1994.
9. M. Nahum, T. M. Eiles, and J. Martinis, Electronic microrefrigerator based on a normal insulator-superconductor, *Appl. Phys. Lett.*, **65**: 3123, 1994.
10. A. Bardas and D. Averin, Peltier effect in normal metal-superconductor (NS) microcontacts, *Phys. Rev. B*, **52**: 12973, 1995.
11. M. M. Leivo, J. E. Pukola, and D. Averin, Efficient Peltier refrigeration by a pair of normal metal/insulator/superconductor junctions, *Appl. Phys. Lett.*, **68**: 1996, 1996.
12. M. Nahum, J. M. Martinis, and S. Cutler, Hot electron microcalorimeters for X-ray and phonon detection, *J. Low Temp. Phys.*, **90**: 733, 1993; A. Reiss, M. Martin, and B. Panneter, Tunnel junction used as the thermometer for ions detection, *J. Low Temp. Phys.*, **93**: 727, 1993.

ATHANASSIOS BARNAS  
State University of New York at  
Stony Brook

## POWER QUALITY

Society is becoming more and more reliant on machines, devices, and processes that are sensitive to the quality of the electrical power they use. At the same time, the generation and distribution system for electrical power is becoming more and more complex and vulnerable to internal and external perturbations that disrupt the quality of power available to the end user. For these reasons, there is a large market developing for high-quality power, and several technologies that can provide such power have been or are being developed. In this article, we discuss some of these power quality issues and provide some insight into one of the technologies proposed to provide high-quality power. The first part of the article focuses on power quality issues and low microSMES (Superconducting Magnetic Energy Storage) technology is inserted into the power flow stream. The second part is devoted to a discussion of the fine points of microSMES technology from a design and performance point of view.

## POWER QUALITY

The term *power quality* encompasses a broad range of technical issues that concern everyone, from the casual computer user to the equipment manufacturer, the industrial plant manager, and the utility transmission engineer. The technical issues associated with power quality at all these levels are complex, and they are often disguised behind misconceptions, vague definitions, and misapplied solutions. The lack of a consistent vocabulary, difficulty in fully characterizing power quality at the ultimate point of use, and confusion over the applicability of available solutions often encourages an adversarial relationship between the user, the equipment manufacturer, and the utility provider.

The technical community as well as the general public is increasingly concerned about power quality factors for a number of interrelated reasons. Some of the issues behind this growing apprehension include:

- Increasing levels of automation and flexibility have caused a proliferation of microprocessor-based controls and power electronic systems that are more sensitive to power quality variations than older electromechanical devices.



- Utility operating procedures designed to minimize sustained interruptions have increased the frequency of short-duration interruptions through wider use of line reclosers and instantaneous breaker operations.
- A growing emphasis on energy conservation has encouraged greater use of high efficiency adjustable-speed motor drives and nonlinear power supplies that generate high levels of harmonics and heating of neutral conductors.
- The increasing sophistication of industrial processes and their reliance on expensive raw materials have greatly multiplied the cost of recovering from disruptions caused by poor power quality.

These factors combine to place a tremendous economic burden on industry, particularly in high-tech operations such as semiconductor wafer fabrication. Estimates of this economic impact on US industry alone range from \$13.3 billion to \$25.6 billion per year [1].

### Power Quality Standards

The definition of power quality terms and standard methods for measuring and characterizing electromagnetic phenomena have received considerable attention in recent years from numerous standards bodies. The principal bodies involved

in establishing international power quality standards are the Institute of Electrical and Electronics Engineers (IEEE) and the International Electrotechnical Commission (IEC). Other North American standards organizations involved with power quality issues include the American National Standards Institute (ANSI), National Electrical Manufacturers Association (NEMA), Underwriters Laboratory (UL), Canadian Standards Association (CSA), and Semiconductor Equipment and Materials International (SEMI). European organizations involved in similar activities include the European Union Standards Organization (CENELEC) and the International Conference on Large High Voltage Electric Systems (CIGRE).

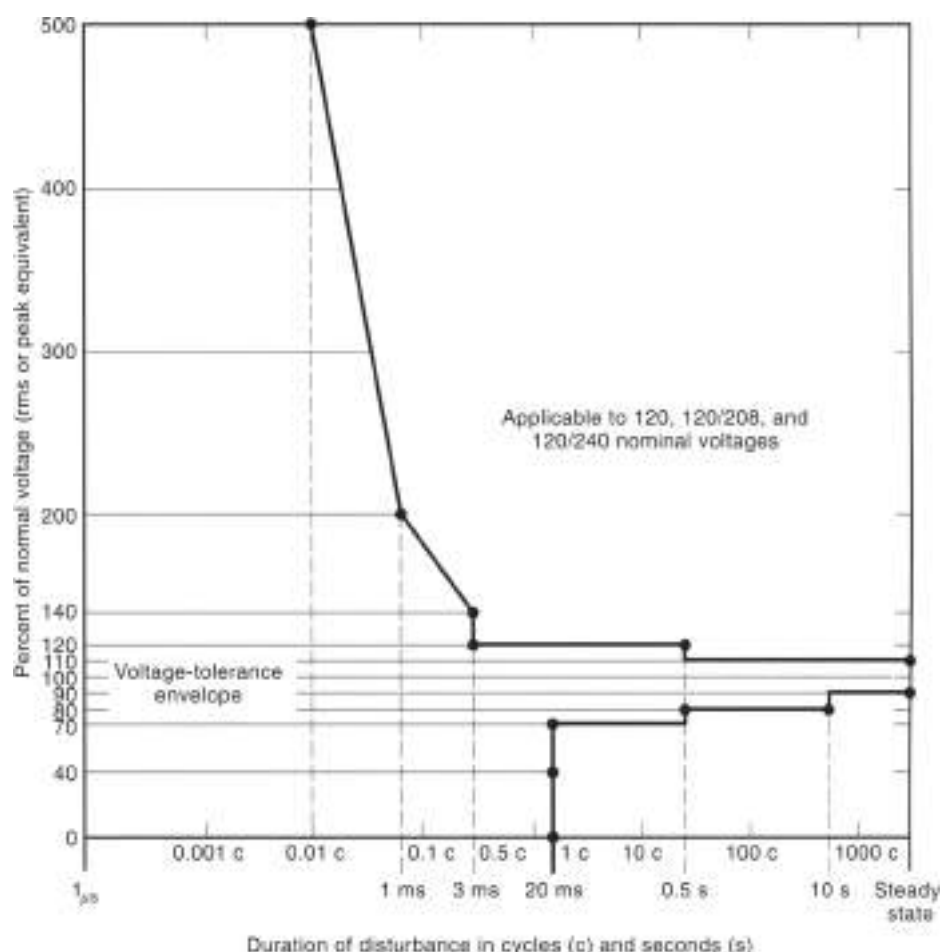
The electromagnetic disturbances defined in the standards and recommended practices produced by these organizations can be classified in several ways, depending on the perspective of the defining body. Table 1 is an example of a particularly comprehensive classification system that not only defines a generally accepted term for each type of disturbance but also gives typical values for spectral content, duration, and magnitude where appropriate [2].

### Power Quality Data

The most frequently referenced method of representing power quality data and equipment compatibility information is the CBEMA curve, originally developed by the Computer Business

Table 1. Classification of Power Quality Disturbances

| Categories                     | Typical Spectral Content       | Typical Duration | Typical Voltage Magnitude |
|--------------------------------|--------------------------------|------------------|---------------------------|
| 1.0 Transients                 |                                |                  |                           |
| 1.1 Impulsive                  |                                |                  |                           |
| 1.1.1 Nanosecond               | 5 ns rise                      | < 50 ns          |                           |
| 1.1.2 Microsecond              | 1 $\mu$ s rise                 | 50 ns-1 ms       |                           |
| 1.1.3 Millisecond              | 0.1 ms rise                    | < 1 $\mu$ s      |                           |
| 1.2 Oscillatory                |                                |                  |                           |
| 1.2.1 Low frequency            | < 5 kHz                        | 0.5-50 ms        | 0-4 pu                    |
| 1.2.2 Medium frequency         | 5-500 kHz                      | 20 $\mu$ s       | 0-8 pu                    |
| 1.2.3 High frequency           | 0.5-5 MHz                      | 5 $\mu$ s        | 0-4 pu                    |
| 2.0 Short-duration variations  |                                |                  |                           |
| 2.1 Instantaneous              |                                |                  |                           |
| 2.1.1 Interruption             |                                | 0.5-30 cycles    | < 0.1 pu                  |
| 2.1.2 Sag (dip)                |                                | 0.5-30 cycles    | 0.1-0.9 pu                |
| 2.1.3 Swell                    |                                | 0.5-30 cycles    | 1.1-1.8 pu                |
| 2.2 Momentary                  |                                |                  |                           |
| 2.2.1 Interruption             |                                | 30 cycles-3 s    | < 0.1 pu                  |
| 2.2.2 Sag (dip)                |                                | 30 cycles-3 s    | 0.1-0.9 pu                |
| 2.2.3 Swell                    |                                | 30 cycles-3 s    | 1.1-1.4 pu                |
| 2.3 Temporary                  |                                |                  |                           |
| 2.3.1 Interruption             |                                | 3 s-1 min        | < 0.1 pu                  |
| 2.3.2 Sag (dip)                |                                | 3 s-1 min        | 0.1-0.9 pu                |
| 2.3.3 Swell                    |                                | 3 s-1 min        | 1.1-1.2 pu                |
| 3.0 Long-duration variations   |                                |                  |                           |
| 3.1 Interruption, sustained    |                                | > 1 min          | 0.0 pu                    |
| 3.2 Undervoltages              |                                | > 1 min          | 0.8-0.9 pu                |
| 3.3 Overvoltages               |                                | > 1 min          | 1.1-1.2 pu                |
| 4.0 Voltage unbalances         |                                | Steady state     | 0.5-2%                    |
| 5.0 Waveform distortion        |                                |                  |                           |
| 5.1 dc offset                  |                                | Steady state     | 0-0.1%                    |
| 5.2 Harmonics                  | 0-100% <sub>THD</sub> harmonic | Steady state     | 0-20%                     |
| 5.3 Interharmonics             | 0-6 kHz                        | Steady state     | 0-2%                      |
| 5.4 Notching                   |                                | Steady state     |                           |
| 5.5 Noise                      | Broadband                      | Steady state     | 0-1%                      |
| 6.0 Voltage Fluctuations       | 25 Hz                          | Intermittent     | 0.1-7%                    |
| 7.0 Power frequency variations |                                | < 10 s           |                           |

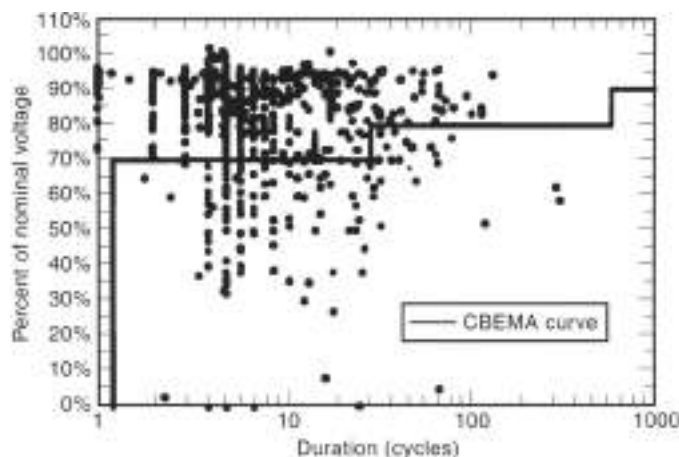


**Figure 1.** Revised voltage tolerance envelope developed by CBEMA to characterize compatibility at the user's load in terms of the amplitude voltage magnitude and duration of the disturbance. This de facto standard was intended to provide a benchmark for measuring the ride-through capability of sensitive electronic equipment against the quality of the available electric supply.

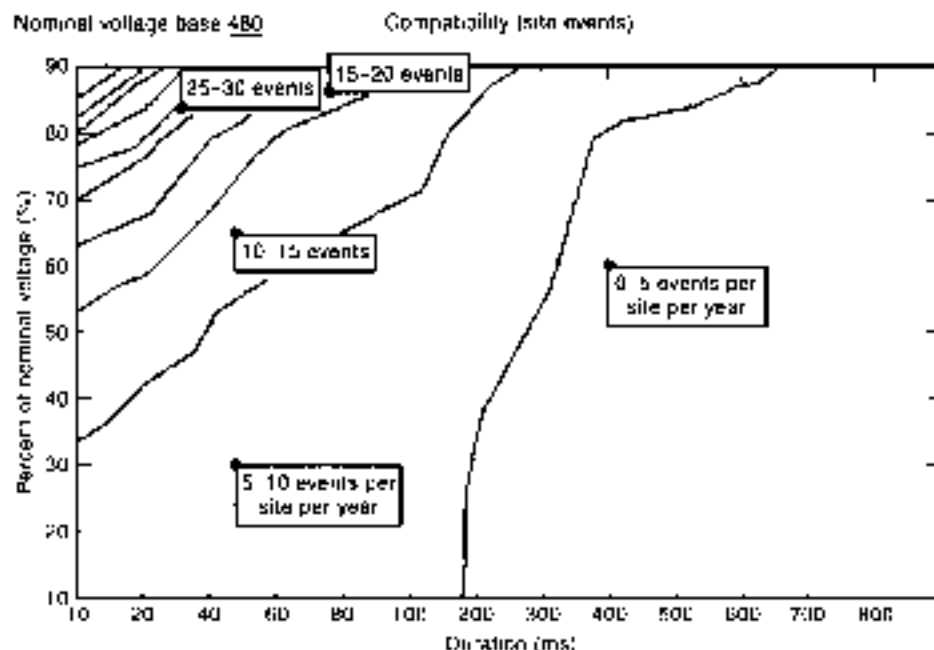
Equipment Manufacturers Association. This organization, now known as the Information Technology Industry Council (ITIC), approved a revised version of the CBEMA curve in 1996. The revised curve, shown in Fig. 1, specifies an envelope of acceptable power quality defined in terms of voltage magnitude and duration. The curve assumes that equipment will ride through disturbances falling within the envelope without any malfunctions. It is assumed that disturbances that fall below the envelope may cause the load equipment to drop out because of a lack of energy, whereas disturbances that fall above the envelope may lead to overvoltage trips, insulation breakdown, and other problems. Unfortunately, the actual voltage tolerance of electronic equipment varies quite widely and rarely approximates the CBEMA curve. Nevertheless, the CBEMA curve has become a de facto reference for evaluating the quality of the electrical supply and the tolerance of sensitive equipment to voltage disturbances.

Figure 2 shows a plot of voltage sag data collected from 15 semiconductor manufacturing sites overlaid on the lower portion of the CBEMA envelope. The disturbance data are a subset of the data collected as part of the EPRI Distribution Power Quality (DPQ) study (3). The DPQ study collected data on a representative sample of utility distribution feeders, with monitors placed on 100 different feeders at 24 geographically dispersed utilities over a period of approximately 2 years. The data shown in Fig. 2 was restricted to data acquired at semiconductor plants to characterize the electrical environment faced by large industrial customers served by dedicated substations.

These sites generally have fewer disturbances than the average substation serving a mix of commercial and residential loads. Even so, the semiconductor plants experienced 1076 disturbances, with 166 of the events lying below the CBEMA curve. This resulted in an average of 5.4 sags and interruptions per



**Figure 2.** Scatterplot of EPRI DPQ study data overlaid on the relevant portion of the revised CBEMA curve. The 1076 disturbance events were acquired at 15 semiconductor manufacturing plants over a period of approximately 2 years. The number of disturbances that fell on or below the CBEMA curve was 166, giving an average of 5.4 events per year per site outside the voltage tolerance envelope.



**Figure 3.** Contour plot of EPRI DPQ study data. The contour lines represent the average number of disturbance events per site that are more severe (having a lower voltage and longer duration) than the magnitude and duration values along the contour line. For example, there were 15 events per year with a magnitude less than 60% of nominal voltage and a duration longer than 20 ms.

year at each site that fell outside the CBEMA voltage tolerance envelope.

The DPQ disturbance data is shown in contour format in Fig. 3 to display the frequency of occurrence for sags and interruptions as a function of their magnitude and duration values (4). This visualization method clearly indicates that the vast majority of the disturbances are relatively small amplitude sags of short duration. Statistical analysis of the DPQ data reveals:

- The average interruption rate was 0.6 per month per site;
- The average sag rate was 4 per month per site (10% < V < 90%);
- The ratio of voltage sags to interruptions was approximately 10:1;
- Most voltage sags had a duration of less than 167 ms (10 cycles at 60 Hz).

From this data, we can conclude that power conditioning equipment providing ride-through protection times of at least 1 s can protect sensitive industrial loads against the vast majority of the electrical disturbances they will encounter.

#### Superconducting Magnetic Energy Storage Configurations

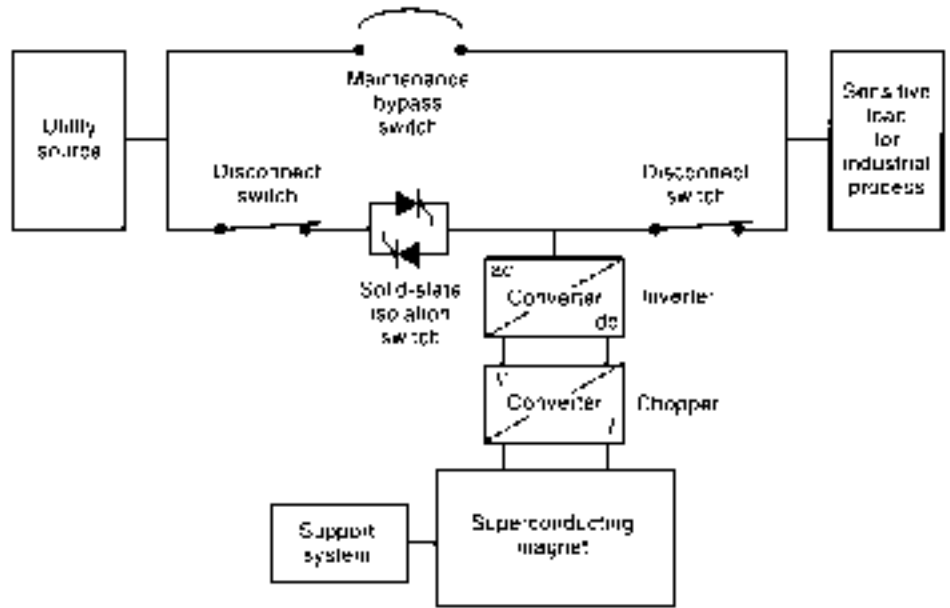
The acronym SMES was coined in the 1970s at the University of Wisconsin. It stands for superconducting magnetic energy storage, a concept whereby energy is stored in the form of a magnetic field. The field is created by a superconducting magnet, which has virtually no electrical resistance when cooled to the boiling point of helium. The lack of electrical resistance means that there are no storage losses; so in principle SMES is a very efficient means of storing energy.

The basic energy storage element of a microSMES system can be connected to the utility and the protected load in a number of ways to maximize protection capabilities for different applications. The operating requirements for the magnet

subsystem are fairly similar for these various configurations, whereas the power electronics interface is quite specific to the target application. Several potential application classes will be briefly discussed in the following sections.

**Shunt-Connected System.** A simplified block diagram of a shunt-connected microSMES application is shown in Fig. 4. This configuration can be used to protect virtually any type of sensitive ac load or industrial process. The load is initially fed from the utility source through the normally closed mechanical disconnect switches on either side of the solid-state isolation switch. When a voltage sag or interruption occurs on the utility source, the solid-state isolation switch is opened to isolate the load from the unreliable utility power. The microSMES unit instantly begins to discharge through the inverter so that the load is seamlessly transferred from the utility source to the energy storage system. If utility power is restored during the discharge, the microSMES unit is synchronized to the grid, and the solid-state isolation switch transfers the load back to the utility source. The microSMES unit is then recharged to full-energy storage capacity in preparation for additional protection cycles.

**Series-Connected System.** A series-connected microSMES unit is shown in the simplified block diagram of Fig. 5. In this application, the microSMES unit is configured to compensate for voltage sags by providing a voltage boost in series with the load. Because the microSMES unit does not support the full load as it does in the shunt-connected configuration, the same amount of energy storage can support a larger electrical load for an equivalent time or the same size load for a longer time. The load is normally fed from the utility source through the solid-state isolation switch bypassing the injection transformer. When a voltage sag is detected, the solid-state switch opens, and the load current flows through the injection transformer. The inverter then injects just enough energy into the transformer to maintain the desired voltage level at the load. The isolation switch closes when nominal utility voltage is restored and the



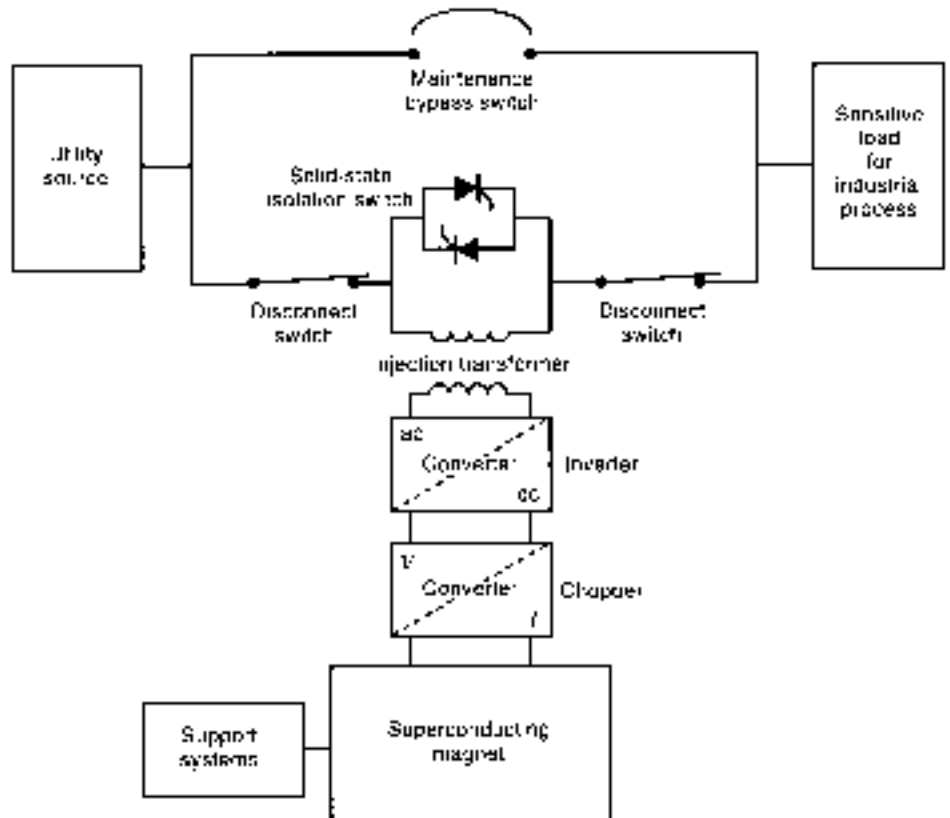
**Figure 4.** Simplified block diagram of a shunt-connected microSMES system. This configuration protects sensitive ac loads against dips and interruptions by disconnecting from the utility service during a disturbance and supplying the load with energy stored in the superconducting magnet.

microSMES unit is recharged and placed in standby mode in preparation for the next event.

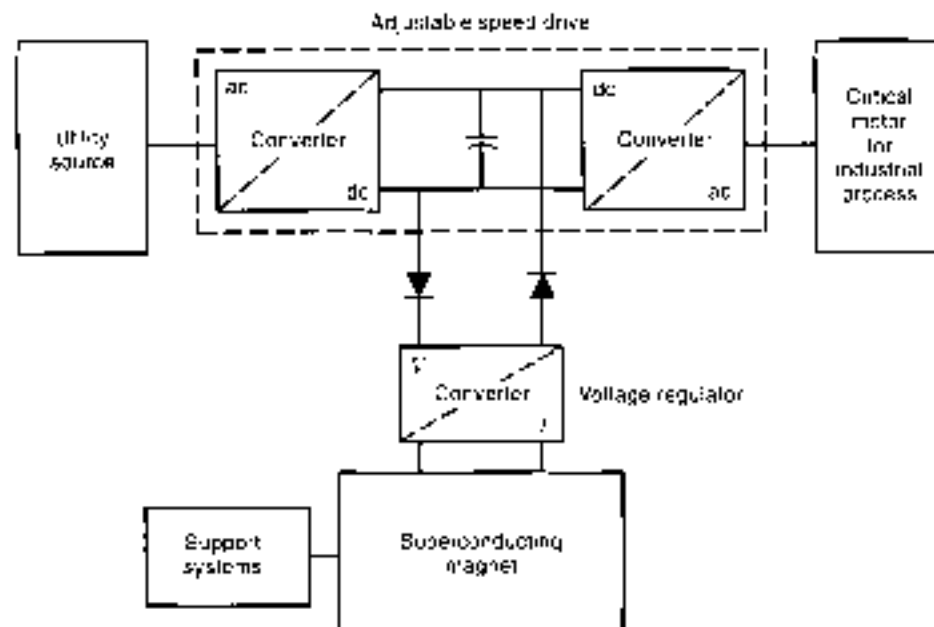
**Motor Drive System.** In adjustable speed motor drive applications, the microSMES unit supplies a dc voltage rather than the ac voltage supplied in shunt- and series-connected applications. As shown in the block diagram of Fig. 6, the dc voltage from the microSMES unit is fed into the dc link between the rectifier and inverter of the motor drive itself. This configuration eliminates the need for a separate inverter for the microSMES

system. The voltage regulator in the microSMES system simply ensures that the dc link capacitor remains fully charged in the event of a loss of utility power feeding the rectifier.

**Hybrid UPS System.** A second configuration requiring dc output from the microSMES is shown in Fig. 7, where the microSMES unit is integrated with a conventional battery Uninterruptible Power Supply (UPS) system. The microSMES voltage regulator monitors the dc link of the UPS and supplies energy from the magnet whenever the link voltage drops because



**Figure 5.** Simplified block diagram of a series-connected microSMES system. This configuration protects sensitive ac loads against voltage dips by boosting the voltage across the injection transformer to compensate for low voltages at the utility source.



**Figure 6.** Simplified block diagram of a microSMES system protecting an adjustable speed motor drive. This configuration eliminates the inverter in the microSMES system because the stored energy in the magnet is supplied directly to the dc link of the motor drive.

of a loss of power at the input rectifier. If the loss of input power persists, the batteries gradually begin to supply energy to the inverter as the stored energy in the microSMES is depleted. This configuration allows the microSMES unit to support the load completely during short disturbances and to slow down the rise time of the battery discharge for longer disturbances. This hybrid configuration can significantly enhance battery performance because battery lifetime is decreased by deep discharges with fast rise times.

### MicroSMES

The remainder of this article concentrates on the system that provides the energy to ride through a power disturbance. In particular, energy stored in a magnetic field is discussed. Magnetic energy storage systems are described in general, and compared with other forms of storage. Details of the design are discussed, with the purpose of identifying the design drivers that affect the economy of such systems. Finally, a generalized algorithm to arrive at a design that minimizes the cost required to store a unit of energy is developed to further illuminate the design issues related to microSMES.

### What is $\mu$ SMES?

Designs for SMES systems have been proposed at several sizes and stored energy scales. Early SMES studies focused on very large systems, which stored thousands of megawatt-hours of energy and were designed to provide peaking power for large utility baseloads. SMES systems storing tens of megawatt-hours have been proposed to provide stabilization of power distribution grids subject to transients that can disrupt the delivery of power over transmission lines. Some applications in the military and elsewhere require hundreds of megawatt-seconds of energy delivered at very high power that can be supplied by SMES systems. Finally, SMES systems storing on the order of 1 MWh to 10 MWh of energy are proposed to provide short-term high-quality power to equipment and processes sensitive

to voltage sags and/or power outages. The latter set of systems as a group is referred to as microSMES, or  $\mu$ SMES.

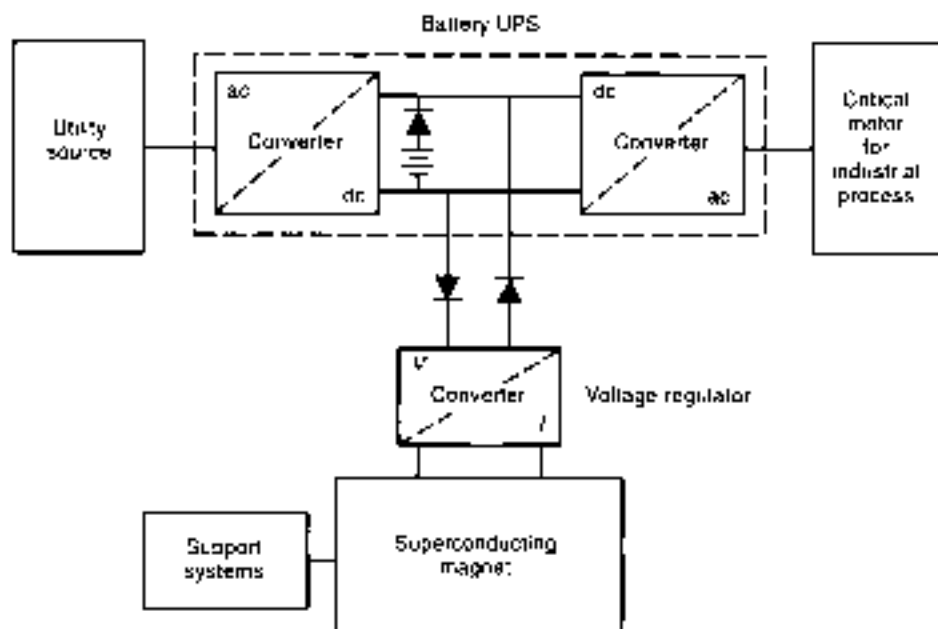
A typical block diagram of a  $\mu$ SMES system is shown in Fig. 6. There is a power electronics subsystem that provides an interface between the energy storage system (the magnet) and the power source/load. Depending on the application, there might be an auxiliary stand-by power source included in the system to handle long power outages. Ancillary subsystems support the magnet. It operates at liquid helium temperature and requires refrigeration and insulating vacuum systems to maintain operating conditions. The heart of the system is a superconducting magnet. This is an electromagnet wound with superconducting wire. When the wire is operated at liquid helium temperature (4.5 K, or  $-450$  F), it will carry an electric current with no resistance. Thus there are no resistive losses, and the storage of energy in the form of a magnetic field becomes possible.

The economics of  $\mu$ SMES are determined by these three major subsystems. Unlike battery storage, where all the cost is in the power electronics, or capacitors, where all the cost is in the energy storage system,  $\mu$ SMES costs are split between the power electronics and the storage, with nontrivial stand-by power needed to run the refrigeration system. Each of these subsystems must be optimized for maximum performance at minimum cost for  $\mu$ SMES to be successful in the marketplace.

Because the power electronics are somewhat common to all types of energy storage, the magnet system and refrigeration system must be optimized for SMES to be competitive with other technologies. For the magnet system, the maximum stored energy must be obtained using the minimum amount of material (superconducting wire). For the refrigeration system, minimizing the heat leaking into the cold space is the key to optimum performance. These issues will be discussed in greater detail later.

### How Does $\mu$ SMES Compare to Other Forms of Energy Storage?

Before describing a  $\mu$ SMES system in greater detail, some points about the technology should be made, with the purpose



**Figure 7.** Simplified block diagram of a hybrid microSMES/battery UPS system. In this configuration, the superconducting magnet supplies energy to the load during short-duration disturbances, and the batteries must be discharged only during sustained outages. The fewer number of cycles and lower discharge depth required from the batteries significantly increases their life time.

of driving out these applications where  $\mu$ SMES makes sense and where it doesn't make sense. A discussion of the desirable and undesirable qualities of SMES follows.

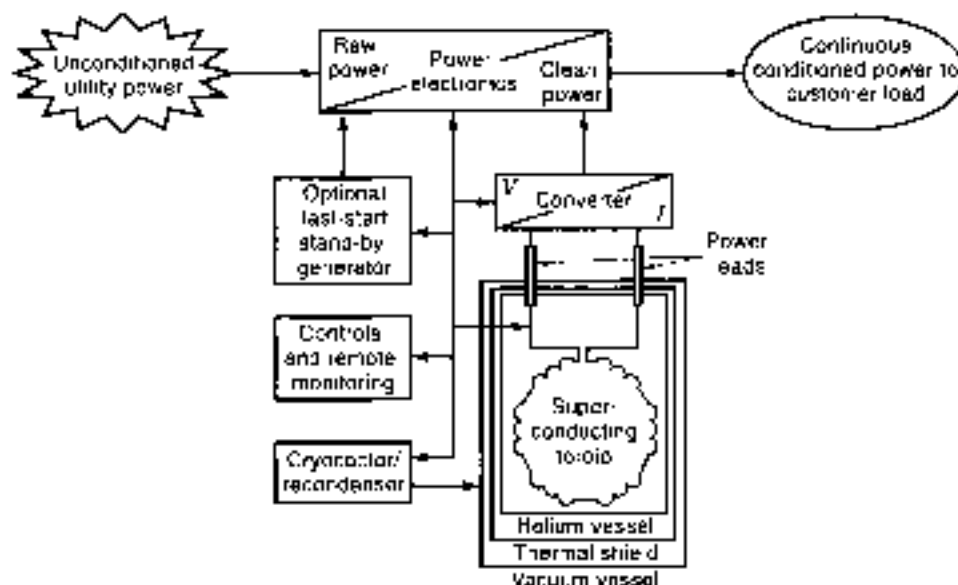
**Positive Attributes.** SMES technology has many virtues. It can provide power at an arbitrary level and do so virtually instantaneously. Furthermore, the life of the device is not limited by the power level or depth of discharge. A SMES system will provide power pulses throughout its design lifetime of up to 30 years with no degradation in performance.

The conversion from magnetic energy to electrical energy is very efficient. Because the magnetic field is created by an electric current circulating in a magnet, it is very straightforward to extract the energy. All that is needed is a voltage applied across the terminals of the magnet. The only loss is in the form

of heat dissipated in the leads and electronic devices in the power electronics system (PCS), which is minimal.

A large fraction of the stored energy can be extracted from the system. The only limitation is that the voltage must be applied to maintain the delivered power. As will be discussed in further detail later, the voltage must rise as the current drops during a discharge. Eventually the voltage will rise to a level where it can break down the insulation on the coil. The discharge must be stopped just before this happens. With proper design, it is possible that this limit is not reached until only a few percent of the original stored energy remains.

The storage system itself requires no maintenance. The only moving parts are in the refrigerator. Maintenance is required on the refrigeration system, but it is possible that it can be performed while the remainder of the system is operating.



**Figure 8.** Block diagram of a microSMES storage system, showing major subsystems. Each subsystem is a major cost driver for SMES technology and must provide optimum performance.

Thus SMES has the potential to have no down time for maintenance.

SMES is generally environmentally benign (see the discussion of fringe magnetic fields). It does not employ hazardous or toxic materials. There are no components requiring recycling or disposal during or at the end of its life.

Finally, the state of charge of the system is easily determined. A simple current or magnetic field measurement is all that is needed to ascertain that the system is charged and ready to deliver power.

**Negative Attributes.** SMES has some undesirable characteristics as well. Some are real, and some are falsely perceived by the market, but nonetheless have prevented SMES from being an overwhelming success to date.

First, SMES is an expensive technology. The superconducting wire used in the magnets is expensive to begin with. The cost of the labor and materials to fabricate the structure of the magnet, the helium containment vessel, the cryostat, and the supporting refrigeration system is and will always be an order of magnitude greater than the cost of batteries storing the same amount of energy.

The technology is esoteric and is perceived to be risky. Superconducting magnets are esoteric for several reasons. Because they operate at such cold temperatures, vacuum insulating technology and helium refrigeration technology are a necessary part of the design. Furthermore, it takes a trivial amount of energy to destroy the superconductive properties of the wire (e.g., the frictional heat generated if two adjacent wires rub together through the distance of  $4/100$ ths of a mil, or the heat generated by the creation of microcracks in an epoxy-impregnated coil when it is energized). If the superconductive property is lost, systems must be in place to protect the wire from overheating and burning out from the excessive resistive heat that is generated. In addition, there can be very high internal voltages in a superconducting magnet that has gone resistive, which can potentially arc through the electrical insulation around the wire and render the magnet useless at worst, or have much reduced capacity at best. These problems are usually encountered when a magnet is first energized, but rarely afterwards. A magnet might require several attempts before it can be fully ramped to its design current and field, but once it has been operated at design conditions, the risk of an unwanted quench virtually vanishes. The risk inherent in the technology is to the producer because if a magnet is not properly designed and fabricated, it becomes apparent only after it has been built. Then it must be scrapped. The risk to an end user is negligible once a magnet has been operated at design conditions. Over the years, the potential problems have come to be understood, and although they cannot be eliminated, they can be mitigated by paying close attention to proper design practice and quality control during fabrication. Indeed, superconducting magnets have reached a mature state of reliability for commercial applications: refer to magnetic resonance imaging (MRI) systems routinely operating in hospitals throughout the world.

It is possible that a SMES system will have fairly large fringe fields (i.e., the magnetic field will extend well beyond the physical limits of the magnet system). This can lead to siting problems or restrictions that can add to the cost. An alternative is to use toroidal magnet systems, where the field is confined totally within the bore of the toroid. This essentially eliminates the fringe field, with the penalty of requiring more su-

perconducting wire to produce the same amount of energy as a solenoid.

The maintenance required for the refrigeration system can be expensive. Depending on the size of the refrigerator required, maintenance can vary from requiring a simple seal replacement to rebuilding a compressor periodically. The cost of this maintenance must be factored into the life cycle cost of the installed system. Typical practice is to provide redundant components such as compressors, with an associated penalty of additional capital cost.

Finally,  $\mu$ SMES suffers from a penalty imposed by the Second Law of Thermodynamics. In order to operate at temperatures close to absolute zero, large amounts of power are required to operate the refrigerator. Typical helium refrigerators require from 500 W to 5000 W of compressor power to remove a single watt of heat load from a helium bath. Thus there are large parasitic power requirements to operate a system. In addition, the current that circulates through the magnet must also circulate through the power electronics system, creating waste heat that must be removed. This waste heat becomes an air conditioning load that represents additional parasitic power.

Because  $\mu$ SMES will always be more expensive than batteries, it should be considered only for applications where batteries cannot compete: power levels higher than 2 MW or so and situations requiring many repetitive discharge cycles lasting on the order of two or three seconds or less. These conditions exploit the positive attributes of  $\mu$ SMES to best advantage.

#### What Does the Storage System Look Like?

The storage subsystem of a  $\mu$ SMES unit consists of a magnet and a cryogenic system. The magnet is a superconducting coil or array of coils. The cryogenic system includes a cryostat that insulates the cold helium vessel from the ambient surroundings, and a helium liquefier.

The magnet may be a single coil or an array of coils. It may be wound in a solenoid configuration or in a toroid. The advantage of a solenoid is that it produces more stored energy per unit conductor than a toroid does, and it is much simpler to fabricate. The advantage of a toroid is that the magnetic field is confined almost completely within the bore of the torus.

It is not within the scope of this article to discuss the art of designing and fabricating superconducting magnets. However, insofar as they relate to the technology of  $\mu$ SMES, the main components of a magnet will be discussed. A superconducting magnet consists of a conductor, wound on some structure, with the conductor isolated from adjacent turns and the supporting structure by some insulation system.

The conductor represents the major materials cost of a  $\mu$ SMES system. It will be either a single monolithic wire, or a cable of smaller wires. At currents up to 2000 A to 3000 A, a single wire will do the job and is easier to fabricate and wind. At higher currents (higher powers), a cable is the preferred approach. The primary reason is that, as the size of wire increases, the ac losses during discharge become significant. These losses are reduced in a cable because of the resistance created at the many contact points between individual wires.

The structure must support the winding against the Lorentz loads created when a conductor is placed in a magnetic field. In a simple solenoid, the structure is trivial because the net force on the coil is zero (excepting its own weight). The forces

act to expand the coil radially outward and compress the coil axially. The winding itself is the structure. In solenoidal arrays, or toroids, there are significant forces that act on the individual coils, which must be reacted with structure. The nature of the forces depend on the geometry and arrangement of the coil array. For a toroid, a centering force acts to compress the toroid into as small a diameter as possible.

The insulation system is key to the economic viability of  $\mu$ SMES because it determines the maximum allowable voltage at which the coil can be discharged. The voltage level determines how much of the stored energy can actually be delivered at a given power level, and the higher the voltage, the lower the current can be, which reduces the refrigeration load. The largest voltages appear across the terminals of the coil, and also from coil to ground. The insulation in these areas must be designed to withstand high voltage stresses for many cycles. The voltages between layers of the coil, and between turns, are generally much less because the terminal voltage gets divided essentially evenly over each layer and turn.

The cryogenic system maintains the operating environment for the superconducting coil, which is at the normal boiling point of liquid helium, or approximately 4.5 K. The components are the helium refrigerator and the cryostat. A key part of the cryostat that affects the economics of  $\mu$ SMES are the leads that connect the cold magnet to the warm power conditioning system.

The complexity, size, and cost of the refrigerator depends on the size of the heat load to the refrigerated space. Larger heat loads, on the order of tens of watts, require a refrigerator based on the Collins cycle. These are commercially available and consist of a compressor and a cold box. The compressor produces a high-pressure stream of gas, which is passed through the cold box. Inside the cold box are counterflow heat exchangers, which cool the high-pressure gas with cold, low-pressure gas returning from the cold space. At several temperature stages, part of the gas stream is diverted through an expansion engine to produce additional cold cooling gas. Only a portion of the gas that is compressed actually is used to refrigerate the magnet. One of the problems with helium refrigerators is that any other contaminants in the gas stream condense out on the heat exchanger surfaces. This includes the oil that must be used to lubricate the compressor. One of the reasons helium refrigerators are so expensive and hard to maintain is that the gas cleaning systems require periodic cleaning.

At smaller heat loads, newer cryocooler technology is preferred. These are small refrigerators based on the Gifford-McMahon cycle, which employs a regenerative heat exchanger to cool the high-pressure gas stream. These devices are also subject to degradation resulting from fouled heat transfer surfaces, but they have the advantage of not requiring cold expansion engines that can wear out, and the regenerator tends to be self-cleaning to some degree. Cryocoolers still require periodic maintenance, but the system can be designed to accommodate a changeout of the cold head without having to shut down the magnet. The disadvantage of cryocoolers is that they are not as efficient as the Collins machines, which is why they are not used for large heat loads.

The cryostat provides a vacuum enclosure for the helium vessel and supports the vessel with low conductivity supports. A thermal shield is placed between the vacuum vessel and helium vessel to intercept heat radiated from the warm surface. The thermal shield is cooled either with liquid nitrogen or

by conduction to a cryocooler. Multilayer insulation is used to further reduce the heat leak. As long as the vacuum is maintained at a pressure of  $10^{-2}$  Pa or lower, the heat leak into the helium vessel can be made very low, with proper design practice.

The current leads to the coil will always be the largest single source of heat into the magnet. The connections, or power leads, cannot be made from materials with low thermal conductivity because they would then have a high electrical resistance. More electrical resistance means more resistive heating, which then ends up in the cold magnet system. The fundamental tradeoff for power leads is the balancing of the heat conducted down the lead against the heat generated by resistance within the lead. Whatever one does to make one component smaller increases the other by roughly the same amount. All is not lost, however; there are three things that can be done to minimize the heat input.

- The first is simply to reduce the current the leads have to carry. The required refrigeration power is directly proportional to the amount of current in the leads. A good magnet design, then, will use the least amount of current possible. The drawbacks are threefold. First, less current in the conductor means more turns of conductor are required to provide the same stored energy. Depending on how the conductor is priced, this can mean additional capital cost for the magnet system. (Note that the total volume of conductor stays about the same, so if the cost of conductor is dominated by bulk material cost, the cost is about the same regardless of current; if the conductor cost is primarily the result of processing, longer lengths of smaller wire will cost more.) The other two drawbacks are risk related. More turns of conductor make it harder to protect the magnet from overheating in the event of a quench. Also, as the current is decreased, it requires more voltage across the coil terminals to extract the same amount of power. As the voltage increases, so does the risk of damaging the coil insulation as a result of voltage breakdown.
- The second way to reduce the heat leak is to employ a high-temperature superconductor for a portion of the leads. The ceramic materials are an exception to the general rule that thermal conductivity goes up as electrical resistivity goes down. Although still a developing technology, high-temperature superconductor (HTS) leads have been built and have demonstrated a fourfold reduction in heat leak over conventional leads. They are complex and lack design maturity at the present time but offer a promising path to reduced heat leak.
- The final way to reduce the heat load caused by the power leads is to short the coil leads with what is called a persistent switch. This is a length of superconducting wire stabilized with copper-nickel alloy instead of pure copper, so it is more resistive when normal. The switch is opened by driving it normal. Allowing it to cool to helium temperature makes it superconducting. When superconducting, the current circulates entirely within the helium vessel. The leads only carry current during a current pulse when the energy is extracted from the  $\mu$ SMES device. As a consequence, the leads can be rated for transient operation only, resulting in a much smaller cross-section and a reduced heat leak.



### What Drives the Design?

The design of a  $\mu$ SMES storage system is fundamentally driven by the need to minimize the cost necessary to provide a specified amount of power for a specified span of time. The cost includes the cost of materials (primarily superconductor), the cost of labor to manufacture the magnet system, and the life-cycle operating cost of the system (the cost of electricity to run the refrigerator dominates). In this section, the design issues and tradeoffs are discussed, and a generic algorithm for producing an optimized design is outlined.

The ultimate design drivers for a  $\mu$ SMES device are the power  $P$  it is required to deliver and the time duration  $t$ ; the power is needed. It is then basic to SMES that the energy stored in the magnet system be at least equal to the product of the required power and the time duration. Other considerations force the total stored energy to exceed this amount.

There are constraints to the design as well. The fact that the voltage must rise as energy is extracted from the magnet has been mentioned. The maximum allowable voltage is one constraint on how much energy can be extracted.

Another constraint to the efficiency of the design is the fact that heat is produced in the windings when the coil is discharged. This is the result of what are commonly referred to as ac losses. They arise when a magnetic field changes with time in a conducting medium. If the ac losses heat the conductor to the point that superconductivity is lost, some of the stored energy is lost to heat dissipated in the windings and cannot be extracted as electrical power.

Another constraint is the stress level in the coil pack. The factors that make a coil efficient in terms of cost are the current density in the conductor, the field that the coil produces, and the size of the coil. Making any of these three larger to improve cost efficiency also increases the stress level in the conductor. Thus there is an upper limit to the economics resulting from the strength of the materials.

The behavior of superconducting material in a magnetic field also places a limit on the performance of  $\mu$ SMES. There is a limit to the amount of current a conductor can carry at any given field. This is called the critical current. The critical current density in the niobium-titanium superconducting alloy is limited to about  $3 \times 10^9$  A/m<sup>2</sup> at 5 T and drops to zero at about 0.5 T.

Finally, although one would like to make the current in the magnet as large as possible so that the discharge voltage is small, the heat leak associated with the current leads gets larger as the current increases. The economics of refrigeration places an upper limit on the value of current.

### What Are the Design Tradeoffs?

The constraints on the design lead to tradeoffs between various cost and risk elements. Quantification of these tradeoffs is the first step toward an algorithm that can be used to produce an optimized design. The ultimate goal of a  $\mu$ SMES design is to provide the necessary power for the necessary time at minimum cost. The cost will include materials, fabrication labor, and life cycle support and maintenance costs. The primary material cost will be cost of the superconductor. The main labor cost will be to wind the coils and assemble the structure and vessels. The largest life cycle cost will be the cost of electricity to run the helium refrigerator and cooling for the power electronics.

**Excess Storage (Cost) vs. Current.** There is a tradeoff between the cost of excess storage in the coil and the cost of refrigeration to remove the heat load resulting from the current leads. The system cost can be represented in a simple fashion by the expression

$$C = a \cdot U_0 + b \cdot I \quad (1)$$

The coefficient  $a$  represents the material and labor cost per unit megajoule of stored energy to build the unit, and  $b$  represents the cost of electricity to refrigerate the heat load due to the current leads.

As energy is extracted from the magnet, the voltage must always rise to maintain constant power as the current in the magnet drops. At some point before all the energy is extracted, the voltage reaches the maximum allowable value  $V_m$  and the discharge must stop. If  $P$  is the power extracted from the unit,  $V$  is the terminal voltage, and  $I$  is the current in the coil, then

$$P = IV \quad (2)$$

The energy  $U$  stored in the magnet at any time is

$$U = \frac{1}{2} LI^2 \quad (3)$$

where  $L$  is the inductance.

During a discharge, the energy stored in the magnet is reduced from the original amount  $U_0$  by the product of power and time  $Pt$ . Thus

$$U = U_0 - Pt \quad (4)$$

Note that for a discharge, the sign of  $P$  is negative (a negative voltage must be applied to the coil to extract energy). If  $t_1$  is the duration the power is needed, then the ratio of the energy extracted to the energy initially stored in the SMES unit can be expressed as

$$-\frac{Pt_1}{U_0} = 1 - \left[ \frac{P}{V_m I_0} \right]^2 \quad (5)$$

Equation (5) was developed by noting that at any time the ratio of  $U$  to  $U_0$  is equal to the square of the ratio of  $I$  to  $I_0$  and that at the end of the discharge, when the current is a minimum, the voltage is at its maximum value (in a negative sense). Note that, for a given allowable maximum voltage, power level, and time duration, as the initial current increases, the required initial stored energy decreases. Depending on the relative values of the cost coefficients  $a$  and  $b$ , there will be a minimum in the total cost that should be sought.

**Excess Storage vs. Power (ac Losses).** The discharge can also be limited by ac losses. If the losses cause the magnet to quench before the end of the discharge time  $t$ , the energy remaining in the magnet is lost to heat deposited in the windings. The factors that determine this limit can be determined as follows.

The ac loss heating  $Q$  is proportional to the square of the field decay rate  $B$ . For the purposes of this article, this can be written as

$$Q = \alpha B^2 \quad (6)$$

Now, the field is proportional to the current. If  $B_p$  is the operating field at the operating current  $I_o$ , then it can be shown that the total heat  $Q$  that is deposited in the conductor during a discharge is equal to

$$Q = \sigma \int_0^{t_q} B^2 dt = \sigma \frac{B_p^2}{4} \frac{P}{I_o} \ln \left[ \frac{I_o}{I_o - P t_q} \right] \quad (7)$$

At some value of  $Q$ , the magnet will quench. This equation defines the parameters that can reduce  $Q$ . Operating at a lower field will reduce  $Q$  but will mean the magnet will be larger (stored energy is proportional to the square of the field integrated over all space). A larger magnet in turn will most likely mean higher costs. Increasing the stored energy also reduces  $Q$  with more cost. The tradeoff here is to determine the minimum size and stored energy which allows a total energy  $P t_q$  to be extracted before the magnet quenches (if it does at all).

**How is the Design Optimized?** Given all these issues related to cost and performance, how does one go about finding the optimum design for a  $\mu$ SMES unit? The ideal approach would be to set down the end requirements  $P$  and  $t_q$ , derive the required initial stored energy  $U_o$ , apply all the constraints, and end with the geometry of the coil defined. Unfortunately, the relationships between field, current density, inductance, and coil geometry do not lend themselves to closed form solutions, which can be manipulated to provide such an approach. Instead, an iterative process must be followed, whereby a trial geometry is selected first and then the resulting  $\mu$ SMES system is checked to see that it satisfies all requirements and constraints. If not, the process is iterated until it does. Then cost is determined as a function of the geometry, and the geometry is varied again until a minimum cost is attained. It is worthwhile to outline in a little more detail the steps involved because it helps to illuminate a little better the design issues that affect the cost/performance of  $\mu$ SMES.

This algorithm finds the minimum wire cost per unit stored energy for a  $\mu$ SMES device. The variables that are free to adjust are the descriptors of the coil geometry, the length, the coil pack thickness, and the radius of the coil. It is common to use the nomenclature shown in Fig. 9, where  $a_i$  is the inner radius of the coil pack,  $a_o$  is the ratio of outer to inner radius, and  $\beta$  is the ratio of coil pack length to diameter. The optimized design will satisfy the constraints that

- the hoop stress in the conductor will be equal to or less than some allowable value  $\sigma_w$ .

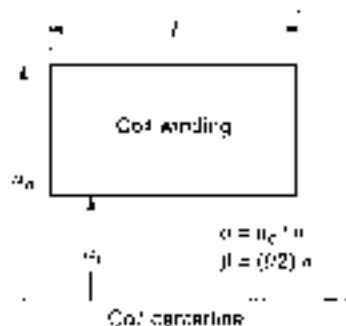


Figure 9. Identification of nomenclature used to describe magnet geometry.

- the absolute value of voltage at the end of a discharge of duration  $t$  and power level  $P$  will be equal to or less than an allowable value  $V_m$ .
- the critical current density in the superconductor will be equal to or less than the critical current density  $J_c$  at the point where the field is highest in the coil pack, and
- the coil will not quench as a result of ac losses before the energy extraction is complete.

The first step is to pick trial values for  $a$  and  $\beta$ . It is assumed that a solenoid design is being pursued. In the case of a toroid, there will be some relationship between the "length" of the coil and the major radius of the torus. Therefore, the major radius becomes a free parameter instead of  $\beta$ .

The next step is to determine the resulting magnetic field, with an assumed value of  $J_c a_i$ . It doesn't matter what the value is at this stage, just something convenient. One of the reasons for computing the field is to compute the stress level in the conductor subsequently. Exactly what detail is needed for the field distribution is a function of the detail needed for the stress calculation. In every case, however, the field at a particular point can be expressed as

$$B = \alpha_i J_c a_i f(\alpha, \beta) \quad (8)$$

The value of  $f$  depends on where it is evaluated, but the point is that  $B$  anywhere is proportional to the overall coil pack current density  $J_c$  and the coil radius  $a_i$ .

The stress is proportional to the product of radius, field, and current density. It can always be written as

$$\sigma = \alpha_i J_c a_i^2 S(\alpha, \beta) \quad (9)$$

The function  $S$  will depend on the level of complexity in the stress analysis. For thin solenoids with an approximately linear field profile through the winding, Wilson provides a closed form expression (5). Given a stress allowable  $\sigma_w$ , the product of current density and radius is then determined:

$$J_c a_i = \sqrt{\frac{\sigma_w}{8 \alpha_i \beta}} \quad (10)$$

Now the field at any point is known. In particular, the peak field in the coil pack is now determined, using Eq. (8).

The next step is to find the value of current density that minimizes the amount of superconductor needed. This value is the critical current density of the superconductor at the peak field. In any coil winding, there is space needed for material other than superconductor, and prudent design calls for some temperature margin for the conductor. If  $\phi$  is the ratio of non-superconductor to superconductor in the wire,  $\lambda$  is the fraction of the coil pack actually occupied by wire, and  $\Delta T_m$  is the temperature margin of the superconductor, then

$$J_c = \frac{\lambda}{1 - \phi} J_c(B_p, T_o + \Delta T_m) \quad (11)$$

Once  $J_c$  is known,  $a_i$  is known.

Given a complete description of the geometry and the current density, the stored energy can be found. Usually the value of  $(L/P)^2$  is determined, either from tables, or from a computer

oxide. In general, it will be proportional to  $a$ , and a function of the shape of the coil; thus,

$$(L_c/a^2) = a_c G(\alpha, B) \quad (12)$$

Once  $(L_c/a^2)$  is known, the stored energy is found from

$$U = \frac{1}{2} \frac{L_c}{a^2} (I_c A)^2 = \frac{1}{2} a_c G(\alpha, B) (I_c a_c 2a^2(\alpha - 1)B)^2 \quad (13)$$

Note that  $U$  is proportional to the fifth power of the coil radius.

At this point, it must be determined whether (a) the energy stored is sufficient to provide the required power for the required time and (b) the ac losses limit the energy extraction or not. If not, a new trial geometry must be selected and the process begun again. If both requirements are met, the algorithm proceeds. The next step is to determine the current, from Eq. (5). Then the cost of wire is determined. Usually, conductor can be estimated at some fixed rate per kiloamp-meter. Thus it is straightforward to scale cost data from a single point.

Once the cost as a function of stored energy is established, it is straightforward, at least in principle, to apply some optimization routine to the algorithm to find the optimum geometry that minimizes cost. The advantage of this algorithm is that, to the extent possible, it automatically finds the conditions of maximum stress and current density for a given field to minimize the amount of conductor needed. The current is set to the minimum value allowed by the voltage limit and the available stored energy (for a given configuration) to minimize the heat load.

## SUMMARY

We have seen how the economics of  $\mu$ SMES are affected by the power conditioning system, the storage system, and the parasitic power required to maintain operating conditions. There are certain power quality applications that benefit most from the advantages of  $\mu$ SMES. These are high-power situations with many repetitive cycles lasting a few seconds or less. The primary design driver for the magnet system is producing the maximum amount of stored energy for minimum cost, subject to constraints that available materials impose. A systematic algorithm to produce a cost-optimized design was developed. It maximizes the efficiency of the superconductor, while satisfying the constraints of stress, ac loss heating, maximum voltage, and allowable current density in the superconductor.

As the market for high-quality power develops and the relative benefits of  $\mu$ SMES as an option become better understood, it is expected that magnetic energy storage will make a major impact on the economics of power quality sensitive processes.

## BIBLIOGRAPHY

1. C. DeWinkel and J. D. Lammere, "Storing power for critical loads," *IEEE Spectrum*, 30(6): 38-42, 1993.
2. R. C. Dugan, M. F. McGranaghan, and H. W. Pasty, *Electrical Power Systems Quality*. New York: McGraw-Hill, 1996.
3. T. E. Greebe, *An Assessment of Distribution System Power Quality*, EPRI Projects RP 3058-1 and 3797-03, April 1985.
4. Draft Standard, *Electric Power System Compatibility with Electronic Process Equipment*, IEEE P1346 Working Group.

† M. N. Wilson, *Superconducting Magnets*, Oxford, UK: Clarendon Press, 1988, pp. 41-47.

Scott D. Beck  
State University of New York at  
Stony Brook  
John C. Zeigler  
Houston Advanced Research Center

## SQUIDS

Since the development of the Superconducting QUantum Interference Device (SQUID) in the late 1960s and its commercial introduction in 1970, SQUID-based instruments have proved to be the most sensitive measurement devices not only for magnetization measurements but also for a number of other electrical measurements. Their device noise (well below 1 mK), frequency response to dc, and low drift permit electromagnetic measurements at levels far below those of conventional techniques.

SQUID instruments consist of a SQUID amplifier or sensor and a detection circuit that transforms the signal of interest into a magnetic flux that is detected by the SQUID sensor. Associated control electronics transform this signal into a room temperature voltage that is available for additional signal processing if needed. The SQUID amplifier and the detection coils are superconducting devices. Thus some type of refrigerant (e.g., liquid helium or nitrogen) or refrigeration device is needed to maintain the SQUID (and detection coil) in the superconducting state. Additional signal-conditioning electronics may be needed to improve the signal-to-noise ratio.

## THE JOSEPHSON EFFECT

SQUIDS combine two phenomena: flux quantization where the flux  $\phi = \oint \mathbf{A}$  penetrating a superconducting loop is quantized in steps of  $\phi_0 = h/2e = 2.068 \cdot 10^{-15}$  Wb, and the Josephson effect (1): electrons tunneling from one superconducting region to another separated by a resistive barrier. Currents smaller than a critical current  $I_c$  can penetrate the barrier (usually called a weak link) with no voltage drop (Fig. 1).

This tunnel circuit is driven by a constant current RF oscillator that is weakly coupled to the SQUID loop. The detected RF output is found to be the periodic function (Fig. 3).

A typical weak link might have a critical current of 10  $\mu$ A. If the loop has a diameter of 2 mm, this is equivalent to several flux quanta. In a superconductor loop interrupted by a weak

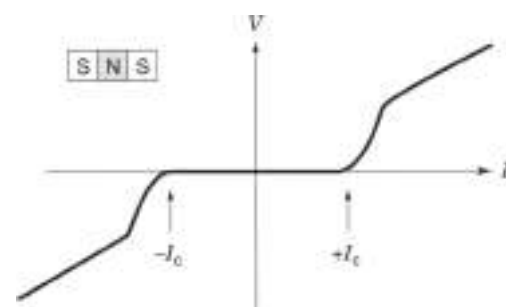


Figure 1. IV curve of a Josephson tunnel junction.

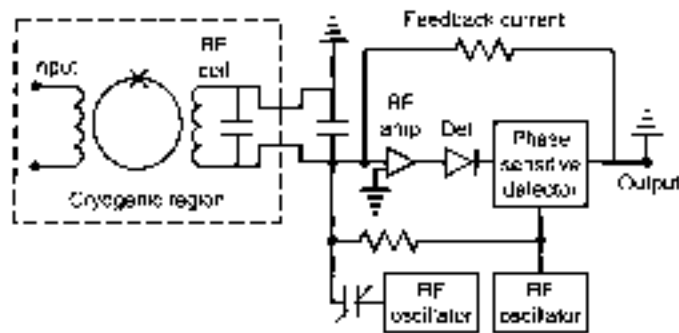


Figure 2. Block diagram of SQUID input and electronics for locked-loop operation.

link Josephson junction, magnetic flux threading through a superconducting loop sets up a current in the loop. As long as the current is below the critical current, the complete loop behaves as if it were superconducting. Any changes in the magnetic flux threading through the loop induce a shielding current that generates a small magnetic field to oppose the change in magnetic flux. The weak link can be a region in which the current flowing is greater than the current needed to drive the superconductor normal  $I_c$ . Details on the Josephson effect and the theory of SQUIDs can be found in References 2, 3, and 4.

## SQUID OPERATION

SQUIDs are operated as either RF (radio frequency) or dc SQUIDs. The prefix RF or dc refers to whether the Josephson junction(s) is biased with an alternating current (RF) or a dc current.

### The RF SQUID

For the RF SQUID (5), flux is normally (inductively) coupled into a SQUID loop containing a single Josephson junction via an input coil (which connects the SQUID to the experiment) and an RF coil that is part of a high- $Q$  resonant circuit to read out current changes due to induced flux in the SQUID loop (Fig. 2).

One way to measure the change in input coil current is to simply count the number of periods it produces in the detected RF output. A more commonly used mode of operation is a feedback scheme (Fig. 2), which locks in on either a peak or a valley in the triangle pattern output from the RF peak detector. A

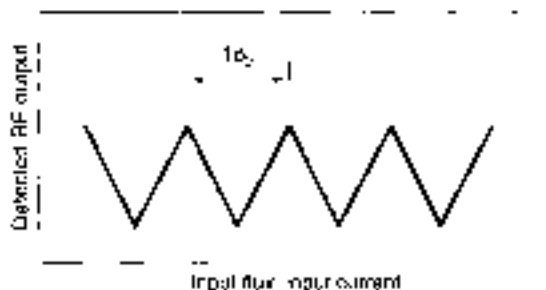


Figure 3. Triangle pattern showing detected output voltage as a function of flux in the SQUID.

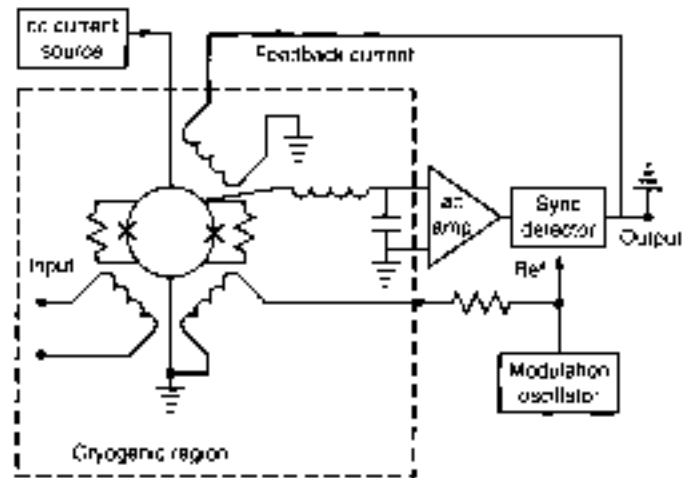


Figure 4. Block diagram of a typical dc SQUID.

feedback flux is applied to the SQUID through the RF coil that just cancels the change in flux from the input coil.

### The dc SQUID

The dc SQUID (Fig. 4) differs from the RF SQUID in the manner of biasing the Josephson junction (dc rather than ac) and the number of junctions (two rather than one).

The dc SQUID is biased with a dc current approximately equal to twice  $I_c$  and develops a dc voltage across the junctions. A change in the flux penetrating the SQUID loop enhances the current through one Josephson junction and reduces the current through the other, driving one junction normally and the other superconducting. This asymmetry, which is periodic in  $\phi_0$ , is used to provide a feedback current that nulls the flux penetrating the SQUID loop. Although total flux within the SQUID loop is in multiples of  $\phi_0$ , by measuring the voltage drop across the feedback resistor, resolutions of external flux changes at the  $10^{-6}$  level can be achieved. The linearity of flux-locked loop SQUID systems are typically better than 1 ppm. Like the RF SQUID, this feedback current (presented as a voltage at the output) is a direct measure of changes in flux applied to the SQUID.

**Control Electronics.** The system output voltage is the voltage drop across the feedback resistor in a negative feedback loop controlled by the SQUID electronics. The feedback signal is generated in response to changes in the output signal of the SQUID sensor. The output of the SQUID sensor is periodic in the field coupled into the SQUID loop. Negative feedback (similar to a phase-locked loop technique) is used to maintain the system operating point at a particular (and arbitrary) flux quantum. When operated in this mode, the system is in a flux-locked loop.

One important factor of SQUID design is such that the feedback electronics be able to follow changes in the shielding currents. If the shielding current changes so fast that the flux in the SQUID loop changes by more than  $\phi_0/2$ , it is possible that the feedback electronics will lag behind the rapidly changing flux. When the electronics finally "catch up," they can lock on an operating point (Fig. 3) different from the original. In this case, the SQUID has "lost lock" because the SQUID has exceeded

the maximum slew rate of the electronics. This places an upper limit on the bandwidth of the system. The typical bandwidth of commercially available SQUID systems is dc to 50+ kHz. Custom electronics have been built extending bandwidths above 5 MHz. Typical slew rates for SQUIDs are in the range of  $10^5$ – $10^6$   $\phi_0/s$ .

Even though one may not need or want to observe rapidly changing signals, situations may arise when ambient noise (e.g., 60 Hz) may determine the slew rate requirements of the system. To recover a signal from such interference, the system must be able to track all signals present at the input, including the noise. When system response is sped up to handle very fast signals, sensitivity to RF interference and spurious transients is also increased. Because the ability to remain locked while subjected to strong electrical transients is greatest when the maximum slew rate is limited (slow), whereas ability to track rapidly varying signals is greatest when the maximum slew rate is greatest (fast), it is desirable to be able to match the maximum slew-rate capability to the measuring situation. As a matter of convenience, many commercial SQUID systems offer user-selectable slew rates along with high-pass and low-pass filters for noise reduction.

### Sensitivity

Because of the varying input impedances of SQUID sensors, the sensitivity of SQUID devices is best discussed in terms of the energy sensitivity.

$$E_N = L_i I_N^2 = \frac{\phi_N^2}{I_N} \quad (1)$$

where  $L_i$  is the input inductance of the device,  $I_N$  is the current noise, and  $\phi_N$  is the flux sensitivity.  $E_N$  is often expressed in terms of Planck's constant  $h = 6.6 \times 10^{-34}$  J/Hz.

The major limiting factor in the noise of a RF SQUID is the bias frequency  $f_b$  used to excite the tank circuit and that RF SQUID noise is proportional to  $1/\sqrt{f_b}$  (6). As  $f_b$  increases, the complexity of the electronics also tends to increase.

The minimum noise energy for a dc SQUID is given by (7)

$$E_N = k_B T \sqrt{L_{junc} C} \quad (2)$$

where  $k_B$  is Boltzmann's constant,  $L_i$  is the inductance of the SQUID loop, and  $C$  is the capacitance of the junction. Substituting appropriate numbers indicates that the minimum noise energy  $E_N$  for a dc SQUID is on the order of  $h/2$ . Devices with sensitivities of  $\sim h$  have been constructed. These extremely low noise levels are achieved by limiting dynamic range and avoiding feedback. The need for practical (useful) devices requires that feedback be used and that the SQUID have a reasonable dynamic range. Commercially available RF SQUIDs have noise levels of  $10^{-22}$  J/Hz, commercial dc SQUIDs are typically  $\sim 10^{-21}$  J/Hz.

In addition to the frequency independent (white) component of system noise, there exists a low-frequency contribution that increases as the frequency decreases. The onset of this  $1/f$  noise can be dependent on the ambient magnetic field when the SQUID sensor is coiled. When coiled in the earth's magnetic field, the point at which the  $1/f$  noise equals the white (frequency independent) noise is typically  $\sim 1$  Hz. Coiling the SQUID sensor in low ambient magnetic fields (less than  $1 \mu T$ )

may improve the  $1/f$  performance by as much as an order of magnitude. A large contribution to this noise in some dc SQUIDs can arise from the presence of the dc current bias. By chopping the dc bias in combination with the conventional flux modulation techniques, it is possible to reduce this added  $1/f$  noise. This ac bias reversal approach (8) separates the original signal waveform from the noise associated with the dc bias and can reduce  $1/f$  noise at very low frequencies.

The major difference between RF and dc SQUIDs is that the dc SQUID offers lower noise. From a historical viewpoint, although the dc SQUID was the first type of SQUID magnetometer made, the early development was with RF SQUIDs because of the difficulty in fabricating two nearly identical Josephson junctions in a single device. With modern thin film fabrication techniques and improvements in control electronics design, the dc SQUID offers clear advantages over the RF SQUID for many applications.

### Limitations on SQUID Technology

It is important to bear in mind several fundamental limitations in designing SQUID-based measurement systems and data reduction algorithms.

1. A fundamental limitation of SQUIDs is that they are sensitive to *relative* (field or current) changes only. This is a consequence of the fact that the output voltage of a SQUID is a periodic function (Fig. 3) of the flux penetrating the SQUID loop. The SQUID is "flux locked" on an arbitrary maximum (or minimum) on the  $V - \Psi$  curve, and the SQUID output is sensitive to flux changes relative to this lock point.
2. A second limitation exists on the system bandwidth. Although the SQUID has an intrinsic bandwidth of several gigahertz, when operated with standard flux-locked loop electronics using ac flux modulation, the maximum bandwidth is typically 50 kHz to 100 kHz. Another limitation is the presence of  $1/f$  noise. High Temperature Superconducting (HTS) SQUIDs (and early commercial Low Temperature Superconducting (LTS) dc SQUIDs) exhibit excess  $1/f$  noise due to critical current fluctuations of the Josephson junctions. This noise can be reduced by reversing the dc bias voltage (no bias). This limits the maximum bandwidths less than half the bias reversal frequency. If the bias reversal frequency is too high, noise can be induced due to voltage spikes in the transformer coupled preamplifier input circuit. Because of this, the maximum bandwidth of present day HTS SQUIDs is  $\sim 30$  kHz. If megahertz bandwidths are required, the ac bias is not used, however, there will be excess noise below 1 kHz.
3. Finally, SQUID magnetometers are vector magnetometers. For a pure magnetometer operating in the earth's magnetic field, a 180° rotation will sweep out a total field change of  $\sim 100 \mu T$ . If the magnetometer has a sensitivity of  $10 \text{ fT}/\sqrt{\text{Hz}}$ , tracking the total field change requires a dynamic range of  $100 \mu T/10 \text{ fT} = 200 \text{ dB}$ , well beyond the capabilities of current electronics. In addition, the rotational speed must not cause the current flowing through the SQUID sensor to exceed its slew rate limitations. An ideal gradiometer is insensitive to a uniform field and would not suffer this dynamic range limitation.

## INPUT CIRCUITS

Whether an RF or dc SQUID, a SQUID system can be considered as a black box that acts like a current- (or flux-) to-voltage amplifier with extremely high gain. In addition, it offers extremely low noise, high dynamic range (>140 dB), excellent linearity (>1:10<sup>7</sup>), and a wide bandwidth that can extend down to dc.

Today, SQUIDs are fabricated as planar devices. In this configuration, the superconducting loop, Josephson junctions and coils (input, feedback, and modulation) are patterned on the same device. Multilayer deposition techniques are used (primarily in LTS devices), and coils are normally in the form of a square washer. The planar configuration leads to quite small devices, occupying only a few cubic millimeters compared to 5 l cm<sup>3</sup> (1.2 cm diam. × 5 cm) for older toroidal RF SQUIDs (9). Another advantage of the planar device is that it is possible to have the detection coils as part of the SQUID sensor, eliminating the need for separate (three-dimensional) detection coils. Such an integrated sensor has the potential to reduce the complexity of multichannel systems significantly.

Although it is possible to couple magnetic flux directly into the SQUID loop, environmental noise considerations (see Fig. 9) make this difficult, if not impossible, in an unshielded environment. In addition, the area of a typical SQUID loop is small (<0.1 mm<sup>2</sup>), and its resulting sensitivity to external flux changes ( $\Delta\Phi = A \cdot \Delta B$ ) is also small. Although a larger loop diameter would increase the SQUID's sensitivity to external flux, it would also make it much more susceptible to environmental noise. For this reason, external flux is normally inductively coupled to the SQUID loop by a flux transformer.

Conceptually, the easiest input circuit to consider for detecting changes in magnetic fields is that of a SQUID sensor connected to a simple superconducting coil (Fig. 5).

Because the total flux in a superconducting loop is conserved, any change in external field (through the signal coil will induce a current in the flux transformer that must satisfy

$$\Delta\Phi = NA\Delta B = (L_{\text{ext}} + L_s)\Delta I \quad (13)$$

where  $\Delta B$  is the change in applied field;  $N$ ,  $A$ , and  $L_{\text{ext}}$  are the number of turns, area, and inductance of the detection coil;  $L_s$  is the inductance of the SQUID input coil, and  $\Delta I$  is the change in current in the superconducting circuit. If the lead inductance is not negligible, it must be added to  $L_{\text{ext}}$  and  $L_s$ .

To calculate the sensitivity and noise level of a simple detection coil system, the inductance of the detection coil must be known. The inductance of a flat, tightly wound, circular multi-

turn loop of superconducting wire is given by (10):

$$4 \times 10^{-7} N^2 \pi r \left[ \ln \left( \frac{2r}{\rho} \right) - 2 \right] \frac{\mu H}{\text{cm turn}^2} \quad (14)$$

where  $r$  is the radius of the detection coil and  $\rho$  is the radius of the (superconducting) wire. Knowing the coil inductance  $L_{\text{ext}}$ , we can rewrite Eq. (13) as

$$\Delta B = (L_{\text{ext}} + L_s)\Delta I/NA \quad (15)$$

Because the SQUID system has an output proportional to the input current, maximum sensitivity is obtained by using the input circuit that provides the maximum current into the SQUID and satisfies all other constraints of the experimental apparatus. For a pure magnetometer, the maximum sensitivity will occur when the impedance of the detection coil matches that of the SQUID sensor ( $L_{\text{ext}} = L_s$ ).

## Detection Coils

Several factors affect the design of the detection coils (11). These include the desired sensitivity of the system, the size and location of the magnetic field source, and the need to match the inductance of the detection coil to that of the SQUID. The ability to separate field patterns caused by sources at different locations and strengths requires a good signal-to-noise ratio. At the same time, one has to find the coil configuration that gives the best spatial resolution. Unfortunately, these two tasks are not independent. For example, increasing the signal coil diameter improves field sensitivity but sacrifices spatial resolution. In practice, system design is restricted by several constraints: the impedance and noise of the SQUID sensors, the size of the dewar, and the number of channels, along with the distribution and strength of noise sources.

It is extremely important for dc response that the detection coil(s) be superconducting. Resistance in the detection circuit has two effects: (1) attenuating the signal and (2) adding Nyquist noise. Resistive attenuation is important only below a frequency  $f_0$ , such that the resistive impedance is equal to the sum of the inductive impedances in the circuit (e.g.,  $f_0 = R/L_s$ , where  $L_s$  is the total inductive impedance of the circuit). Resistive noise is important only if it becomes comparable to other noise sources or the signal (>10<sup>-20</sup> J/Hz for biomagnetism, <10<sup>-26</sup> J/Hz for geophysics). For a SQUID with  $R_N = 10^{-25}$  J/Hz, the total resistance of the circuit, including any joints, must be less than  $1.4 \times 10^{-18}$  Ω (12). Thus it is very important that all solder joints, press-fits, or connections have as low a joint resistance as possible.

Figure 6 displays a variety of detection coils. The magnetometer (Fig. 6(a)) responds to the changes in the field penetrating the coil. More complicated coil configurations provide the advantage of discriminating against unwanted background fields from distant sources while retaining sensitivity to nearby sources.

Because of the present inability to make flexible wire or make true superconducting joints in HTS materials, three-dimensional HTS coil structures (e.g., Figs. 6(b,d-f)) are not possible. Present day HTS magnetometers are fabricated as planar devices and are available only as pure magnetometers (Fig. 6(a)) and planar gradiometers (Fig. 6(c)). As a result, commercially available HTS devices are currently in

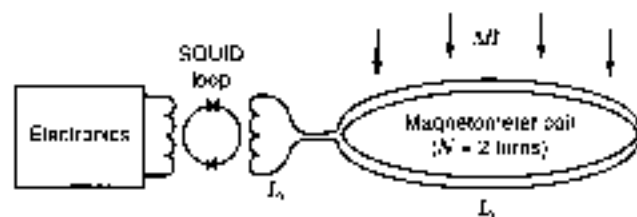
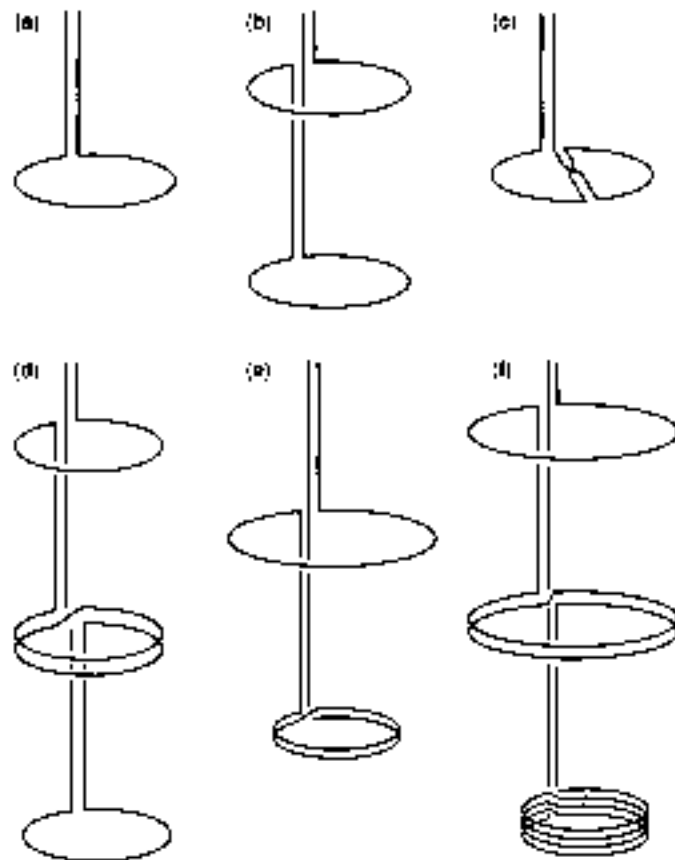


Figure 5. Schematic diagram of typical SQUID input circuit.



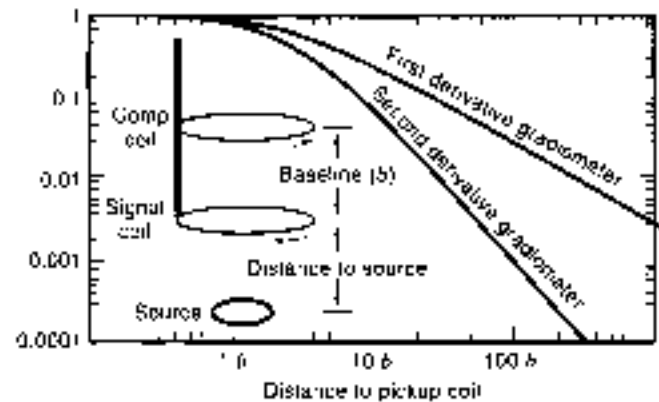
**Figure 6.** (a) magnetometer, (b) first derivative gradiometer, (c) planar gradiometer, (d) second derivative gradiometer, (e) first derivative asymmetric gradiometer, (f) second derivative asymmetric gradiometer. Courtesy of S. J. Williamson

the form of magnetic sensing rather than current sensing devices.

### Gradiometers

Magnetometers are extremely sensitive to the outside environment. This may be acceptable if one is measuring ambient fields. If what is to be measured is close to the detection coil and weak, outside interference may prevent measurements at SQUID sensitivities. If the measurement is of a magnetic source close to the detection coil, a gradiometer coil may be preferred. The field of a magnetic dipole is inversely proportional to the cube of the distance between the dipole and the sensor. It follows that the field from a distant source is relatively uniform in direction and magnitude at the sensor. If we connect in series two identical and exactly parallel loops wound in opposite senses, separated by a distance  $b$  (the baseline), we obtain a coil [Fig. 6(b)] that will reject uniform fields.

Because the response of a single coil to a magnetic dipole goes as  $1/r^3$ , an object that is much closer to one coil than the other will couple better to the closer coil than the more distant. Sources that are relatively distant will couple equally into both coils. For objects that are closer than  $0.3b$ , the gradiometer acts as a pure magnetometer, while rejecting more than 99% of the influence of objects more than  $300b$  distant (Fig. 7). In essence, the gradiometer acts as a compensated magnetometer. It is possible to use two gradiometers connected in series opposition [Fig. 6(d)] to further minimize the response of the system



**Figure 7.** Response of gradient coils relative to magnetometer response ( $r^3$  suppressed):

to distant sources. This can be extended to higher orders by connecting in series opposition two second-order gradiometers, etc. Doing so, however, reduces the sensitivity of the instrument to the signal of interest and may not significantly improve the signal-to-noise ratio.

Rejection of distant noise sources depends on having a precise match (or balance as it is sometimes referred to) between the number of area-turns in the coils. A symmetric gradiometer [Fig. 6(b)] requires that  $N_1A_1 = N_2A_2$ , where  $N$  is the number of turns and  $A$  is the area of the signal and compensation coils, respectively. An asymmetric design [Fig. 6(e,f)] has the advantage that the inductance ( $L_1$ ) of the signal coils is much greater than the compensation coils ( $L_2$ ); greater sensitivity is achieved than with a symmetric design. Another advantage is that the signal coil diameter is reduced, leading to potentially higher spatial resolution. The optimum conditions for the number of turns on an asymmetric signal coil is given by (18):

$$(L_1 + L_2 - L_3 + L_4) - N_2 \frac{\partial}{\partial N_1} (L_1 + L_2 + L_3 + L_4) = 0 \quad (18)$$

If the gradiometer is perfectly made (balanced), it will reject uniform fields. However, if one coil has a larger effective diameter than the other, the response will not be that of a perfect gradiometer, but that of a gradiometer in series with a magnetometer. Mathematically, the balance,  $\beta$ , can be defined as  $V_1 \times G = \beta \times H$ , where  $V_1$  is the system response,  $G$  is the coil's response to a gradient field, and  $H$  is the applied uniform field. Typically, coil forms used to wind gradiometers can be machined (grooved) to achieve balances  $\beta$  that range from 0.01 to 0.001. Planar devices, through photolithography, can achieve lower levels—a factor of 10 or better. Superconducting trim tabs placed within the detection coils can improve  $\beta$  to the parts per million level. High degrees of balance can allow a SQUID gradiometer to operate in relatively large (millitesla) ambient fields while maintaining sensitivities in the tens of femtoTesla.

For multichannel systems (such as are used in biomagnetism), it is not possible to use externally adjustable trim tabs—each tab tends to interfere with the others. The use of electronic balancing (14) can provide balance ratios at the parts per million level. In this situation, portions of (additional) magnetometer reference channel response are summed electronically with the gradiometers' input to balance out its effective

magnetometer response. The use of eight-element tensor arrays as reference channels can further improve external noise rejection. The major advantage of electronic balancing is significant improvement in immunity to low-frequency environment noise.

## REFRIGERATION

The superconducting nature of SQUIDs require them to operate well below their superconducting transition temperature (9.3 K for Nb and 93 K for  $YBa_2Cu_3O_{7-x}$ ). The thermal environment for the SQUID sensor and detection coil has typically been liquid helium or liquid nitrogen contained in a vacuum-insulated vessel known as a dewar (Fig. 8). The cryogen hold time depends on the boil-off rate (heat load) and the inner vessel volume.

The major heat load on dewars is the result of thermal conduction down the neck tube and a magnetometer probe along with black body radiation. The space between the inner and outer walls is evacuated to prevent thermal conduction between room temperature and the cryogen chamber. Within the vacuum space, a thermal shield (anchored to the neck tube) acts to reduce heat transfer by thermal (black body) radiation. The thermal shield either can be vapor cooled—using the enthalpy of the evaporating helium or nitrogen gas—or have the shield thermally connected to a liquid nitrogen reservoir. Dewars with removable sections use liquid-nitrogen-cooled shields.

If the experiment involves measurements interior to the dewar, then a metallic dewar is preferable. Metallic dewars offer significant shielding from environmental noise at frequencies above 10 Hz to 100 Hz. If the system is to measure magnetic fields exterior to the dewar, the dewar must be magnetically transparent, and metallic construction is not appropriate.

Dewars for external field measurements are normally constructed of nonmetallic, low-susceptibility materials to minimize their magnetic interactions with the SQUID sensors and detection coils. Materials used are typically glass-fiber epoxy composites such as G-10. In an effort to get the detection coils as close as possible to the object being measured, a "tailed" design is often used. This decreases the forces on the bottom of the dewar and allows the use of thinner end pieces (closer tail spacing). Dewars for biomagnetic measurements often have curved tails to get closer to the head.

The major advantage of high-temperature superconductivity is the simplified cryogenics and reduced spacing between cryogenic regions and room temperature. The thermal load (due to conduction and black body radiation) is less, and the heat capacity of what needs to be cooled is larger (implying smaller temperature variations for a given heat load). Because the latent heat/unit volume of liquid nitrogen is 60 times larger than liquid helium, hold times become months rather than days for an equivalently sized dewar.

## Closed Cycle Refrigeration

As an alternative to the use of liquid cryogens, closed cycle refrigeration would be desirable for several reasons. These include reduction of operating costs, use in remote locations, operation in nonvertical orientations, avoiding interruptions in cryogen deliveries, safety, and the convenience of not having to transfer every few days. Although one mechanically cooled SQUID system has been built (15), the inherent vibration and magnetic signature of present day closed cycle refrigerators prevent them from widespread use. The development of pulse tube refrigerators (16) offers promise for magnetometer operation without cryogens.

## Environmental Noise

The greatest obstacle to SQUID measurements is external noise sources. If the object being measured is within the cryostat (such as is typical in most laboratory experiments), metallic shielding can minimize external noise (e.g., act as a low-pass eddy current shield). Superconducting shields essentially eliminate all external field variations. This assumes that any electrical inputs to the experimental region have been appropriately filtered. Powerline or microprocessor clock frequencies can severely degrade performance.

When measuring external fields, the SQUID magnetometer must operate in an environment—the magnetic field of the earth—that can be ten orders of magnitude greater than its sensitivity (Fig. 9). The magnetic field at the surface of the earth is generated by a number of sources. There exists a background field of  $\sim 50 \mu\text{T}$  with a daily variation of  $10.1 \mu\text{T}$ . In addition, there is a contribution (below 1 Hz) from the interaction of the solar wind with the magnetosphere. The remaining contributions to external magnetic fields are primarily man-made. These can be caused by structural steel and other localized magnetic materials such as furniture and instruments that distort the earth's field and result in field gradients; moving vehicles that generate transient fields; electric motors; elevators; radio, television, and microwave transmitters; and the ever-present powerline electromagnetic field and its harmonics.

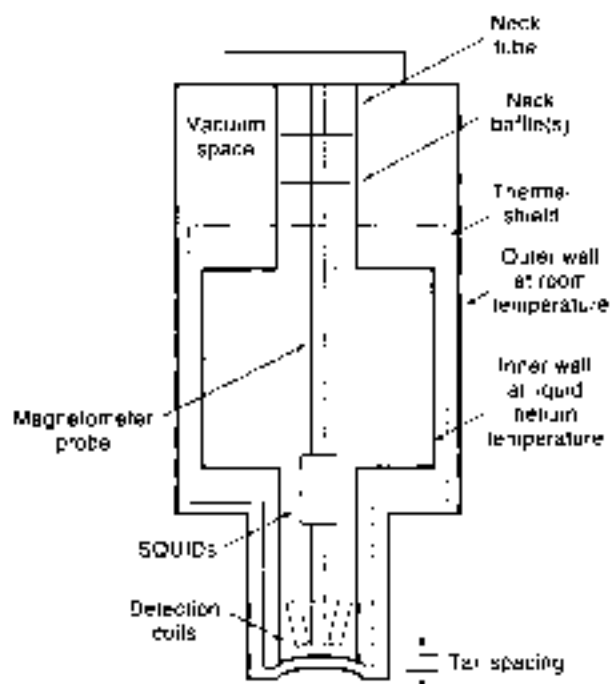


Figure 8. Typical design of a fiberglass dewar used for biomagnetic measurements.



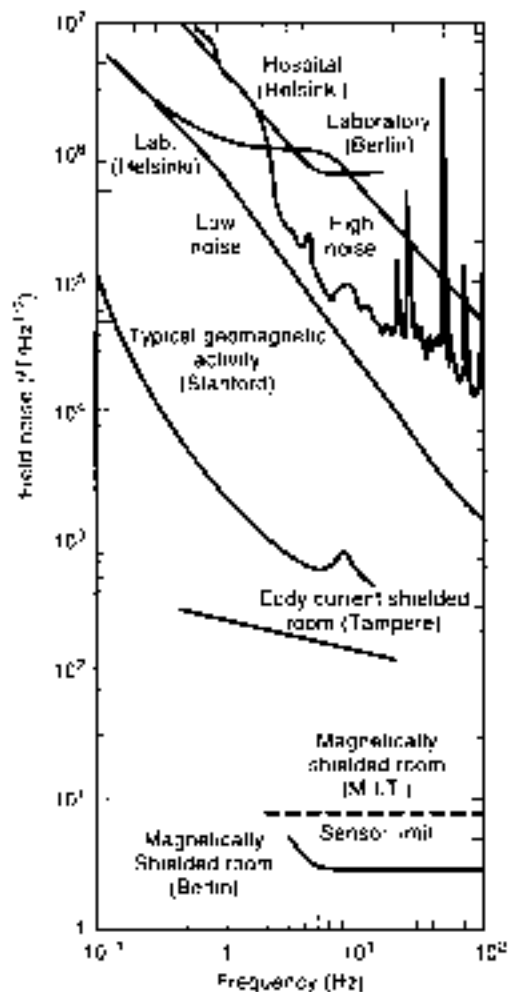


Figure 9. Rms field noise spectra in various environments as a function of frequency. Courtesy of S. J. Williamson.

### Noise Reduction

One method to attenuate external noise sources is with an eddy current shield that generates fields that act to cancel the externally applied fields within the conducting material. The shielding effect is determined by skin depth  $\lambda$ . For a sinusoidal varying wave

$$\lambda = \sqrt{\rho / \pi \mu_0 f} \quad (7)$$

where  $f$  is the frequency of the applied field,  $\rho$  is the electrical resistivity, and  $\mu_0$  is the magnetic permeability of free space. In situations where the wall thickness  $t \ll \lambda$ , external fields are attenuated by

$$\frac{H}{H_0} = \frac{1}{1 + (2\pi f L_0 R)^2} \quad (8)$$

where  $L_0$  is the inductance of the enclosure and  $R$  is the resistance along the path of current flow. Unfortunately, induced currents in the shield generate noise. For a cylindrical shape,

$$R_{\text{enc}} = \sqrt{\frac{64\pi k_B T l}{hd_0}} \quad (9)$$

where  $h$  is the length and  $d_0$  is the diameter of the can. The cut-off frequency is given by  $f_{\text{cut-off}} \approx \rho / 4\pi d_0^2$ . Because of noise considerations, eddy current shields that are to be placed near the detection coils should be made from relatively poor conductors such as BeCu.

### Shielded Rooms

Another approach is to use eddy current shielding to shield the entire measurement system. An eddy current room constructed with 2 cm high purity aluminum walls can achieve shielding >40 dB at 60 Hz with improved performance at higher frequencies. The equivalent field noise is less than 200 fT/√Hz at frequencies above 1 Hz.

The need for shielding at lower frequencies has led to the use of magnetically shielded rooms (MSR). In the situation where  $t \gg \lambda$ , the attenuation goes as  $(t/\lambda)^2$ . Using pure eddy current shielding would require wall thicknesses that could exceed 1 m or more (below 1 Hz). For a ferromagnetic material, the permeability of the material ( $\mu = \mu_0(1 + \chi)$ ) replaces  $\mu_0$  in Eq. (8). The shielding is due to the fact that flux prefers the path with the highest permeability. Because magnetically "soft" materials (e.g., mu-metal<sup>®</sup>) can have permeabilities that exceed  $10^5$ , the external magnetic flux is routed around the walls, avoiding the interior. The use of multiple shields can act to further shield the interior of a MSR. For the six-layer Berlin MSR (Fig. 9), shielding factors exceeded  $10^3$  at frequencies above 0.01 Hz with noise levels below 3 fT/√Hz.

### APPLICATIONS

A large number of applications (Fig. 10) configure the SQUID as a magnetometer. SQUIDs can also be configured to measure a wide variety of electromagnetic properties (Fig. 11).

The state of the art in materials processing limits the variety of superconducting input circuits that can be used with LTS SQUIDs. As already mentioned, there is no existing method for making superconducting connections to SQUIDs with HTS wire. As a result, commercially available HTS devices are currently in the form of magnetic sensing (Fig. 6(b)) rather than current sensing devices (Fig. 6(a,c-f)).

#### Laboratory Applications (4,8,17)

Table 1 shows typical capabilities of commercially available SQUID-based instruments. The number in the parenthesis refers to the corresponding Fig. 11.

**Current.** One common use of a SQUID is as an ammeter (Fig. 11(a)). The input can be connected to an experiment at liquid helium temperatures or to room temperature. If the signal is to be inductively coupled to a detection coil that is connected to the SQUID input, then the circuit must be superconducting if dc response is desired.

If the measurement is of a current that passes through the detection coil, a toroidal geometry for the detection coil has the advantage of extremely good coupling to the source while rejecting contributions due to external sources. Because the measurement is inductive, there is no loading of the current-generating elements.

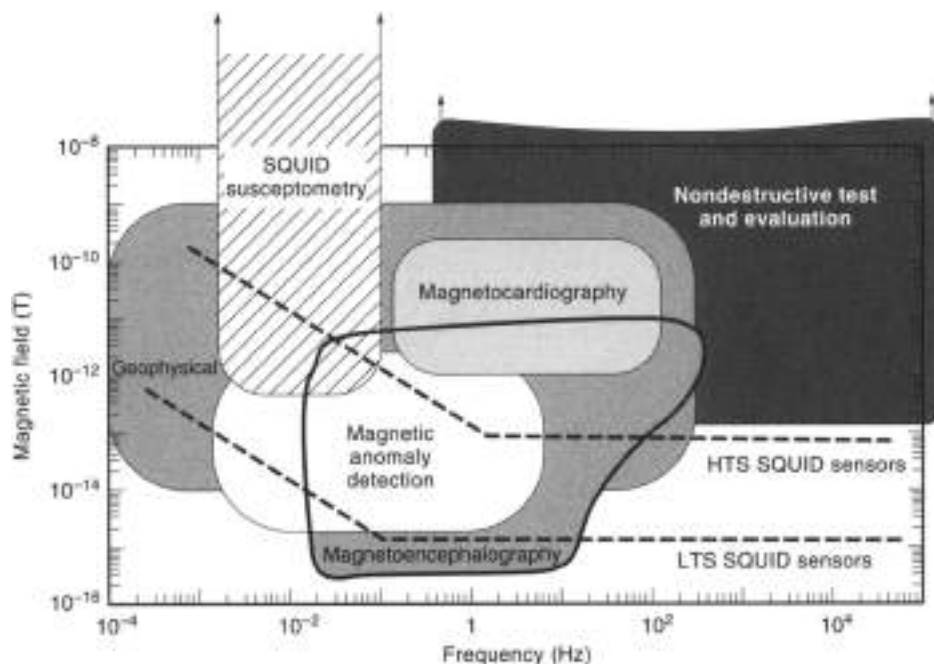
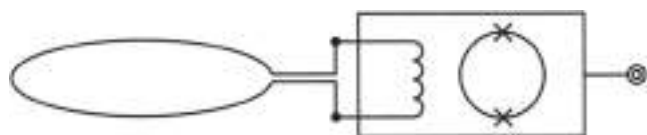


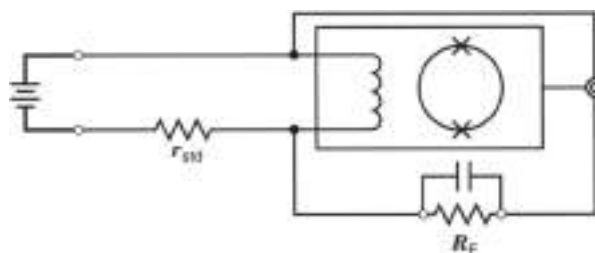
Figure 10. Field sensitivities and bandwidths typical of various applications. The lines indicate the sensitivity of commercially available SQUIDs.



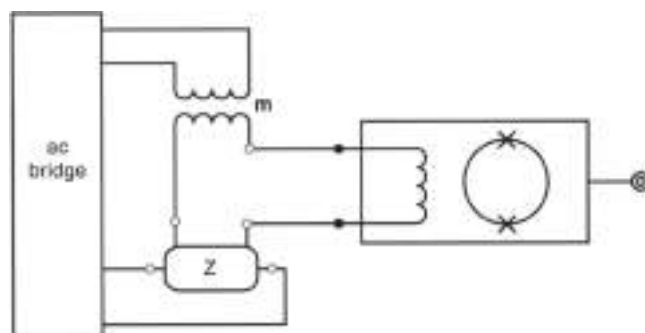
(a) ac and dc current



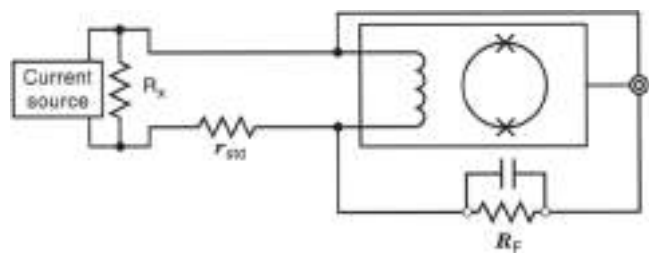
(b) Magnetic field



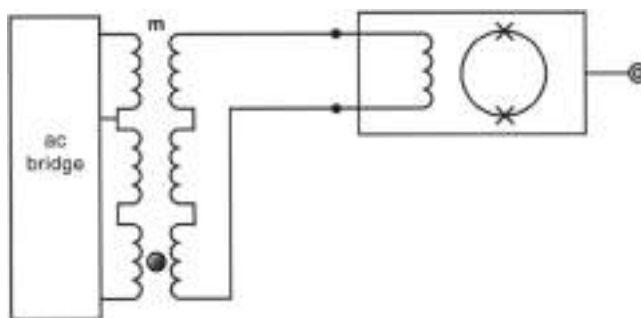
(c) dc voltage



(e) ac resistance/inductance bridge



(d) dc resistance



(f) ac Mutual Inductance (susceptibility bridge)

Figure 11. (a) ac and dc current, (b) magnetic field, (c) dc voltage, (d) dc resistance, (e) ac resistance/conductance bridge, (f) ac mutual inductance (susceptibility bridge).

Table 1. Typical Capabilities of SQUID-based Instruments

| Measurement                         | Sensitivity     |
|-------------------------------------|-----------------|
| Current [Fig. 11(a)]                | $10^{-6}$ A/√Hz |
| Magnetic fields [Fig. 11(b)]        | $10^{-6}$ T/√Hz |
| dc voltage [Fig. 11(c)]             | $10^{-6}$ V     |
| dc resistance [Fig. 11(d)]          | $10^{-6}$ Ω     |
| Mutual/self inductance [Fig. 11(e)] | $10^{-6}$ H     |
| Magnetic moment [Fig. 11(f)]        | $10^{-6}$ emu   |

**Voltage.** Typically the majority of applications use superconducting circuits. There are, however, a number of applications where resistive circuits are used. One example is the detection of extremely small voltages or resistances (Fig. 12).

When a voltage  $V_1$  is applied across the input terminals, a current is generated in the SQUID input coil. In this situation, the feedback current ( $I_F$ ) that would normally be applied to the SQUID loop is fed back via  $R_F$  through  $r_s$  until the voltage drop across  $r_s$  is equal to  $V_1$  and there is no net current through the SQUID.  $V_2$  measures the voltage drop across  $R_F$  and  $r_s$  with  $V_1 = V_2(r_s/R_F + r_s)$ . The voltage gain of the system is determined by the ratio of  $R_F/r_s$ . Typical values for  $R_F$  and  $r_s$  are 3 and 80  $\mu\Omega$ , respectively, giving a voltage gain of  $10^3$ . The standard resistor  $r_s$  is typically at 4.2 K. The voltage source, however, may be at a completely different temperature.

The input noise of a SQUID picovoltmeter ( $\sim 10^{-14}$  V) is a function of the source resistance and temperature ( $K \propto T_1$ ), the inherent voltage noise (due to  $r_s$ ), and the inherent current noise of the SQUID. Measurement of the Johnson noise in a resistor ( $V^2 = 4 k_B T R \Delta f$  where  $\Delta f$  is the bandwidth of the measurement) can be used to determine absolute temperature. Commercially available LTS SQUIDS have equivalent device temperatures  $< 1$   $\mu$ K and are suitable for noise thermometry.

With the addition of an appropriate current source (Fig. 11d), it is possible to measure resistance. Resolutions of  $10^{-12}$   $\Omega$  can be achieved for  $R_s < 10^{-2}$   $\Omega$ . Other applications of picovoltmeters include measurements of thermopower, thermal electromotive forces (teofs) (thermocouples), and infrared bolometers.

**Ac Susceptibility (4).** The SQUID can also be used as the null detector in an ac bridge circuit (Fig. 13) to measure both re-

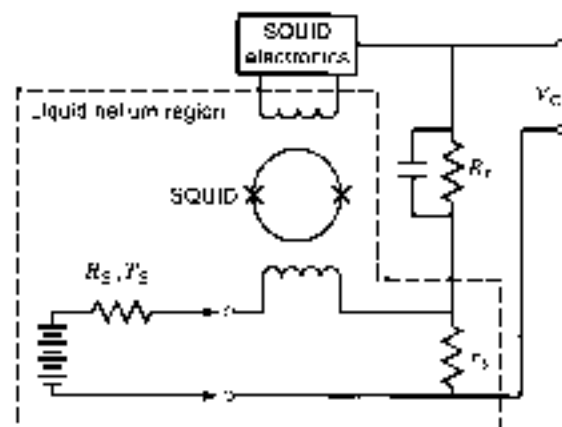


Figure 12. Block diagram of SQUID picovolt measuring system. Negative feedback is applied to the cryogenic input circuit through the voltage divider formed by  $R_F$  and  $r_s$ .

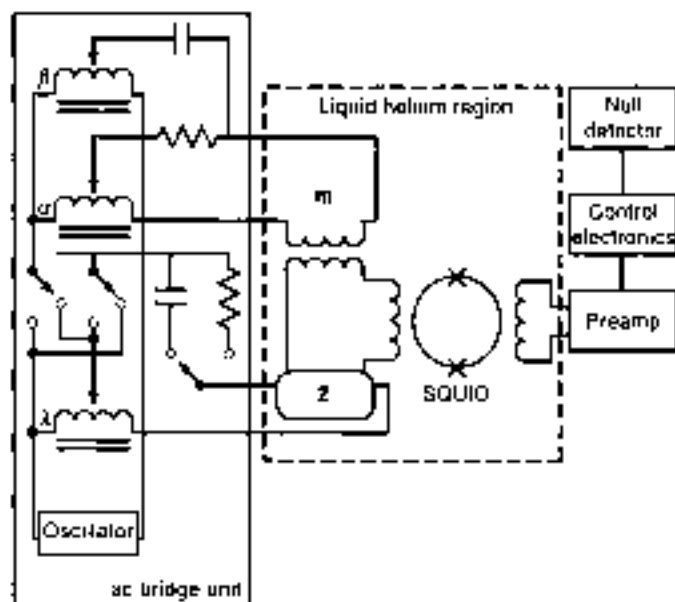


Figure 13. Block diagram of ac bridge

sistive and reactive components of a complex impedance. The unknown impedance  $Z$  is excited by a current generated by an oscillator voltage, which is attenuated by a precision ratio transformer  $\lambda$ . The difference between the voltage developed across the unknown impedance  $Z$  and that developed in the secondary of a nulling mutual inductor  $m$  is applied to the input of the SQUID circuit. The primary current in  $m$  is proportional to the oscillator voltage and defined by the setting of the ratio transformer  $\alpha$ . An additional reactive current is supplied by a second ratio transformer  $\beta$ , which causes the primary current to be passed through a capacitor rather than a resistor, thus generating a 90° phase shift in the voltage applied to  $m$ . The amplified off-balance signal, which appears at the output of the SQUID control electronics, can be displayed by means of a lock-in amplifier tuned to the oscillator frequency.

As mentioned earlier, the sensitivity is limited inherently by Johnson noise in the resistive components of the unknown (including the potential connections) and by the device noise of the SQUID sensor. Assuming  $I_C = 1$   $\mu$ A/√Hz, such a system is capable of measuring  $R$  between  $10^{-10}$  to 0.5  $\Omega$  and  $L$  (self and mutual) between  $10^{-2}$  and  $10^{-8}$  H. Using a current comparator as the bridge excitation  $V_1$ , a 0.1 ppm resistance bridge can be constructed.

**SQUID Magnetometer/Susceptometers (16).** Instead of using a secondary ac excitation coil (Fig. 11(f)), a dc field can be used to magnetize samples. Typically the field is fixed, and the sample is moved into the detection coil's region of sensitivity. The change in detected magnetization is directly proportional to the magnetic moment of the sample. Because of the superconducting nature of SQUID input circuits, true dc response is possible.

Commonly referred to as SQUID magnetometers, these systems are properly SQUID susceptometers. They have a homogeneous superconducting magnet to create a very uniform field over the entire sample measuring region and the superconducting pick-up loops. The magnet induces a moment allowing a measurement of magnetic susceptibility. The superconducting detection loop array is rigidly mounted in the center of the magnet (17). This array is configured as a gradient coil to

reject external noise sources. The detection coil geometry determines what mathematical algorithm is used to calculate the net magnetization. Oppositely paired Helmholtz, first and second derivatives have all been successfully used. Coupling two axial channels of differing gradient order can significantly improve noise rejection.

Sensitivities of  $10^{-17}$  amp have been achieved, even at applied fields of 9 T. Placement of secondary excitation coils can allow ac susceptibility measurements to be made in the presence of a significant dc bias field. Variable temperature capability (1.7 K to 800 K) is achieved by placing a reentrant cryostat within the detection coils.

**Other Laboratory Applications.** NMR signals (18) can be measured by placing a sample (e.g., protons or  $^{19}\text{F}$ ) in the center of SQUID detection coils and either sweeping the external field or applying an  $\omega$  excitation to the sample. The same experimental concept can be used to measure electron paramagnetic resonance (EPR) signals. SQUIDs have been used for more esoteric applications including temperature measurements with resolution near 10 K to 12 K. SQUIDs have also been used to measure position for gravity wave detectors with sub Angstrom resolution and tests of Einstein's General Theory of Relativity. Because SQUID magnetometers are vector devices, they can detect rotations better than  $10^{-4}$  arc-seconds in the earth's magnetic field. SQUIDs have been used in searches for dark matter such as Weak Interacting Massive Particles (WIMPs) along with attempts at detecting magnetic monopoles and free quarks.

#### Geophysical Applications (19,20)

SQUID magnetometers are used to measure the earth's magnetic field (Fig. 9) at frequencies ranging between 1 kHz and  $10^{-4}$  Hz. A technique known as magnetotellurics (21) can be used to determine the electrical conductivity distribution of the earth's crust by measuring the earth's electric and magnetic field. Because the earth is a good electrical conductor compared to the air, the electrical field generated in the ionosphere (as a result of solar wind) is reflected at the earth's surface, with components of both the electric and magnetic field decaying as they penetrate into the earth. The decay length or skin depth  $\delta = 500\sqrt{\rho T}$ , where  $\rho$  is the electrical resistivity of the earth and  $\tau$  is the period of the electromagnetic wave.

In magnetotellurics, the electric field (as a function of frequency) is related to the magnetic field via an impedance tensor where  $\mathbf{E}(\omega) = \mathbf{Z}_{\text{eff}} \mathbf{H}(\omega)$ . The impedance tensor  $\mathbf{Z}_{\text{eff}}$  contains four complex elements  $Z_{xx}$ ,  $Z_{yy}$ ,  $Z_{xy}$ , and  $Z_{yx}$  and is related to the resistivity by  $\rho(\omega) = 0.2 |Z_{ij}(\omega)|^2$  where  $Z_{\text{eff}}$  has units of mV/km-A.

Magnetic anomaly detection uses the five unique spatial components of  $\nabla B$  to locate a magnetic dipole uniquely. This has potential uses in mineralogical surveys and detection of unexploded ordnance.

#### Nondestructive Test and Evaluation (22,23)

Magnetic-sensing techniques such as eddy current testing have been used for many years to detect flaws in structures. A major limitation on their sensitivity is the skin depth (Eq. 17) of metallic materials. Because SQUID sensors have true dc response and superior sensitivity, they can see "deeper" into metallic structures. dc response also means that they can de-

**Table 2. NDE Measurement Techniques**

| Imaging  |  |
|--|--|
| Intrinsic currents   |  |
| Remnant magnetization  |  |
| Flux-induced perturbations in applied currents   |  |
| Johnson noise in metals  |  |
| Eddy currents in an applied ac field (flows)   |  |
| Embedded magnetic sensors  |  |
| Hysteretic magnetization due to:   |  |
| Cyclic stress-strain   |  |
| Simultaneous dc & ac magnetic fields   |  |
| Magnetization of paramagnetic, diamagnetic and ferrimagnetic materials in dc magnetic fields |  |

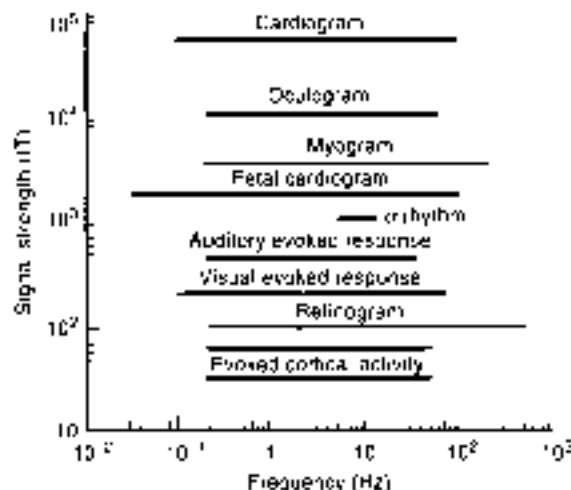
tect remnant magnetization—without the need for externally applied magnetic fields. Their flat frequency response and zero phase distortion allows for a wide range of applications. One potential application of SQUIDs is in detection of stress or corrosion in reinforcing rods used in bridges, aircraft runways, or buildings. Table 2 shows some of the measurement techniques that can be used with SQUID sensors.

SQUID magnetometers have been used to make noncontact measurements of timing circuits (24)—one instrument has better than 10  $\mu\text{m}$  resolution (25). Such instruments with MHz bandwidths could be used for circuit board and integrated circuit (IC) mapping.

#### Medical Applications of SQUIDs (23,26–29)

The use of bioelectric signals as a diagnostic tool is well known in medicine (e.g., the electrocardiogram (ECG) for the heart and the electroencephalogram (EEG) for the brain). The electrical activity that produces the surface electrical activity that is measured by EEG and ECG also produces magnetic fields. The analogous magnetic measurements are known as the magnetoencephalogram (MEG) and the magnetoecardiogram (MEC). Other physiological processes also generate electrical activity with analogous magnetic fields (Fig. 14).

Magnetic fields from active electrical sources in the body can be measured passively and external to the body by placing the magnetometer in close proximity to the body's surface. It has been shown that a population of neurons in the brain can be



**Figure 14.** Typical amplitudes and frequency ranges for various bio-magnetic signals.

Table 3. Medical Applications of SQUIDs

|   |
|---|
| Studies of the brain—neuro-magnetism        |
| Epilepsy                                    |
| Pre-surgical cortical function mapping      |
| Drug development and testing                |
| Stroke                                      |
| Alzheimer's                                 |
| Neurovascular disorders                     |
| Prenatal brain disorders                    |
| Performance evaluation                      |
| Studies of the heart—magnetoencephalography |
| Arrhythmia                                  |
| Heart muscle damage                         |
| Fetal cardiography                          |
| Other medical applications                  |
| Studies of the stomach—gastroenterology     |
| Intestinal ischemia                         |
| Noninvasive in vivo magnetic liver biopsies |
| Lung function and clearance studies         |
| Nerve damage assessment                     |

modeled as a current dipole that generates a well-defined magnetic field profile. Mapping of these field profiles can be used to infer the location of the equivalent active dipole site regions to within millimeters. Using evoked response techniques, the location of signal pathways and information processing centers in the brain can be mapped at different delay times (latencies) following the stimulus.

There are also magnetic measurements for which there are no electrical analogs. These are measurements of static magnetic fields produced by ferromagnetic materials ingested into the body and measurements of the magnetic susceptibility of materials in the body. In particular, information on the quantity and depth of diamagnetic or paramagnetic materials (such as iron stored in the liver) can be obtained by using magnetizing and detection coils of differing sizes in the same instrument and measuring the induced field as a function of distance. This technique is already being used to monitor patients suffering from iron overload diseases such as thalassemia and hemochromatosis.

The development of the SQUID has allowed the development of noninvasive clinical measurements of biomagnetic fields. The use of gradiometers can allow measurements to be made in unshielded environments at sensitivities below  $20 \text{ (fT)} \sqrt{\text{Hz}}$ . Typically, however, neuromagnetic measurements are made in room-sized MSEs that will allow measurements of the magnetic field of the brain over the entire surface of the head ( $>150$  positions simultaneously). Table 3 gives some of the areas in which SQUID magnetometers are currently being used in medical research.

## BIBLIOGRAPHY

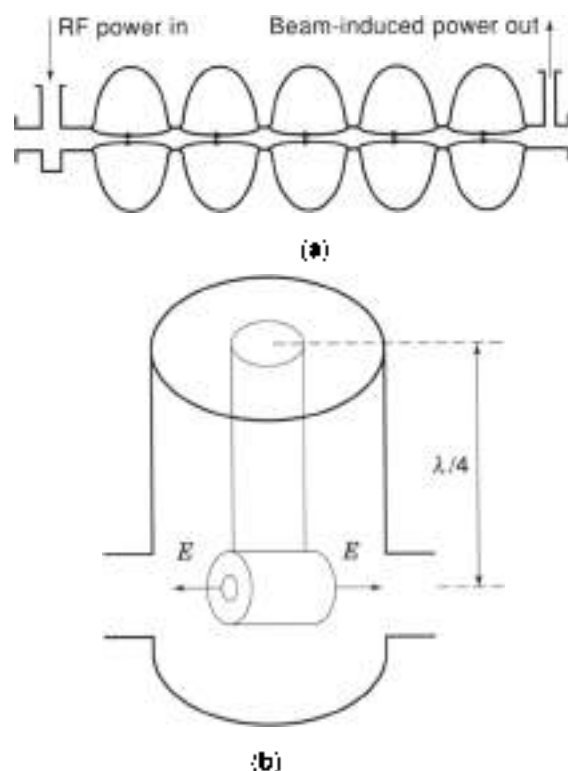
- B. D. Josephson, Possible new effect in superconductive tunneling, *Phys. Lett.*, **1**: 251–253, 1965.
- J. Clarke, SQUID fundamentals, in H. Weinstock (ed.), *SQUID Sensors: Fundamentals, Fabrications and Applications*, Dordrecht: Kluwer, 1997.
- T. Van Duzer and C. W. Turner, *Principles of Superconductive Devices and Circuits*, New York: Plenum, 1981.
- T. P. Orlando and K. A. Delin, *Fundamentals of Applied Superconductorivity*, Reading, MA: Addison-Wesley, 1991.
- R. P. Giffard, R. A. Webb, and J. C. Wheatley, Principles and methods of low-frequency electric and magnetic measurements using an off-biased point-contact superconducting device, *J. Low Temp. Phys.*, **6**: 503–510, 1972.
- R. J. Franke et al., Fully engineered high performance UHF SQUID magnetometer, *Cryogenics*, **21**: 501–506, 1981.
- C. D. Triche and J. Clarke, DC SQUID: noise and optimization, *J. Low Temp. Phys.*, **20**: 301–331, 1982.
- M. B. Sammonds and R. P. Giffard, *Apparatus for reducing low frequency noise to dc biased SQUIDs*, US Patent No. 4,369,612, 1983.
- R. L. Fagaly, *Superconducting magnetometers and instrumentation*, *Sci. Prog., Oxford*, **71**: 191–201, 1987.
- E. W. Grover, *Inductance Calculations, Working Formulas and Tables*, New York: Dover, 1962.
- J. P. Wikswa, Jr., Optimization of SQUID differential magnetometers, *AIP Conference Proc.*, **44**: 145–149, 1978.
- R. B. Stephens and R. L. Fagaly, High temperature superconductors for SQUID detection coils, *Cryogenics*, **31**: 988–992, 1991.
- R. Huisman et al., Multi-SQUID devices and their applications, in D. F. Brewer (ed.), *Progress in Low Temperature Physics*, vol. XII, Amsterdam: Elsevier, 1989.
- J. Véhs, SQUID gradiometers in real environments, in H. Weinstock (ed.), *SQUID Sensors: Fundamentals, Fabrications and Applications*, Dordrecht: Kluwer, 1997.
- D. S. Buchanan, D. N. Paulson, and S. J. Williamson, Instrumentation for clinical applications of neuro-magnetism, in R. W. Fast (ed.), *Advances in Cryogenic Engineering*, vol. 33, New York: Plenum, 1988, pp. 97–106.
- C. Heiden, Pulse tube refrigerators: A cooling option, in H. Weinstock (ed.), *SQUID Sensors: Fundamentals, Fabrications and Applications*, Dordrecht: Kluwer, 1997.
- R. E. Sarasonski, Superconducting instrumentation, *Cryogenics*, **17**: 671–679, 1977.
- R. A. Webb, New technique of improved low-temperature SQUID NMR measurements, *Rev. Sci. Instrum.*, **48**: 1585, 1977.
- J. Clarke, Geophysical Applications of SQUIDs, *IEEE Trans. Mag. MAG-19*: 249–256, 1983.
- H. Weinstock and W. C. Overton, Jr. (eds.), *SQUID Applications to Geophysics*, Tulsa: Soc. of Exploration Geophysicists, 1981.
- K. Vozoff, The magnetotelluric method in the exploration of sedimentary basins, *Geophysics*, **37**: 98–114, 1972.
- G. B. Donaldson, SQUIDs for everything else, in H. Weinstock and M. Nisenoff (eds.), *Superconducting Electronics*, New York: Springer-Verlag, 1989.
- J. P. Wikswa, Jr., SQUID magnetometers for biomagnetism and non-destructive testing: Important questions and initial answers, *IEEE Trans. Appl. Supercond.*, **5**: 74–120, 1995.
- R. L. Fagaly, SQUID detection of electronic circuits, *IEEE Trans. Magn.*, **MAG-25**: 1216–1218, 1989.
- J. Kutley, Imaging magnetic fields, *IEEE Spectrum*, **33**: 12: 40–49, 1996.
- S. J. Williamson and L. Kaufman, Biomagnetism, *J. Magn. Mag. Mat.*, **22**: 129–202, 1981.
- G.-L. Koenig, S. J. Williamson, and L. Kaufman, Diamagnetic instrumentation, *Rev. Sci. Instrum.*, **63**: 1818–1846, 1992.
- R. L. Fagaly, Neuro-magnetic instrumentation, in S. Sato (ed.), *Advances in Neurology*, Vol. 54: *Magnetoencephalography*, New York: Raven Press, 1991.
- C. Anne et al., *Advances in Biomagnetism Research*, Biomag96, New York: Springer-Verlag, 1997.

ROBERT L. FAGALY  
Tristan Technology

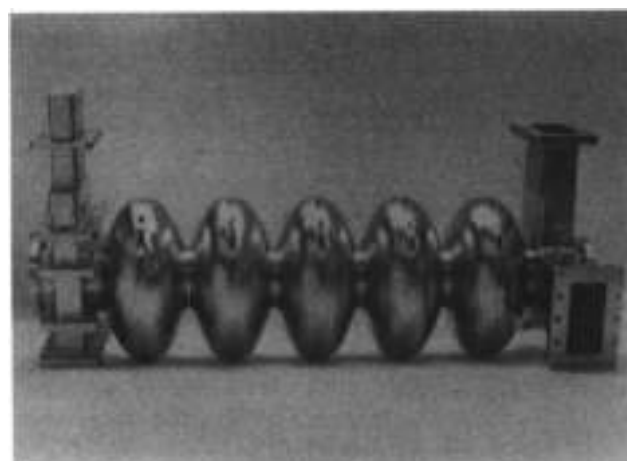
## SUPERCONDUCTING CAVITY RESONATORS

A key component of the modern particle accelerator (59) is the device that imparts energy to the charged particles. This is an electromagnetic radio frequency (RF) cavity resonating at microwave frequencies typically between 50 MHz and 3000 MHz. Traditionally, accelerating devices are normal conducting cavities typically made from copper (59). One of the main incentives for using superconducting cavities is that the dissipation in the walls of the copper structure can be substantially reduced. This is especially beneficial for accelerators that operate in a continuous wave (CW) mode or at a high duty factor (e.g., >1 percent). Superconducting cavities economically provide high CW operating fields. Another benefit is that superconducting cavities can be designed to have a large beam aperture which reduces the beam-cavity interactions, allowing higher beam quality and higher beam current.

There are two distinct types of superconducting cavities, depending on the velocity of the particles. The first category is for accelerating charged particles that move at nearly the speed of light, such as electrons in a high-energy linear accelerator [e.g., at TJNAP (1) at Jefferson Lab in Newport News, VA] or a storage ring [e.g., LEP (2) at CERN in Switzerland]. The second type is for particles that move at a small fraction (e.g., 0.01 to 0.9) of the speed of light, such as the heavy ions emerging from



**Figure 1.** (a) An accelerating structure for velocity of light particles. The resonant frequency for superconducting structures is typically between 350 MHz and 3000 MHz. The cell length is half a wavelength  $\lambda/2$  long. The phase of the electric field on the axis of each cell is shown for the accelerating mode. Ports outside the cell region are for input power couplers and higher order mode power output couplers. In many applications, the power induced by the beam in higher order modes has to be removed by output couplers. (b) An accelerating structure for low velocity particles, such as heavy ions. The resonant frequency is typically between 50 MHz and 150 MHz. The accelerating gaps are  $\beta\lambda/2$  in length, where  $\beta = v/c$  and  $v$  is the velocity of the heavy ions.



(a)



(b)

**Figure 2.** (a) Five-cell 1.5 GHz niobium cavity developed at Cornell, now used at TJNAP. (b) A quarter-wave resonator from niobium developed for the JAERI/Tokai, Japan, heavy ion linac.

a de high-voltage Van de Graaff accelerator, ATLAS (3) at Argonne National Lab, Argonne, IL is the longest-running heavy ion accelerator facility. Figure 1(a) is a sketch of the typical superconducting accelerating structure of the first type, and Fig. 2(a) is a corresponding photograph (4). There are five accelerating cells that resonate in the  $TM_{010}$  mode of the cylindrical cavity. As the particle traverses each half-wavelength ( $\lambda/2$ ) accelerating gap in half a radio-frequency (RF) period, it sees a longitudinal electric field pointing in the same direction for continuous acceleration. Figure 1(b) is a sketch for a structure for low velocity particles, and Fig. 2(b) is a corresponding photograph (5). A coaxial transmission line a quarter wavelength long resonates in the TEM mode. A drift tube is suspended from the end of the hollow center conductor. The structure has two accelerating cells between the ends of the drift tube and the beam hole openings located in the outer conductor of the coax. The accelerating gap is  $\beta\lambda/2$ , where  $\beta = v/c$ . Since  $\beta$  is small,  $\lambda$  must be chosen to be large, to achieve a useful acceleration. Therefore a low resonant frequency is chosen, typically 100 MHz. The

wavelength also sets the height of the quarter-wave resonator. The example of Fig. 2 has a manageable height of less than one meter.

Large-scale application of superconducting cavities to electron and ion accelerators is now established at many laboratories around the world (6). These accelerators provide high-energy electron and positron beams for elementary particle research, medium-energy electron beams for nuclear physics research, low-energy, heavy ion beams for nuclear research, and high-quality electron beams for free electron lasers. Altogether more than 500 meters of superconducting cavities have been installed worldwide and successfully operated at accelerating fields up to 6 MV/m to provide a total of more than 2.5 GV for a variety of accelerators.

The two most salient characteristics of an accelerating cavity are its average accelerating field,  $E_{acc}$ , and the quality factor  $Q_0$ . The typical accelerating field at which  $\beta = 1$  superconducting cavities are now operated routinely is  $E_{acc} = 5$  MV/m, and the typical  $Q_0$  value is  $2 \times 10^9$ . The corresponding numbers for low-velocity structures are 3 MV/m and  $10^9$ . Accelerating fields as high as 40 MV/m and  $Q_0$  values as high as  $10^{11}$  have been reached in high-performance  $\beta = 1$  superconducting test cavities.

The strongest incentive to use superconducting cavities in accelerators that operate in a continuous-wave (CW) mode, or at a high duty factor ( $> 1\%$ ), for CW operation, the power dissipation in the walls of a structure built from normal conducting material (such as copper) is substantial. Therefore the typical CW operating field for a copper cavity is usually kept below 1 MV/m. The microwave surface resistance of a superconductor is typically five orders of magnitude lower than that of copper, and therefore the  $Q_0$  is five orders of magnitude higher. For applications demanding high CW voltage, such as increasing the energy of electron storage rings, the advantage of superconducting cavities becomes clear. Since the dissipated power increases with the square of the operating field, only superconducting cavities can economically provide the needed voltage. For example, LEP requires 2.5 GV to double its energy from 50 GeV to 100 GeV per beam. If copper cavities were to be used, both the capital cost of the klystrons and the ac power operating cost would become prohibitive at the higher accelerating field. Several MW/m of ac power would be required to operate a copper cavity at 5 MV/m. There are also practical limits to dissipating high power in the walls of a copper cavity. When more than 100 kW is dissipated in a copper cell, the surface temperatures exceed 100 C, causing vacuum degradation, stresses, and metal fatigue due to thermal expansion. High accelerating fields (4–100 MV/m) can be produced in copper cavities, but only for microseconds, and the peak RF power needed (59) becomes enormous (many 100 MW).

Apart from the general advantages of reduced RF capital and reduced RF associated operating costs, superconductivity offers certain special advantages that stem from the low cavity wall losses. Because of the low power dissipation at high accelerating field, one can afford to make the beam hole of superconducting cavity much larger than for a normal conducting cavity. The large beam hole substantially reduces the beam-cavity interaction (or wake fields (59)), allowing better beam quality and higher current for improving the precision and reaction rates of physics experiments. For the intense proton beams, where scraping of the proton beam tails is a major worry because of radon activation of the accelerator, the wide beam hole greatly reduces the risk of beam-loss-induced radioactivity.

## RF Superconductivity Basics

The remarkable properties of superconductivity are attributed to the condensation of charge carriers into Cooper pairs, which move frictionlessly. At  $T = 0$  K, all charge carriers are condensed. At higher temperatures, some carriers are unpaired; the fraction of unpaired carriers increases exponentially with temperature, as  $e^{-\Delta/kT}$ , until none of the carriers are paired above  $T_c$ . Here  $2\Delta$  is the energy gap of the superconductor, the energy needed to break up the pairs. In this simplified picture, known as the *London two-fluid model*, when a dc field is turned on, the pairs carry all the current, shielding the applied field from the normal electrons. Electrical resistance vanishes.

In the case of RF currents, however, dissipation does occur for all  $T > 0$  K, albeit very small compared to the normal conducting state. While the Cooper pairs move frictionlessly, they do have inertial mass. For high-frequency currents to flow, forces must be applied to bring about alternating directions of flow. Hence an ac electric field will be present in the skin layer, and it will continually accelerate and decelerate the normal carriers, leading to dissipation proportional to the square of the RF frequency. The two-fluid model provides a simple explanation for the quadratic frequency and the exponential temperature dependence of the RF surface resistance. The power dissipated is proportional to the internal electric field (proportional to the RF frequency) and to the normal component of the current. The "normal" component of the current, being proportional to the interior electric field, gives another factor proportional to the RF frequency. The normal component of the current also depends on the number of carriers thermally excited across the gap  $2\Delta$  and is given by the Boltzmann factor  $e^{-\Delta/kT}$ .

Besides the phenomenally low RF surface resistance, other important fundamental aspects are the maximum surface fields that can be tolerated without increasing the microwave surface resistance substantially or without causing a breakdown of superconductivity. The accelerating field,  $E_{acc}$ , is proportional to the peak surface RF electric field ( $E_{pk}$ ), as well as the peak surface RF magnetic surface field ( $H_{pk}$ ).

The ultimate limit to the accelerating field is the RF critical magnetic field. Theoretically, this is equal to the superheating critical magnetic field. In the Ginzburg-Landau phenomenological theory of superconductivity (7), surface energy considerations lead to estimates for superheating critical field in terms of the thermodynamic critical field,  $H_c$ , and the Ginzburg-Landau parameter,  $\kappa$ , as follows:

$$\begin{aligned} H_{c1} &\approx 0.69 H_c & \text{for } \kappa \ll 1 \\ H_{c1} &\approx 1.2 H_c & \text{for } \kappa \approx 1 \\ H_{c1} &\approx 0.75 H_c & \text{for } \kappa \gg 1 \end{aligned} \quad (1)$$

For the most commonly used superconductor, niobium,  $H_{c1}$  is about 200 mT, which translates to a maximum accelerating field of 55 MV/m for a typical  $\beta = 1$  niobium structure and roughly 30 MV/m for a  $\beta < 1$  niobium structure.

Typically, cavity performance is, however, significantly below the theoretically expected surface field. One important phenomenon that limits the achievable RF magnetic field is "thermal breakdown" of superconductivity, originating at submillimeter-size regions of high RF loss, called "defects." When the temperature outside the defect exceeds the superconducting transition temperature,  $T_c$ , the losses increase, because

large regions become normal conducting. Several measures have been developed to overcome thermal breakdown, such as (a) improving the thermal conductivity of niobium by purification or (b) using thin films of niobium (or lead) on a copper substrate cavity.

In the early stages of the development of superconducting cavities, a major performance limitation was the phenomenon of "multipacting." This is a resonant process in which a large number of electrons builds up within a small region of the cavity surface due to the fact that the secondary electron emission coefficient of the surface is greater than unity. The avalanche absorbs RF power, making it impossible to raise the fields by increasing the incident RF power. The electrons impact the cavity walls, leading to a large temperature rise and eventually to thermal breakdown. With the invention of the spherical cavity shape (8) and later the elliptical cavity shape (9), multipacting is no longer a significant problem for velocity-of-light structures. Multipacting is still an impediment for structures for low-velocity particles but can be reduced by long periods of exposure to high RF power, called *conditioning*, during which the secondary electron emission is reduced by long-term electron bombardment.

In contrast to the magnetic field limit, there is no known theoretical limit to the tolerable surface electric field. Continuous wave electric fields up to 145 MV/m (10) and pulsed electric fields up to 220 MV/m (11) have been imposed on a superconducting niobium cavity surface without any catastrophic effects. However, at high electric fields, an important limitation to the performance of superconducting cavities arises from the emission of electrons from high-electric-field regions of the cavity. Power is absorbed by the electrons and deposited as heat when electrons impact the cavity walls. If the emission grows intense, it can even initiate thermal breakdown. There have been extensive studies about the nature of field emission sites as well as development of techniques to avoid emission sites and to destroy them (12).

For low-velocity accelerators, there is an important additional performance consideration. Ambient acoustic noise (microphonics) excites mechanical vibrational modes of the cavity, causing the resonant frequency to vary. The resonant cavities are extended, loaded structures (e.g., drift tubes supported by pipes) and generally have reduced mechanical stability. The cavity RF phase must be synchronized with an RF clock. This requires rapidly tuning the cavity to cancel the effects of acoustically induced mechanical distortions (13).

**Figures of Merit for a Superconducting Cavity**

We show how to calculate the important physical quantities, such as resonant frequency, accelerating field, peak electric and magnetic fields, power dissipation, the quality factor  $Q_c$ , and shunt impedance for a simple cavity, the cylindrically symmetric pillbox. The treatment is basic to both normal conducting and superconducting cavities (see CAVITY RESONATORS). We also work out illustrative values. Similar analytic calculations can be carried out for a coaxial TEM quarter wave resonator, as illustrative of an accelerating structure for low-velocity particles. Only simple structures can be calculated analytically. For real structures with beam holes, it is necessary to use field computation codes, such as (a) URMEL (14) for cylindrically symmetric structures and (b) MAFIA (15) for 3-D geometries.

For a cylinder of length  $d$  and radius  $R$ , the electric ( $E_z$ ) and magnetic ( $H_\theta$ ) fields for the standing wave  $TM_{010}$  mode are

$$E_z = E_0 J_0 \left( \frac{2.405 \rho}{R} \right) e^{i\omega t}, \quad H_\theta = -i \frac{c}{\sqrt{\epsilon_0}} E_0 J_1 \left( \frac{2.405 \rho}{R} \right) e^{i\omega t} \quad (2)$$

where all other field components are 0.  $J_0$  and  $J_1$  are Bessel functions of the radial coordinate. The angular resonant frequency  $\omega = 2\pi f$  is given by

$$\omega_{TM0} = \frac{2.405 c}{R} \quad (3)$$

Note that the resonant frequency,  $f$ , is independent of the cavity length.

Assume an electron traveling nearly at the speed of light ( $v$ ). It enters the cavity at time  $t = 0$  and leaves at a time  $t = d/c$ . To receive the maximum kick from the cavity, the time it takes the particle to traverse the cavity is to equal one-half of an RF period, that is

$$t = \frac{d}{c} = \frac{1}{2} T_{RF} = \frac{\pi}{\omega} \quad (4)$$

Under this condition, the electron always sees a field pointing in the same direction. The accelerating voltage ( $V_{acc}$ ) for a cavity is

$$V_{acc} = \left| \int_{z=0}^{z=d} E_z dz \right| \quad (5)$$

For an electron accelerator with energy  $> 10$  MeV, it is sufficiently accurate to use  $v = c$ , so that  $t(z) = z/c$ . Thus

$$V_{acc} = \left| \int_{z=0}^{z=d} E_z(z) e^{i\omega t(z)} dz \right| \quad (6)$$

$$V_{acc} = E_0 \left| \int_{z=0}^{z=d} e^{i\omega z/c} dz \right| = d E_0 \frac{\sin \left( \frac{\omega d}{2c} \right)}{\frac{\omega d}{2c}} = d E_0 T \quad (7)$$

Here  $T$  is referred to as the "transit time factor." At 1.5 GHz we have  $d = \pi r / \omega = 10$  cm, and Eq. (7) simplifies to

$$V_{acc} = 0.064 \pi \cdot E_0$$

The average accelerating electric field ( $E_{acc}$ ) is given by

$$E_{acc} = \frac{V_{acc}}{d} = \frac{2E_0}{\pi} \quad (8)$$

Here  $E$  has the dimensions of V/m.

**Peak Surface Fields**

To achieve a high accelerating field in a cavity, it is important to minimize the design values of the peak fields to the accelerating field. For the  $TM_{010}$  mode in a pillbox cavity we have

$$E_{pk} = E_0, \quad H_{pk} = \frac{c}{\sqrt{\epsilon_0}} J_1(1.8411) E_0 = \frac{K_1}{6.47 \Omega} \quad (9)$$



Thus we obtain the following ratios:

$$\frac{E_{\text{pk}}}{E_{\text{acc}}} = \frac{\pi}{2} = 1.6, \quad \frac{H_{\text{pk}}}{E_{\text{acc}}} = 2430 \frac{\text{A/m}}{\text{MV/m}} = 3.05 \frac{\text{mT}}{\text{MV/m}} \quad (10)$$

The units for magnetic field used are Tesla.

**Power Dissipation and  $Q_0$**

In order to support the electromagnetic fields, currents flow within a thin surface layer of the cavity walls. If the surface resistance is  $R_s$ , the power dissipated (unit area  $P_s$ ) due to Joule heating is

$$P_s = \frac{1}{2} R_s H^2 \quad (11)$$

The quality,  $Q_0$ , is related to the power dissipation by the definition of  $Q_0$ :

$$Q_0 = \omega \frac{\text{Energy stored}}{\text{Power dissipated}} = \frac{\omega U}{P_s} \quad (12)$$

where  $U$  is the stored energy and  $P_s$  is the dissipated power.

The total energy in the cavity and the power dissipated are

$$U = \frac{1}{2} \mu_0 \int_V |H|^2 dV, \quad P_s = \frac{1}{2} R_s \int_S |H|^2 dS \quad (13)$$

where the integral is taken over the volume of the cavity. Thus

$$Q_0 = \frac{\omega \mu_0 \int_V |H|^2 dV}{R_s \int_S |H|^2 dS}, \quad Q_0 = \frac{G}{R_s}, \quad G = \frac{\omega \mu_0 \int_V |H|^2 dV}{\int_S |H|^2 dS} \quad (14)$$

Here  $G$  is called the *geometry factor*. It only depends on the cavity shape and not its size. For the  $\text{TM}_{010}$  mode in a pillbox cavity, we obtain

$$U = \frac{\pi \epsilon_0 E_0^2}{2} \int_0^d (2.405)^2 r dr \quad (15)$$

$$P_s = \frac{\pi R_s E_0^2}{\omega} \int_0^d (2.405)^2 (H/R) \cdot d \quad (16)$$

$$G = \frac{453 \text{M}\Omega}{(R^2 - R d)} \quad (17)$$

Combining Eqs. (13) and (14), we find that in order to obtain the maximum accelerating voltage from the cavity, we require

$$\frac{R}{d} = \frac{2.405}{\tau} \quad (18)$$

so that  $G = 257 \Omega$ . A typical observed surface resistance for a well-prepared superconducting Nb cavity is  $R_s = 20 \text{ n}\Omega$ . Thus we have a  $Q_0$  value of

$$Q_0 = \frac{G}{R_s} = 1.3 \times 10^{10} \quad (19)$$

For a typical cavity length of  $d = 10 \text{ cm}$  (at 1.5 GHz), we obtain  $R = 7.65 \text{ cm}$ . For an accelerating voltage of 1 MV, we obtain the

following results.

$$\begin{aligned} E_{\text{acc}} &= \frac{V_{\text{acc}}}{d} = 10 \text{ MV/m} \\ E_{\text{pk}} &= E_0 = \frac{\pi}{2} E_{\text{acc}} = 15.7 \text{ MV/m} \\ I_{\text{pk}} &= 2430 \frac{\text{A/m}}{\text{MV/m}} E_{\text{acc}} = 24.3 \text{ kA/m} = 30.5 \text{ mT} \\ U &= \frac{\pi \epsilon_0 E_0^2}{2} \int_0^d (2.405)^2 r dr = 0.54 \text{ J} \\ P_s &= \frac{\omega U}{Q_0} = 0.4 \text{ W} \end{aligned} \quad (20)$$

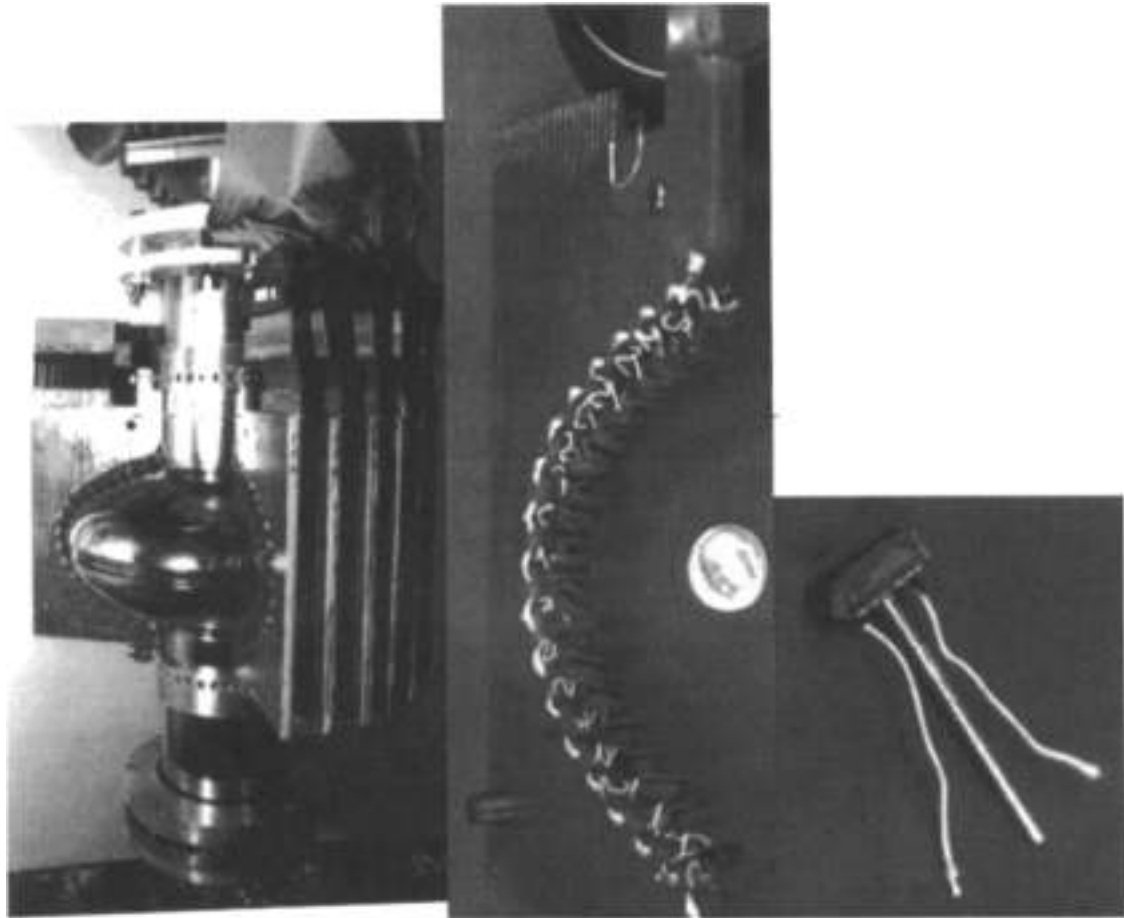
The performance of a superconducting cavity is evaluated by measuring the  $Q_0$  as a function of the cavity field level. This gives information on the average behavior of the RF surface.

**Thermometry Based Diagnostics**

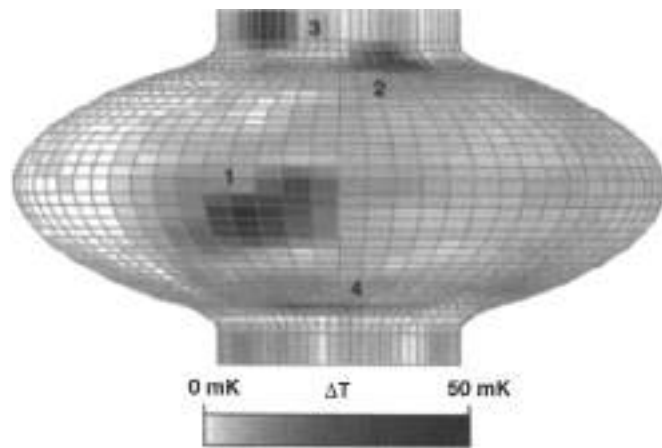
To resolve the local distribution of RF losses from various mechanisms described above, temperature mapping is used as a diagnostic technique. A chain of rotating carbon thermometers, or an array of fixed thermometers, samples the temperature of the outer wall of the cavity. Temperature mapping with carbon thermometers has played a key role in improving the understanding of mechanisms that lead to residual resistance, multipacting, thermal breakdown, and field emission. Carbon makes a sensitive thermometer at liquid helium temperatures because, as a semiconductor, its resistance increases exponentially with decreasing temperature. Above the superfluid temperature (2.17 K), temperature increments of the cavity wall of a few mK can be easily detected. A single rotating arm bearing 10 to 20 thermometers per cell is appropriate for locating stable field emitters or thermal defects in sizable structures, such as a multicell cavity. For temperature mapping in superfluid helium, thermometers need to be isolated from the superfluid bath so that movable elements do not provide good sensitivity. A large array of fixed thermometers is preferred. These are brought in intimate contact with the cavity wall by the use of spring loaded contacts. Grease applied between the cavity wall and the thermometer element improves heat transfer and keeps the superfluid away. Due to the large number of thermometers and leads, the fixed method is suitable for investigations with single-cell cavities. An example of a fixed thermometry system is shown in Fig. 3, and a typical temperature map is shown in Fig. 4 (16).

**Refrigerator Requirements**

Although the power dissipated in the superconducting cavity is very small, the losses will be dissipated in the liquid He bath. Together with the static heat leak to the cryostat, these losses comprise the cryogenic loss. Typically the ac power needed to operate the refrigerator is larger than the dissipated power in 2K liquid He by a factor of 750. One part of this factor comes from the technical efficiency ( $\eta$ ) of the refrigerator, typically  $\eta = 0.2$  for a large system, and the other part comes from the



**Figure 3.** (a) A single-cell niobium cavity surrounded by an array of 8700 carbon thermometers that make close contact with the outer wall of the cavity. (b) There are 18 thermometers placed on each individual board that is contoured to closely follow the cavity profile. (c) A single thermometer consists of a 100  $\Omega$  carbon resistor embedded in an epoxy housing. It is held by a spring-loaded pin inserted into holes in the board. The surface of the thermometer is ground so that the carbon element is exposed and subsequently covered with a thin layer of varnish to provide electrical insulation. The leads are made of a low-thermal-conductivity alloy, such as manganin.



**Figure 4.** Temperature map at 40 mT of a single-cell 1.5 GHz cavity showing heating at a defect site near the cavity equator (labeled 1) and field erosion sites (labeled 2, 3, and 4) near the cavity ends.

Coupled efficiency  $\eta_c$ , which at 2 K is

$$\eta_c = \frac{300 - 2}{2} \quad (21)$$

At 10 MV/m the required refrigerator ac power due to the RF loss would be 300 W for the case of a single-cell 1.5 GHz cavity. For a copper cavity of the same geometry, with a typical  $R_s = 8 \text{ m}\Omega$ , the RF power dissipation in the cavity would be 60 kW for an accelerating field  $E_{acc} = 10 \text{ MV/m}$ . Furthermore, the ac wall power will be a factor of 2 higher because of the typical klystron efficiency. Thus the ac power cost of running a copper cavity in CW mode would be several hundred times higher than the cost for an Nb cavity.

**Shunt Impedance**

An important quantity used to characterize the losses in a cavity at a given accelerating voltage is the shunt impedance  $R_{sh}$

**Table 1. Figures of Merit for the Cornell/CBEAF 5-Cell Cavity**

|                                  |               |
|----------------------------------|---------------|
| $Q$                              | 290 41        |
| $R/Q$ (per 5-cell cavity)        | 480 $\Omega$  |
| $F_{\text{cell}}/R_{\text{sur}}$ | 2.0           |
| $H_{\text{cell}}/E_{\text{acc}}$ | 4.7 MT/(MV/m) |

\*Data taken from Ref. 4

as typified by a parallel RLC circuit:

$$R_0 = \frac{V_{\text{acc}}^2}{P} \tag{22}$$

in which case  $P$  = power dissipated and  $V_{\text{acc}}$  is the acceleration voltage. Hence the shunt impedance is in  $\Omega$ .

Ideally the shunt impedance should be large for the accelerating mode so that the dissipated power is small. This is particularly important for copper cavities, where the wall power dissipation is a major issue and we wish to have as large an accelerating field as possible. For the  $TM_{010}$  mode pillbox cavity and  $R_0$  of 20  $n\Omega$  we have

$$R_0 = \frac{4V_{\text{acc}}^2}{\pi^2 R_{\text{cell}} n Q_c^2 [2.405/R(R-1)]} = 2.5 \times 10^{12} \Omega \tag{23}$$

Note that the ratio of  $R_0/Q$  is given by

$$\frac{R_0}{Q_0} = \frac{V_{\text{acc}}^2}{n I^2} \tag{24}$$

which is independent of the surface resistance. For the pillbox  $TM_{010}$  mode we have

$$\frac{R_0}{Q_0} = 150 \Omega \frac{d}{R} = 146 \Omega \tag{25}$$

By applying computer codes to determine electromagnetic fields, the computed figures of merit for the Cornell/TJNAF 5-cell cavity are given in Table 1. Note that due to the presence of the beam holes the shunt impedance is reduced and the peak surface fields are enhanced, relative to the pillbox case. For a realistic cavity shape,  $R/Q_0$  is lowered due to the presence of the beam holes, typically by a factor of 2.

**RF Surface Resistance**

Based on the very successful BCS theory (17), expressions for the superconducting surface impedance have been worked out by Mattis and Bardeen (18). These expressions involve material parameters, such as the London penetration depth  $\lambda_L$ , the coherence distance  $\xi_0$ , the Fermi velocity  $V_F$ , and the electron mean free path  $l$ . They are in a rather difficult form to obtain general formulas to work with. Computer programs have been written—for example, by Turneaure (19) and Halbritter (20). Figure 5 gives the results from Halbritter's programs for niobium and lead and  $Nb_3Sn$ . Table 2 gives the material parameters used for the calculations. Calculations from the theory agree well with experimentally measured  $R_s$  for  $T/T_c < 0.3$ . At lower temperatures the residual resistance term dominates.

A simplified form of the temperature dependence of  $Nb$  for  $T_c/T > 2$  and for frequencies much smaller than  $2\Delta/h$  is

**Table 2. Material Parameters Used for the Calculations of Fig. 5**

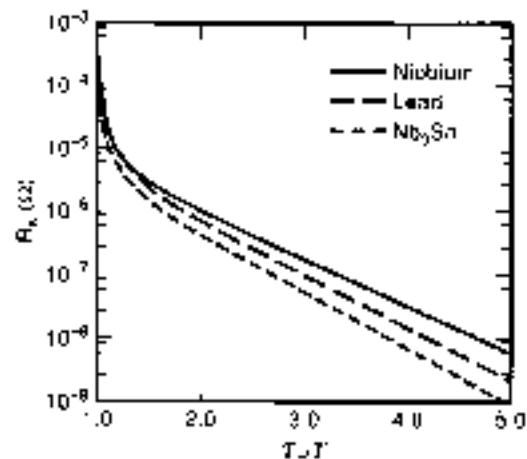
| Material Parameter                             | Pb     | Nb   | $Nb_3Sn$ |
|--|--------|------|----------|
| $T_c$ (K)                                      | 7.19   | 9.20 | 18.00    |
| Energy gap, $\Delta$ (eV)                      | 2.10   | 1.86 | 2.25     |
| Penetration depth $\lambda_L$ ( $\text{\AA}$ ) | 260    | 360  | 600      |
| Coherence length $\xi_0$ ( $\text{\AA}$ )      | 1110   | 840  | 60       |
| Mean free path $l$ ( $\text{\AA}$ )            | 10,000 | 500  | 10       |

$10^{12}$   $\Omega$  is

$$R_s = A(T/T_c)^2 \exp(-\Delta/T)/k_B T + R_0 \tag{26}$$

Here  $k_B$  is Boltzmann's constant. The second term,  $R_0$ , is called the residual resistance. Typical  $R_0$  values for  $Nb$  cavities fall in the range from  $10^{-7}$  to  $10^{-8} \Omega$ . The record for the lowest surface resistance is  $1.2 \times 10^{-9} \Omega$  (21). For comparison, the surface resistance of copper at 1.5 GHz is 3 m $\Omega$ .

The operating temperature of a superconducting cavity is usually chosen so that the first term in Eq. (26) is reduced to an economically tolerable value.  $R_0$ , referred to as the residual resistance, is influenced by several factors. Some of the sources are extraneous to the superconducting surface—for example, lossy joints between components of the structure. Other factors originate at the superconducting surface. A well-understood and controllable source of residual loss is trapped dc magnetic flux from insufficient shielding of the earth's magnetic field, or other dc magnetic fields in the vicinity of the cavity. To get the highest  $Q_0$ , a superconducting cavity must be well-shielded from the earth's field. Typically, at 1 GHz,  $R_0$  is 10  $\mu\Omega$ /mT (22). Another important residual loss mechanism arises when the hydrogen dissolved in bulk niobium precipitates as a lossy hydride at the  $Nb$  surface (23). This residual loss is a subtle effect that depends on the rate of cooldown and the amount of other interstitial impurities present in niobium. The effect can be severe enough to lower the  $Q_0$  to  $10^7$  depending on the amount of hydrogen dissolved and the cooldown rate of the cavity. More than 2 ppm wt of hydrogen can be dangerous.



**Figure 6.** Theoretical surface resistance at 1.5 GHz of lead, niobium and  $Nb_3Sn$  as calculated from Halbritter's program (20). The values used for the material parameters are given in Table 2.

### Cavity Fabrication and Surface Preparation

Niobium cavities can be constructed from sheet niobium using the techniques of forming (e.g., deep drawing or spinning) followed by electron beam welding (24). Another method is to deposit a thin niobium film onto a preformed copper cavity substrate (25). The copper cavity is made in essentially the same way as the sheet niobium cavity, except for surface preparation before film deposition. If the cavity has more than one cell, the cells need to be tuned relative to each other, by adjusting the dimensions, so that the accelerating field is the same for each cell. Dimensional variations between cells are sufficient to alter the field profile substantially. Typical fabrication tolerances are in the range of a few tenths of a  $\mu\text{m}$ .

The purity of niobium used is important, both in terms of bulk impurity content and in terms of inclusions from manufacturing steps, such as rolling. Inclusions on the RF surface play the role of normal conducting sites for thermal breakdown. Dissolved impurities serve as scattering sites for the electrons not condensed into Cooper pairs. These impurities lower the thermal conductivity, impede the heat transfer to the helium, and limit the maximum tolerable surface magnetic field before the onset of thermal breakdown. The accompanying decrease in electrical conductivity, or the  $RRR$  value, serves as a convenient measure of the purity of the metal. The formal definition of  $RRR$  is

$$RRR = \left( \frac{\text{resistivity at 300 K}}{\text{residual resistivity at low temperature (normal state)}} \right) \quad (27)$$

Here low temperature means the temperature at which the dc resistivity in the normal state becomes residual. A convenient relationship between thermal conductivity and  $RRR$  for niobium is

$$k \approx 0.25 RRR \left( \frac{\text{W}}{\text{mK}} \right) \quad (28)$$

This relationship can be derived from the Wiedemann-Franz Law (26) and from the ratio of the superconducting to normal conducting state thermal conductivities (27). To achieve the optimum RF performance, the surface of the cavity must be prepared to approach as close as possible the ideal. Microscopic contaminants can limit the performance, either by thermal breakdown or by field emission (28). A clean RF surface

is achieved by chemically etching away a surface layer, rinsing thoroughly with ultrapure water, and then taking precautions so that no contaminants come in contact with the clean RF surface. The resistivity of the water should be close to theoretically pure (18  $M\Omega\text{-cm}$ ), and the water should be filtered to eliminate particles larger than 1  $\mu\text{m}$ . After etching, water is recirculated for several hours through the cavity in series with the water purification system so as to continuously and thoroughly remove any chemical and particulate residue from the niobium surface. For a review of cavity fabrication and preparation procedures, see Ref. 28.

Many laboratories have found that the RF surface can be made even cleaner if chemistry is followed by high-pressure rinsing (HPR) of the cavity with ultrapure water (29). At T.J. NAF for example, water at a pressure of 70 bar to 80 bar is sprayed through stainless steel nozzles each having a 0.3 mm diameter orifice (30). The potent jets of water are scanned across all parts of the RF surface to dislodge and sweep away microscopic contaminants that have adhered to the surface.

After rinsing, the cavity is transported into a dust-free clean room where the water is drained. The cavity surface thus only comes in contact with filtered air. The level of cleanliness required is comparable to that in the semiconductor industry where a clean room environment of Class 10–100 is routine. Class 10 refers to the number of particles of size 0.5 microns or larger in one cubic foot of air. The surface of the cavity must be dried before the cavity is evacuated, placed inside a cryostat, and cooled down for RF tests. During final assembly, the laboratory workers in the vicinity need to wear special particulate-free clothing and follow strict protocols to reduce particulate generation.

During these various stages of cavity production and preparation, there are many opportunities for defects to enter the cavity. Therefore great care must be exercised during the manufacture of sheet metal, deep drawing of cups, electron beam welding, chemical etching, rinsing, drying, and insertion of coupling devices, as well as in the final attachment of the cavity to the vacuum system of the test stand or the accelerator. Two examples of defects that caused thermal breakdown are shown in Fig. 6 (31). On a statistical basis, we expect that the number of defects increases with cavity surface area.

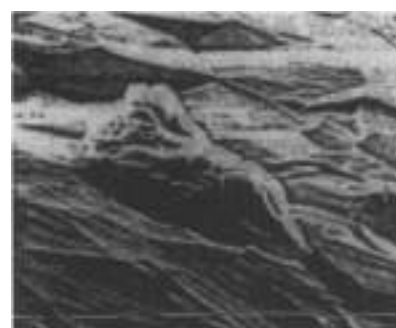
### Overcoming Field Emission

The temperature mapping diagnostic technique for superconducting cavities shows that emission arises from particular spots, called "emitters," located in high-electric-field regions.

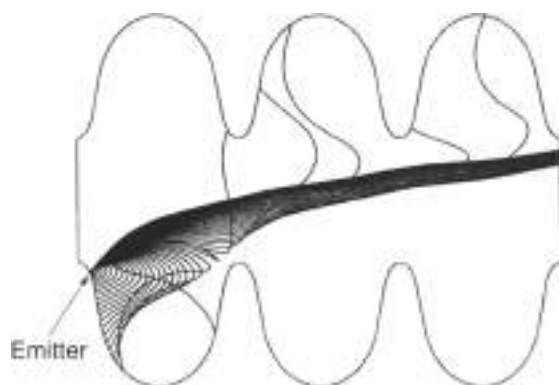
**Figure 6.** SEM micrographs of defects that caused thermal breakdown. (a) A chemical or drying stain 440  $\mu\text{m}$  in diameter. The small crystal on the right side contains K, Cl, and P. This defect quenched at  $E_{\text{th}} = 3.4 \text{ MV/m}$ . (b) A 30  $\mu\text{m}$  crystal containing S, Ca, Cl, and K. This defect quenched at  $E_{\text{th}} = 10.7 \text{ MV/m}$ . These defects were located by temperature maps. (Courtesy of CERN.)



(a)



(b)

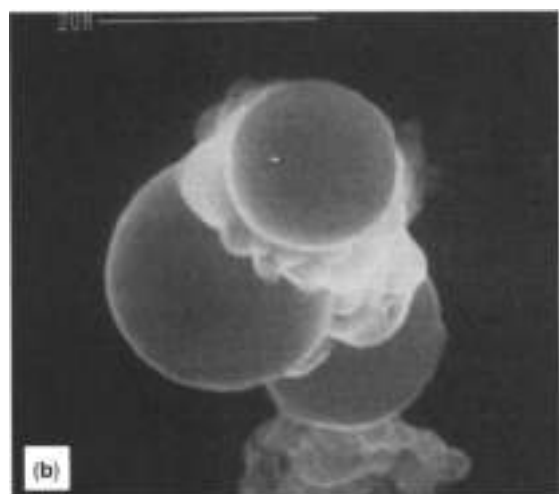
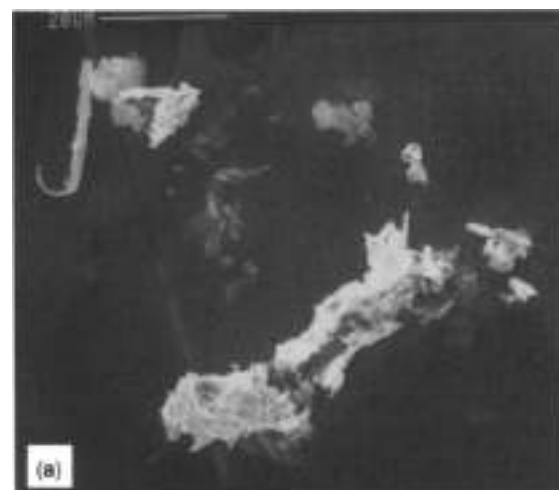


**Figure 7.** Calculated electron trajectories in a 3-cell 1.5 GHz cavity operating at  $E_{pk} = 50$  MV/m. The emitter is located in the end cell, where the surface electric field is 44 MV/m. Note that a significant number of field-emitted electrons bend back and strike the wall near the emitter where they are accelerated through the cavity structure and could produce unwanted "dark current" that may be accelerated in adjacent cavities.

The electrons that emerge from the emitters travel in the RF fields of the cavity and impact the surface (Fig. 7). Some electrons may be captured in the axial fields and accelerated along with the beam. These produce unwanted "dark current," which may spoil the beam quality or input the walls of adjacent cavities. The pattern of temperature rise as a function of position along a given meridian contains implicit information about the location and characteristics of the source. The power deposited by the impacting electrons depends on the trajectory as well as on the intrinsic properties of the emitter.

In their basic theory of field emission (32), Fowler and Nordheim (FN) showed that in the presence of an electric field, electrons tunnel out of the metal into the vacuum, because of their quantum wave-like nature. However, a comparison with the observed currents reveals that, at a given field, emission is substantially higher than the FN predictions. Traditionally, the excess has been attributed to a "field enhancement factor," which is believed to be related to the physical properties of the emitter discussed below. Both RF and dc studies reveal that emitters are micron- to sub-micron-size contaminant particles (33). Figure 8 shows an example of a region of emitting particles found in a multi-cell cavity (33). The properties of the emitter that lead to enhanced emission are (a) the microgeometry of the particle (34), (b) the nature of condensed gases or adsorbates on the surface of the particle (35), and (c) the interface between the particle and the underlying metal RF surface (36). Accordingly, a high level of cleanliness is necessary for cavity surface preparation. Field emission free performance has been achieved with HPR (30). Recently, many 9-cell 1.3 GHz structures were prepared at DESY by using HPR (37). A sample of their results is shown in Fig. 11.

When raising the RF electric field in a superconducting cavity for the first time, the field emission often decreases abruptly: the cavity is said to "process" or "condition." There has been much progress in characterizing processed emitters at a microscopic level using techniques such as SEM, EDX, Auger, and AFM. These studies reveal that emitter processing is an explosive event that accompanies what we usually refer to as a "spark" or a "discharge," or the "electrical breakdown" of the

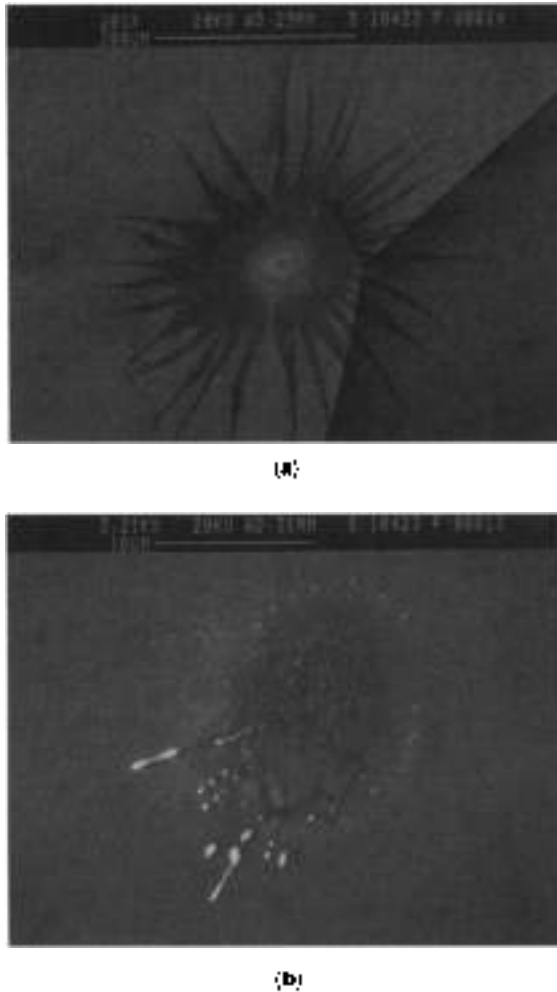


**Figure 8.** (a) SEM micrograph of field emitting particles. Note the cluster of small spherical balls which indicate that a part of the site melted. EDX analysis shows that the particles are stainless steel. Note also the jagged microgeometry of the particles believed responsible for field enhancement. (b) The melted cluster is expanded.

insulating vacuum (38). Figure 9 shows a typical SEM micrograph of an exploded emitting site (39).

To reach the highest accelerating fields, the highest thermal conductivity is essential to avoid thermal breakdown, and a high level of cleanliness is essential to avoid field emission. High-pressure water rinsing is a very successful cleaning technique to avoid field emission. In multicell structures with large surface area, there is always a significant probability that a few emitters will eventually find their way on to the cavity surface. There is also the danger of dust falling into cavities during installation of power coupling devices as well as during installing of the structure into the accelerator.

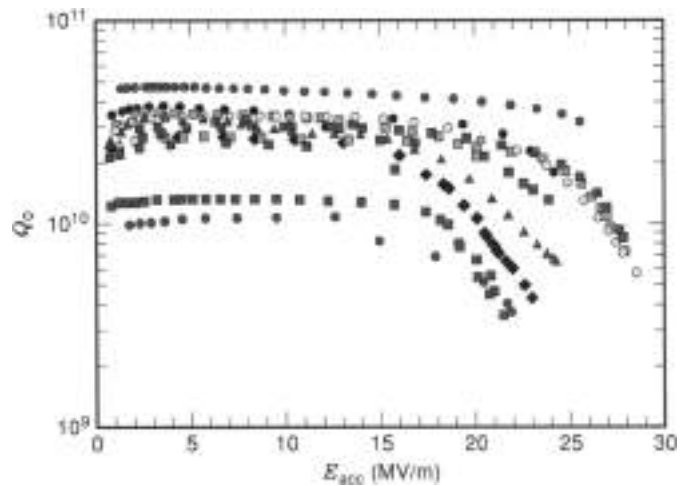
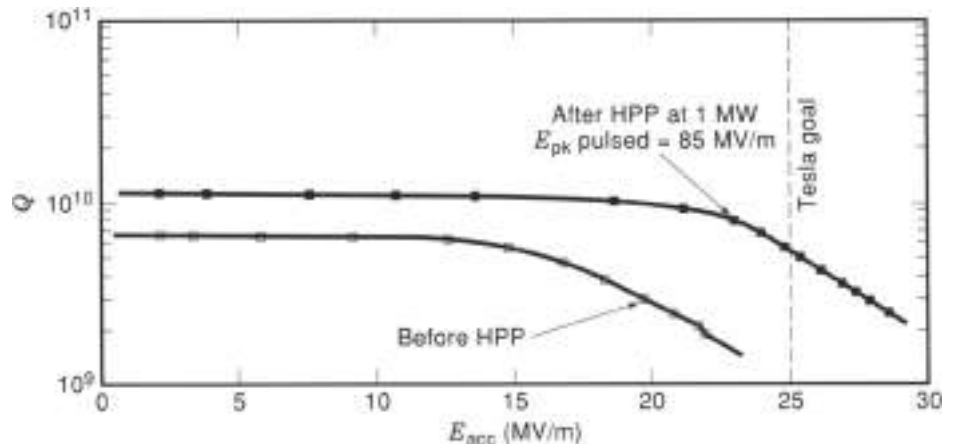
A technique that eliminates field emitters *in situ* is high pulsed power RF processing (HPP) (40). The essential idea is to raise the surface electric field at the emitter as high as possible, even if for a very short time (a few milliseconds). As the field rises, the emission current rises exponentially to the level at which melting, evaporation, gas evolution, plasma formation, and ultimately a microdischarge (RF spark) take place. The ensuing



**Figure 8.** SEM pictures of the processed site found at the location predicted via temperature maps: (a) low magnification, (b) high magnification of crater region within the starburst of (a). The mollen splashes in the crater region were found to contain indium, presumably from the indium wire seals used to make vacuum joints.

explosive event destroys the emitter. An important benefit of HPP is that the technique can be applied to recover cavities after their final installation. It can also be used to recover the performance of cavities which may be accidentally contaminated, as, for example, in a vacuum mishap. To achieve emission-free

**Figure 10.** Performance of a 5-cell, 1.3-GHz axial cavity improved by HPP. Before HPP, the maximum field was limited by heavy field emission to  $E_{acc} = 22$  MV/m. After applying 1 MW of power and reaching  $E_{pk} = 30$  MV/m in the pulsed mode, the field emission was processed away and  $E_{acc} = 28$  MV/m was possible in the CW mode.



**Figure 11.** High performance of several five- to 13-cell cavities achieved by high-pressure rinsing.

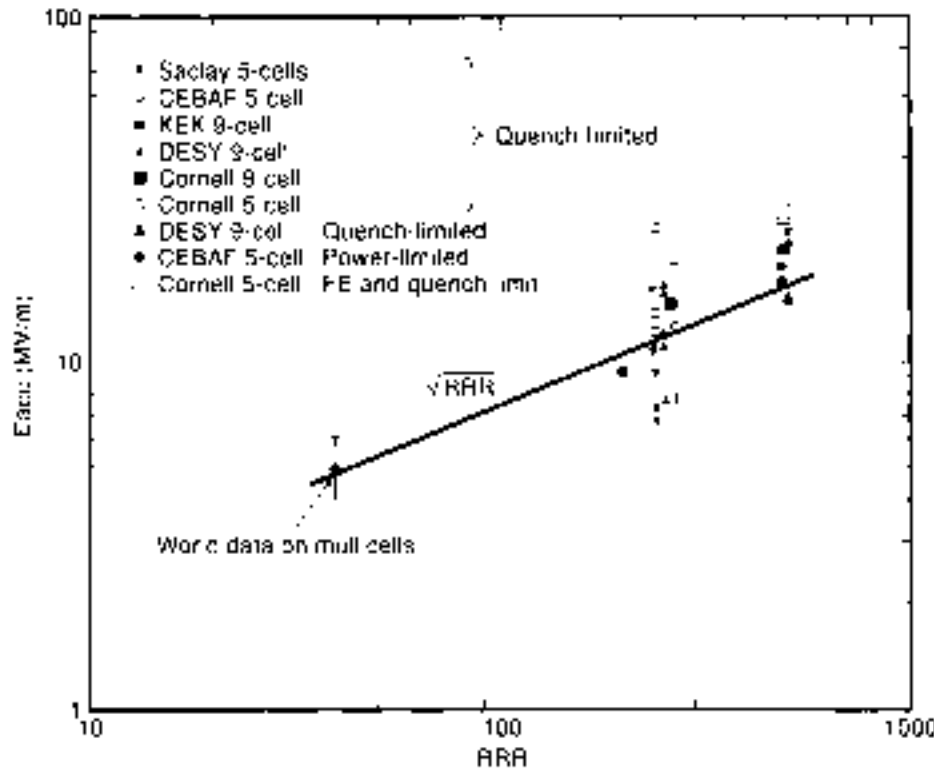
performance at a desired  $E_{acc}$ , processing must be carried out at  $\approx 2 \times E_{acc}$ . Figure 10 shows the improvement in performance achieved by HPP (41). [Recently, many 9-cell 1.3-GHz structures were prepared at DESY by using HPR techniques (37). A sample of their results is shown in Fig. 11. Occasionally it is possible to achieve field emission free performance, as shown by the best curve of Fig. 11.]

**Overcoming Thermal Breakdown**

The most effective cure for thermal breakdown caused by nm to sub-mm size defects is to (a) use better quality material that is free of such defects or (b) to raise the thermal conductivity of the medium so that remaining defects will be able to tolerate more power before driving the neighboring superconductor into the normal state (42). A simple analysis of the thermal breakdown shows that the maximum magnetic surface field is given by

$$H_{max} = \sqrt{\frac{4kTY_s}{\pi R_s} \dots} \quad \text{i.e., } H_{max} \propto \sqrt{k} \propto \sqrt{RHT} \quad (29)$$

Here  $k$  is the thermal conductivity,  $T_s$  the superconducting transition temperature,  $T_f$  is the bath temperature,  $a$  is the radius



**Figure 12.** A summary of the results of multi-cell cavities [39-43] showing the importance of high  $RRR$  coupled with emission reduction techniques such as HPU and HIR. The line shows  $\sqrt{RRR}$  dependence expected from the simple theory of thermal breakdown.

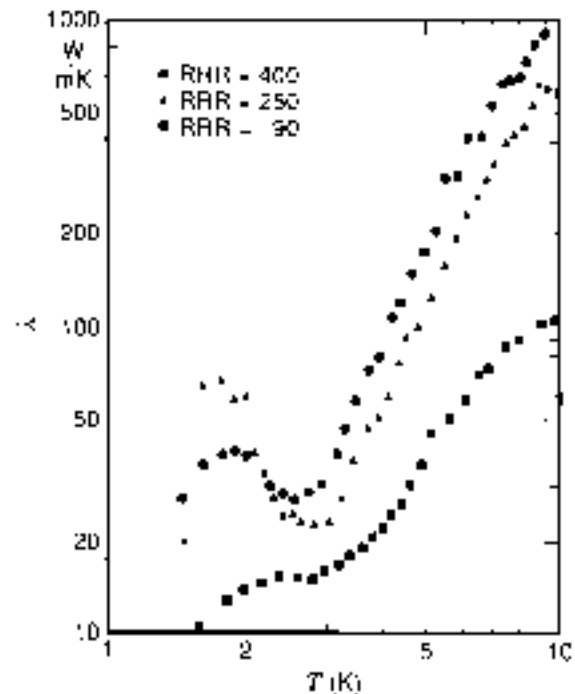
of the defect, and  $R_0$  is the surface resistance of the defect. This dependence on  $RRR$  is supported by detail numerical simulations of thermal breakdown, as well as by experiments on cavities made from Nb of different  $RRR$  (Fig. 12).

Figure 13 shows the thermal conductivity of three samples of niobium that have different histories of heat treatment [43]. The common feature of all three curves is the sharp drop below  $T_c = 9.2$  K, as more and more electrons condense into Cooper pairs. At the higher temperatures ( $4$  K  $< T < T_c$ ), a significant, though small, fraction of electrons is not frozen into Cooper pairs and can carry heat effectively, provided that the electron impurity scattering is low. Since the temperature in the neighborhood of the defect is between the bath temperature and  $T_c$ , the high-temperature thermal conductivity is the most important and has the strongest effect on thermal breakdown. The higher the  $RRR$ , the higher the thermal conductivity in this temperature range.

Below 4 K, as electrons condense into Cooper pair, electron-phonon scattering also decreases. As a result, the thermal conductivity from phonons begins to increase, leading to the phonon peak near 2 K. With decreasing temperature, the number of phonons decreases  $\propto T^3$ . The value of the phonon conductivity maximum is limited by phonon scattering from lattice imperfections, of which the grain boundary density is the most important. If the crystal grains of niobium are very large (e.g., because of annealing at high temperature), one observes a large phonon peak, as shown in the thermal conductivity behavior of the sample with  $RRR = 250$ , which was annealed at 1400 C. Since the phonon peak is at about 2 K, it does not help to thermally stabilize defects that heat up in the RF field.

The light, interstitially dissolved impurities have the strongest effect on the  $RRR$ . Among these, oxygen is dominant. The other interstitials are carbon, nitrogen, and hydrogen. Among the metallic impurities, tantalum is found in the

highest concentration (typically 500 ppm by weight) since all naturally occurring ores contain some tantalum. This impurity level is not a problem since tantalum is a substitutional impurity and does not substantially affect the electronic properties. However,  $T_c$  can become a problem if clustering occurs. The



**Figure 13.** Thermal conductivity ( $\kappa$ ) of niobium with  $RRR = 90$  (as received),  $RRR = 400$  after post-purification with yttrium, and  $RRR = 250$  after annealing the post purified sample for 6 hours at 1400 C. (Courtesy of Wuppertal.)

Table 3. Expected  $RRR$  for 1 ppm wt of Major Impurities

| Element          | $RRR$ |
|------------------|-------|
| H                | 2640  |
| N                | 4230  |
| C                | 4380  |
| O                | 5580  |
| Ta (1000 ppm wt) | 1140  |

Note that the effect of Ta is given in terms of 1000 ppm wt.

Nb used to fabricate a cavity has been checked, for example, by an eddy current scanning technique (44) to look for large defects such as T<sub>1</sub> clusters. Next in abundance are the higher-temperature, refractory elements, such as tungsten, zirconium, hafnium, and titanium, usually found at the level of 10–50 ppm wt. The electron-scattering effectiveness of the various impurities are shown in Table 3 in terms of their effect on the  $RRR$  (45).

To obtain the net  $RRR$ , one must add the resistance contributions for each impurity element in parallel. The contributions of the phonons is always present, so that the highest theoretical  $RRR$  for niobium is 35,000 (46). Experimentally, the highest  $RRR$  ever achieved in a niobium sample was 28,000 (47).

The most convenient method to obtain high-purity niobium for superconducting cavities is to remove the interstitials during the electron-beam melting stages of the ingot. Multiple melts and progressive improvements in the furnace chamber vacuum have led to a steady increase in the  $RRR$  of commercial niobium over the last decade from 30, typical of commonly available "reactor grade" niobium, to 300 (48). The  $RRR$  of commercially available Nb continues to improve. Recently, niobium sheet of  $RRR = 500–700$  became available from a Russian source (49).

If RF surface magnetic fields higher than 50 mT are desired on a consistently reproducible basis, the thermal conductivity of the niobium must be improved to  $RRR > 300$ . In one method called *post-purification*, the purity of the niobium is increased by solid-state gettering of oxygen using yttrium (50) or titanium (51) at high temperature. The foreign metal is vapor-deposited on the niobium surface. In the same step, the high temperature decreases the diffusion time of the oxygen in niobium. Over a few hours, oxygen is trapped in the deposited getter layer. If yttrium is used, the best temperature is 1200–1250 °C because both the vapor pressure of yttrium and the diffusion rate of oxygen in niobium are sufficiently high. If titanium is used, temperatures of 1350–1400 °C are required because of the lower vapor pressure of titanium. Typically during post-purification, the  $RRR$  improves by a factor of two in a few hours. An important disadvantage of the post-purification is that the yield strength of niobium falls substantially due to the high-temperature treatment. Also, titanium diffusion into the bulk along grain boundaries demands additional etching.

### Nb<sub>2</sub>Cu Cavities

As we mentioned, thin Nb films on higher-thermal-conductivity copper is another way to avoid thermal breakdown. The technique of sputter coating niobium has been developed by CERN for 350 MHz structures and applied successfully to hundreds of structures (25). In the most successful coating method to date, thin film deposition is carried out by cylindrical magnetron sputtering. Before the coating stage, the copper cavity

is degreased, chemically polished ( $\approx 20 \mu\text{m}$ ), rinsed with high purity, dust-free water and alcohol, and dried under clean laminar air flow. After bakeout of the copper cavity to reach a good vacuum, a typical coating time is 4 hours. The coating thickness is a few microns at a substrate temperature of 180–200 °C. The  $RRR$  of the deposited niobium serves as one of the monitors of film quality. The sputtering rate and substrate temperature are optimized to reach an  $RRR$  greater than 20. Note that the low  $RRR$  relative to bulk niobium is not a problem because the film is very thin. The rod-like grains of the niobium film are up to 1  $\mu\text{m}$  long and 10–150 nm in diameter. When studied with transmission electron microscopy, the individual grains show a high density of defects, consisting of dislocations and point defect agglomerates (52). The distance between two defects varies from 2 to 20 nm. The onset  $T_c$  of a deposited film is 9.6 K, but the transition width is larger than for bulk niobium (typically a few tenths of K). The large transition width (5 K in some cases) is indicative of poor film quality.

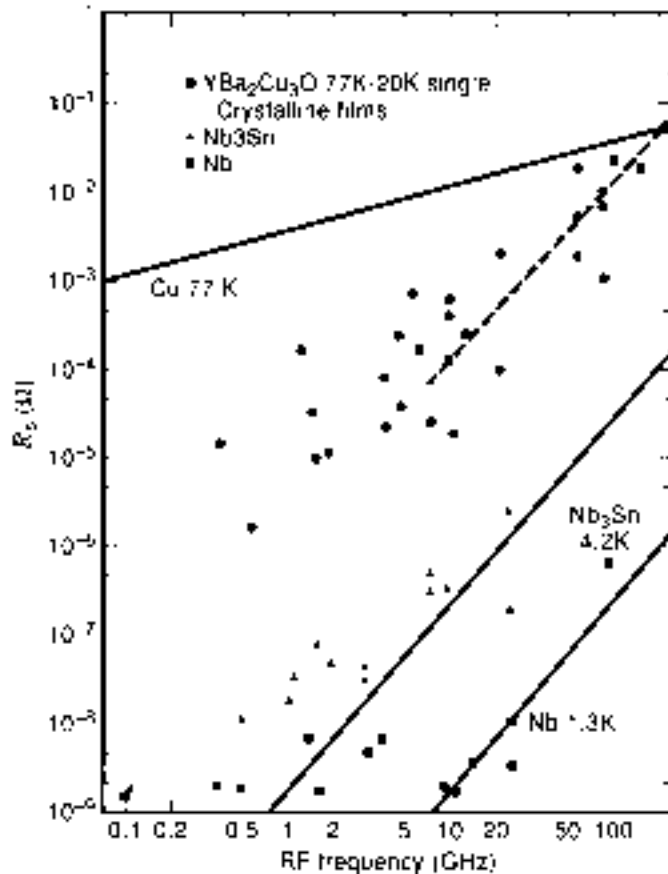
Although  $Q_0$  values  $\approx 10^{10}$  are obtained at low fields, the RF losses of Nb<sub>2</sub>Cu cavities increase steadily with field. This effect is attributed to intergrain losses in the niobium films, which become more severe at higher frequency. Recently (53), there is evidence to show that impurities buried in the films can also account for increased losses at high fields.

### Future Directions, New Materials

Based on the fundamental aspects, for a material to be useful in accelerators, the primary requirements are a high transition temperature and a high superheating critical magnetic field. Among the elemental superconductors, niobium has the highest  $T_c$ . While lead, coated on to a copper cavity, has been very useful in early studies and heavy-ion accelerator applications, the higher  $T_c$  and  $H_c$  has made niobium the more attractive choice. Technical considerations, such as ease of fabrication and the ability to achieve uniformly good material properties over a large surface area, have also proven favorable for niobium. The realm of superconducting compounds has been much less explored because of technical complexities that govern compound formation. In looking at candidates, such as Nb<sub>3</sub>Sn, Nb<sub>3</sub>N, and the new high-temperature superconductors (HTS), such as YBa<sub>2</sub>Cu<sub>3</sub>O<sub>7</sub>, it is important to select a material for which the desired compound phase is stable over a broad composition range. With this criterion, formulation of the compound may prove more tolerant to variations in experimental conditions, which in turn would make it possible to achieve the desired single phase over a large surface area. With a  $T_c$  of 16 K, Nb<sub>3</sub>Sn is the most successful compound explored to date (54). At low fields, residual resistance values comparable to niobium have been achieved. However, the maximum fields reached to date are far lower than those for sheet niobium cavities. The new HTS are even further from the performance level desired for application to accelerators. Figure 14 compares the measured RF surface resistance at low fields of several superconductors: HTS, Nb<sub>3</sub>Sn, and Nb (55). The surface resistance of Nb at 1.3 K, Nb<sub>3</sub>Sn at 4.2 K, and copper at 77 K are included for comparison.

Early enthusiasm over the remarkable strides made in the transition temperature of HTS are now tempered with difficulties in achieving useful properties, such as a high critical current density. The coherence lengths of the cuprates are very short: 17 Å within the copper-oxygen planes and 3 Å perpendicular to the planes, respectively. There is also a large anisotropy





**Figure 14.** Measured surface resistance of HTS compared to the same for Nb (1.3 K) and Nb<sub>3</sub>Sn (4.2 K). In the case of HTS the resistance is quoted at 77 K if it is already residual; if the resistance is still decreasing with temperature, the residual is obtained at 20 K. The solid lines show the calculated surface resistance of copper at 77 K, Nb<sub>3</sub>Sn at 4.2 K and Nb at 1.3 K.

of the magnetic and electrical properties between the *c* axis and the *ab* planes, with superior behavior when the current flow is in the *ab* plane. To produce good-quality HTS films, it is therefore necessary to orient the grains so that the *c* axis is normal to the RF surface everywhere. This restriction will be a significant challenge for realizing HTS in existing accelerating cavity shapes. It is also essential to have the right stoichiometry and oxygen content. Because of the short coherence length, transport properties are extremely sensitive to minute defects, such as grain boundaries and their associated imperfections. Decoupling of superconducting grains is believed to occur because the coherence lengths approach the scale of the grain boundary thickness, forming only weak links between individual grains. As a result, the intergrain critical current is two to three orders of magnitude lower than intragrain critical current. Even at a clean grain boundary, the scale of the disorder that exists from breaking up of a unit cell can exceed the coherence length, especially in the direction of the *c* axis (58).

## CONCLUSION

Even at the modest fraction of the ultimate potential, many attractive applications are now in place, and new ones are forthcoming. As our understanding of field-limiting mechanisms con-

tinues to improve, new techniques emerge to further advance gradients, such as high-purity niobium to raise the thermal conductivity, high-pressure rinsing to provide cleaner, field-emission-free surfaces, and high-pulsed-power processing to destroy residual emitters. The new techniques for bulk niobium cavities have demonstrated that gradients can be improved to between 20 and 30 MV/m in multicell structures. If such gradients can be reliably achieved exciting new applications on the horizon, such as the TeV electron-positron linear collider (57) or a multi-TeV muon collider (58).

## BIBLIOGRAPHY

1. J. Pascale, in H. Bunin (ed.), *Proc. 7th Workshop HF Supercond.*, Gif-sur-Yvette, France, CEA/SaClay 96-080, I, 1995, p. 173.
2. G. Geschonke, in B. Bunin (ed.), *Proc. 7th Workshop RF Supercond.*, Gif-sur-Yvette, France, CEA/SaClay 96-080/I, 1995, p. 143.
3. L. M. Bollinger, *Annu. Rev. Nucl. Particle Sci.*, **35**, 475, 1987.
4. H. Sundeim, *IEEE Trans. Nucl. Sci.*, **NS-32**: 3570, 1985.
5. S. Takeuchi, in K. W. Shepard (ed.), *Proc. 3rd Workshop RF Supercond.*, Argonne National Laboratory, Argonne, IL, ANL-PHY-89-1, 1988, p. 429.
6. H. Padamsee, K. Shepard, and R. Sundeim, *Annu. Rev. Nucl. Particle Sci.*, **43**: 635, 1993.
7. V. L. Ginsburg and L. D. Landau, *Zh. Eksperim. i Teor. Fizika*, **20**: 1064, 1950.
8. U. Klein, and D. Proch, in J. S. McCarthy and R. R. Whitney (eds.), *Proc. Conf. Future Possibilities Electric Accelerators*, Charlottesville, University of Virginia, P N1-17, 1979.
9. F. Kneise, H. Vinson, and J. Halbritter, *Nucl. Instrum. Meth.*, **188**: 660, 1981.
10. D. Meffat, in Y. Kajima (ed.), *Proc. 4th Workshop RF Supercond.*, KEK, Tsukuba, Japan, Rep. 89-21, 1990, p. 4-5.
11. J. Delayen and K. W. Shepard, *Appl. Phys. Lett.*, **57**(5): 514, 1990.
12. H. Padamsee, in *High Voltage Vacuum Insulation*, Latham (ed.), New York: Academic Press, 1995, p. 431.
13. J. Drinyen, in Y. Kajima (ed.), *Proceedings of the 4th Workshop on RF Superconductors*, KEK, Tsukuba, Japan, Rep. 89-21, 1990, p. 249.
14. U. Lauthner, U. van Rienen, and T. Weiland, DESY M-87-03, 1989.
15. R. Klein, DESY M-86-07, 1987.
16. J. Knobloch and H. Meier, *Rev. Sci. Instrum.*, **65**(11): 3721, 1994.
17. J. Bardeen, L. N. Cooper, and J. R. Schrieffer, *Phys. Rev.*, **108**: 1175, 1957.
18. D. C. Mattis and J. Bardeen, *Phys. Rev.*, **111**: 412, 1953.
19. J. P. Turney, Ph.D. thesis, Stanford University, 1967; see also J. M. Poore, and L. Martin (ed.), *Methods of Experimental Physics*, Vol. 11, 541, 1974.
20. J. Halbritter, *Z. Physik*, **238**: 466, 1970.
21. See the survey of low field residual resistance values in J. P. Turney, *Proc. Appl. Supercond. Conf.*, Annapolis, 1972, p. 621.
22. C. Volder et al., in E. H. Menke et al. (eds.), *Proc. 1992 European Particle Accelerator Conf.*, Editions Frontiers, 1992, p. 1295.
23. B. Bontz and R. Roeth, in D. Proch (ed.), *Proc. 5th Workshop RF Supercond.*, DESY, Hamburg, Germany, DESY M-92-01, 1991, p. 210.
24. J. Kirchgasler, in K. W. Shepard (ed.), *Proc. 3rd Workshop RF Supercond.*, Argonne National Laboratory, Argonne, IL, ANL-PHY-89-1, 1988, p. 503.
25. C. Benvenuti, in D. Proch (ed.), *Proc. 5th Workshop RF Supercond.*, DESY, Hamburg, Germany, DESY M-92-01, 1991, p. 189.

26. See Aschroff and Mermin, *Solid State Physics*, W. B. Saunders Company, 20, 1976.
27. L. P. Kadanoff and P. C. Martin, *Phys. Rev.* **124**, 670 (1961).
28. H. Padamsee, J. Knobloch, and T. Hays, *RF Supercond. Accelerators*, Wiley (1988).
29. P. Bernard et al., in H. Henken et al. (eds.), *Proc. 3rd Eur. Particle Accelerator Conf.*, Editions Frontières, 1992, p. 1269.
30. P. Kneisel, B. Lewis, and L. Turlington, in R. M. Sundelin (ed.), *Proc. 6th Workshop RF Superconductivity*, CERN, Newport News, VA, 1994, p. 624.
31. H. Padamsee, J. Tuckmantel, and W. Weingarten, *IEEE Trans. Magn.* **Mag-19** 1306, 1993.
32. R. H. Fowler and L. Nordheim, *Proc. R. Soc. London*, **A119**: 173, 1924.
33. J. Knobloch et al., in *Proc. 1995 Particle Accelerator Conf.*, Dallas, 1995, p. 1623.
34. M. Jimenez et al., *J. Phys. D: Appl. Phys.* **27**, p. 1036 (1994).
35. Q. S. Shu et al., in *IEEE Trans. Magn.* **25**, 1669, 1989.
36. N. S. Xu, in *High Voltage Vacuum Insulation*, Ed. B. V. Latham, p. 116, Academic Press 1995.
37. W. D. Moeller and M. Pakeker, in S. Myers et al. (ed.), *Proc. 5th European Particle Accelerator Conf.*, Barcelona, Spain, IUPP Publishing, Bristol, 1996, p. 2015.
38. D. Moffat et al., *Particle Accel.*, **40**: 85, 1992.
39. J. Gruber et al., *Nucl. Instrum. Meth. Phys. Rev.*, **A 350**: 582, 1994.
40. J. Gruber et al., *Nucl. Instrum. Meth. Phys. Rev.*, **A 360**: 572, 1994.
41. C. Crawford et al., *Particle Accel.*, **49**: 1, 1995.
42. H. Padamsee, *Proc. 1st Workshop RF Supercond.*, KFK, Karlsruhe, Germany, KFK-3019, 1980, p. 145.
43. G. Mueller, in K. W. Shepard (ed.), *Proc. 3rd Workshop RF Supercond.*, Argonne National Laboratories, Argonne, IL, ANL-PHY-88-1, 1986, p. 321.
44. W. Singer, in *Proc. of the 6th Workshop on RF Supercond.* Abano Terme, Italy, Ed. E. Palmieri 1997 in press.
45. K. K. Schulze, *J. Metals*, **33**: 33, 1981.
46. K. Schulze, Niobium, in H. Sturstedt (ed.), *Proc. Int. Symp.*, San Francisco, The Metallurgical Society of AIME, 1991, p. 163.
47. A. Gladun et al., *J. Low Temp. Phys.*, **27**: 373, 1977.
48. H. Padamsee, in R. M. Sundelin (ed.), *Proc. 6th Workshop RF Supercond.*, CERN, Newport News, VA, 1994, p. 515.
49. A. V. Elgutin, et al., in D. Proch (ed.), *Proc. 5th Workshop RF Supercond.*, DESY, Hamburg, Germany, DESY-M-92-01, 1991, pp. 354, 426.
50. H. Padamsee, in H. Lengeler (ed.), *Proc. 2nd Workshop RF Supercond.*, CERN, Geneva, Switzerland, CERN, 1994, p. 390.
51. P. Kneisel, *J. Low Temp. Metals*, **139**: 179, 1968.
52. C. Durand and W. Weingarten, *IEEE Trans. Appl. Supercond.* **6**: 1107, 1995.
53. S. Cultrini, in *Proc. 8th Workshop RF Supercond.* Abano Terme, Italy, Ed. E. Palmieri, 1997 in press.
54. M. Peiniger, in K. W. Shepard (ed.), *Proc. 3rd Workshop RF Supercond.*, Argonne National Laboratory, Argonne, IL, ANL-PHY-86-1, 1986, p. 503.
55. D. Busch et al., in R. M. Sundelin (ed.), *Proc. 6th Workshop RF Supercond.*, CERN, Newport News, VA, 1994, p. 173.
56. G. Mueller, in Y. Kojima (ed.), *Proc. 4th Workshop RF Supercond.*, KEK, Tsukuba, Japan, Rep. 89-21, 1990, p. 267.
57. B. Brinkmann, in *Proceedings of the 1995 Particle Accelerator Conference*, Vol. No. 96CH35845, 1995, p. 674.
58. *LHC: Collider. A Feasibility Study*, BNL-62503, 1996.
59. P. Wilson, in *Physics of High Energy Particle Accelerators: Feasibility*

*School*, 1981, no. 87, American Institute of Physics 1982.

HASAN PADAMSEE  
Cornell University

## SUPERCONDUCTING CYCLOTRONS AND COMPACT SYNCHROTRON LIGHT SOURCES

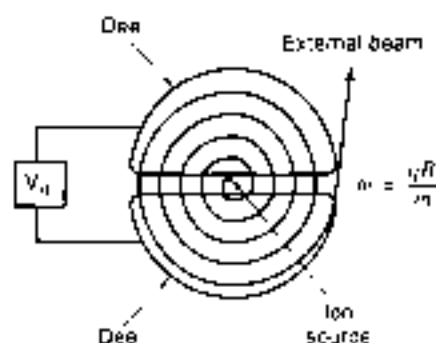
Superconducting magnets are the enabling technical component of the highest-energy particle accelerators ("atom-smashers") constructed and planned. These accelerators require both brute force and precision in the magnetic fields, which guide the particle beam on a roughly circular orbit, sometimes many kilometers in diameter, in order to return it to a radio-frequency accelerating structure for repeated energy boosts. The tera-electron volt (TeV) energy regime would not have been reached without a technology able to produce magnetic fields in the 5 T to 10 T range, accurate to about 1 part in 10,000, and economical enough to build in kilometer lengths. In this regime, superconducting magnets are the clear technology of choice.

The ability of superconducting magnets to produce strong, accurately shaped magnetic fields over large volumes also motivates their use in lower-energy accelerators, namely cyclotrons and electron synchrotrons, where normal magnets are also an option. Cyclotrons, because of their continuous beams, are usually the technology of choice for nuclear research in the hundred mega-electron volts per nucleon range. Such cyclotrons involve large magnets, in the 1000 ton range for magnets of room temperature design. Cyclotrons are also widely used in the commercial production of radionuclides and are used to a small degree for cancer therapy. The weight-reducing, cost-saving advantage of superconducting magnets (in the 100 ton range) has made them the design of choice in most recent nuclear physics projects. Superconducting magnets have also been used in a few situations for isotope production and for cancer therapy but have thus far not been widely adopted for either of these applications.

The short wavelength optical radiation emitted by electron synchrotrons has in the past two decades become a centrally important element in many research fields and in potentially important industrial applications. The need for sources compatible with on-line use in production facilities has led to the development of compact synchrotron light sources, whose compactness is enabled by the strong magnetic fields available from superconducting magnets.

## SUPERCONDUCTING CYCLOTRONS

The cyclotron was the earliest of the circular particle accelerators, earning a Nobel prize for E. O. Lawrence in 1936. First the cyclotron and then the synchro-cyclotron provided the beams for "energy-frontier" physics studies in the 1930s and 1940s. In the late 1950s, the intrinsic cost advantage of the ring geometry of the synchrotron (versus the pancake geometry of the cyclotron) shifted energy-frontier studies to the newly evolving synchrotron. However, because of a series of significant technical innovations (sector and spiral-ridge focusing and superconducting magnets), the cyclotron has retained



**Figure 1.** Conceptual drawing of the original cyclotron. The constant magnetic field is into the page. (Whereas the beam spirals outward in a cyclotron, in the synchrotron discussed later the beam stays at a constant radius and the energy increases as the magnetic field increases.)

several important regions of superiority and has not faded from the scene in the fashion of a number of other pioneering accelerator systems (e.g., the synchro-cyclotron and the betatron).

Figure 1 shows the key features of the classic cyclotron, the essence of which is quite eloquently summed up by a child's description: "they (the ions) start at the center and go round and round and come out at the edge." The "round and round" is the result of the transverse force exerted on the ions by the magnetic field, and the spiral path is the result of acceleration provided by a radiofrequency voltage applied to the "dees." The "cyclotron equation"  $\omega = qB/m = 2\pi f$  gives the orbital frequency  $f$  of a particle in a magnetic field  $B$ , where  $q$  and  $m$  are the charge and mass of the accelerated ion. The cyclotron equation reveals a gift of nature: taking  $q$ ,  $B$ , and  $m$  as constant, the orbital frequency  $f$  does not depend on the speed of the ion because the radius of the orbit increases just in proportion to the speed, so that the time per rotation is constant. Limits to the cyclotron process resulting from frequency errors from the relativistic increase in ion mass as the speed approaches the speed of light as well as deviations from the constant frequency ("isochronous" magnetic field shape in order to provide stabilizing "focusing" forces to keep ions on the design trajectory) were realized at an early time, and modifications were introduced. First the radio frequency (RF) was modulated to give the frequency modulated or synchro-cyclotron. Later strong and weak azimuthal regions in the magnetic field (hills and valleys or sectors) were introduced to provide both focusing to hold the ions near the design path and isochronism so that the ions remain in step with the accelerating voltage and can be accelerated to high energies limited only by the bending power of the magnetic field. The energy limit of the cyclotron thus evolved into a cost limit, and a very natural further evolution of the cyclotron developed. As superconductors lowered the cost of magnetic bending power, the superconducting cyclotron pushed the cost limit to higher energies.

The start-in-the-center-and-go-to-the-outside character of the cyclotron leads, in fact, to an exceptionally large benefit from the use of high magnetic fields. The field must be provided over the complete pole of area  $\pi R^2$ , where  $R$  is the maximum radius. The magnetic flux  $\Phi$  which determines the cross-sectional area of the return path steel, is field  $\times$  area (i.e.,  $\Phi = B\pi R^2$ ). Expressing the flux in terms of the bending power  $BR$  then gives

$$\Phi = BR\pi R^2 = \pi BR^3/B$$

(i.e., for fixed bending power, total magnetic flux varies as  $1/B$ ). The length of the magnetic return path also varies as  $1/B$ , giving  $1/B^2$  scaling for the total volume of steel. Superconducting cyclotrons are then lighter by a very large factor ( $\times 10$  to  $\times 20$ ) than room temperature cyclotrons with the same maximum bending power.

Table 1 lists the present superconducting cyclotrons of the world, including several now (1998) under construction. Accelerators of this type have assumed a centrally important role in medium-energy nuclear physics—the 5 T in bending power of the K1200 cyclotron at the National Superconducting Cyclotron Laboratory at Michigan State University, for example, makes it currently (1998) the world's highest-energy cyclotron, whereas its 300 ton weight is minor compared to the 4000 tons of the largest room temperature cyclotron. (Many beams from the K1200 reach energies in excess of 10 GeV/ion making it also the highest-energy continuous wave (CW) accelerator of any type.) The large superconducting cyclotrons at Catania, Chalk River, Groningen, and Texas A&M form, together with the K1200, a central core of overall world research capability in medium-energy nuclear physics.

The 1.5 T in bending power cancer therapy cyclotron at Detroit's Harper Hospital and the twelve 0.5 T in PET (positron emission tomography) cyclotrons manufactured by Oxford Instruments for leading medical centers in many countries have also demonstrated the feasibility and effectiveness of the superconducting cyclotron in situations where sophisticated technical support is much less available than in a major physics research center.

The coils used in most superconducting cyclotrons are cylindrically symmetric so that the very high "hoop stress" caused by the high field density and high current density are largely uniform with azimuth except for variations caused by the sector structure of the iron. Earlier superconducting cyclotrons used "cryogenically stable" helium-bath-cooled coils (i.e., coils with liquid helium in direct contact with each strand of conductor and with sufficient copper in the conductor so that if a section of conductor ceases to superconduct, the resistive power loss in the copper is nevertheless low enough for the conductor to recoil as a result of the presence of the liquid helium rather than go into a runaway thermal excursion). Two of the more recent superconducting cyclotrons, at Harper Hospital and at Groningen, in contrast, use an "intrinsically-stable," epoxy-potted winding where cooling is supplied only at the external surfaces of the coil. In this type of coil, it is very important to have a tightly packed, fully impregnated winding because sudden wire motion at the scale of 20  $\mu$ m to 40  $\mu$ m will release heat sufficient to drive the conductor normal.

The last two cyclotrons in the Table 1 list (Munich and Riken) are under construction; both attempt to carry superconducting technology an additional step by introducing noncircular coils, which preferentially strengthen the hills in the magnet compared to the valleys. Such a coil configuration is necessary if the cyclotron is to operate at particle velocities  $\beta v \approx 0.6$ , where axial focusing becomes a critical limiting parameter and open, low field regions are necessary between the hills ("separated sectors") to obtain sufficient axial focusing. The cyclotron at Riken is to produce beams up to 400 MeV/nucleon ( $\beta v = 0.7$ ); therefore, the technology of superconducting sector magnets is a central element of the project.

The superconducting cyclotron project at Munich also involves the feature of noncircular coils, but the dominating novel characteristic of this cyclotron is its superconducting RF system

Table 1. Superconducting Cyclotrons

| Name   | Location  | Purpose            | Magnet<br>Bending<br>Power<br>(T · m) | Final<br>Beam | Sector # | Type                |
|--------|---|--------------------|---------------------------------------|---------------|----------|---------------------|
| K500   | NSCL, Michigan State University,<br>East Lansing MI | Nuclear research   | 3.2                                   | 4992          | 3        | Compact             |
| TASSC  | Chalk River, Ontario Canada                         | Nuclear research   | 3.3                                   | 3995          | 4        | Compact             |
| K500   | Texas A&M, College Station TX                       | Nuclear research   | 3.2                                   | 4998          | 4        | Compact             |
| K1200  | NSCL, Michigan State University,<br>East Lansing MI | Nuclear research   | 4.8                                   | 4998          | 3        | Compact             |
| K100   | Harper Hospital, Detroit MI                         | Cancer therapy     | 1.4                                   | 4999          | 3        | Compact             |
| OSCAR  | Oxford Instruments, Oxford, UK                      | Isotope production | 0.6                                   | 6991          | 3        | Compact             |
| AGOR   | Groningen, Netherlands                              | Nuclear research   | 3.5                                   | 4961          | 3        | Compact             |
| K800   | Caracas, Italy                                      | Nuclear research   | 4.2                                   | 5985          | 3        | Compact             |
| Munich | Munich, Germany                                     | Prototype SRC      | 1.5                                   | -             | 12       | Separated<br>sector |
| SRC    | Riken, Japan  | Nuclear research   | 6.4                                   | -             | 6        | Separated<br>sector |

(in addition to superconducting sector magnets). Superconducting RF is not only very attractive conceptually because the RF system is normally the dominant power dissipation element in a cyclotron but also very difficult because superconducting RF currents will quench in the presence of dc magnetic fields of the level of about  $10^{-2}$  T. It is very difficult to achieve an almost field-free region in close proximity to magnets of the field strength needed in cyclotrons. The Munich project thus uses an unusual cyclotron structure known as a separated orbit cyclotron, which has been discussed in the literature for many years but previously never built. In such a cyclotron, sufficient RF voltage is provided to achieve some stipulated separation between the last and next-to-last orbits so that extraction of the beam is facilitated. In the Munich example, superconducting cavities provide sufficient energy gain per turn to give a 4 cm separation between final turns, and the cavity system has performed impressively. The magnet in such a cyclotron is, however, quite complicated, with a series of  $N$  separate magnets on every turn, where  $N$  is the sector number. This gives 240 magnets for the case of the 12 sector, 20 turn Munich prototype. The problem of appropriately adjusting the current in all these magnets without inducing quenches caused by misaligned beam, has lead to a prolonged, presently incomplete period of debugging of the Munich system.

#### APPLICATIONS AND CHARACTERISTICS OF SYNCHROTRON RADIATION

A relativistic charged particle beam deflected by a magnetic field emits electromagnetic radiation known as *synchrotron radiation*. The radiation from a light electron is more intense than from a heavier proton. Therefore practical synchrotron light sources are electron (or positron) storage rings. These machines are the brightest continuous sources of vacuum ultraviolet and X-ray radiation available.

The total power  $P$  radiated by an electron beam of energy  $E$  and current  $I$  when deflected in a magnetic field  $B$  is, in practical units,

$$P[\text{kW}] = 26.6E^4[\text{GeV}]^4/I[\text{A}]$$

A most useful attribute of synchrotron radiation is its high degree of natural collimation. The power is emitted within a cone of width  $1/\gamma$  around the direction of the electron beam, where  $\gamma$  is the energy of the electron divided by its rest mass (0.511 MeV). Hence for a typical source with energy 1 GeV, the synchrotron radiation is confined to a fan in the plane of the electron orbit, diverging with an opening angle of less than 1 mrad. Practical devices that store  $\sim 1$  GeV electron beams of  $\sim 100$  mA and deflect them in magnetic fields of  $\sim 1$  T therefore produce tens of kilowatts of collimated radiation, emanating from a line source that can be less than a millimeter in diameter. Large facilities for the production and use of synchrotron radiation have been constructed at research centers in most of the industrialized countries of the world. Compact machines allow use of this radiation to be extended to industrial applications.

Different applications of synchrotron radiation require flux from different parts of the electromagnetic spectrum. The spectrum emitted by a synchrotron source is characterized by the *critical energy*  $\epsilon_c$ . Half of the power radiated from the source is within a broad spectrum below the critical energy, and half is within an exponentially falling spectrum above  $\epsilon_c$ . The critical energy is related to the electron energy and deflecting field by

$$\epsilon_c[\text{keV}] = 0.665E^2[\text{GeV}]^2/B[\text{T}]$$

Thus, through proper selection of the electron beam energy and deflecting field strength, the output of a synchrotron source can be centered in the spectral range of interest.

The bend radius  $\rho$  of an electron in a magnetic field is given by

$$\rho[\text{m}] = E[\text{GeV}]/(0.3B[\text{T}])$$

Combining these expressions leads to an expression for the bend radius in terms of the critical energy and magnetic field

$$\rho[\text{m}] = 4.06\epsilon_c^{-1/2}[\text{keV}]/B^{3/2}[\text{T}]$$

This formula makes explicit the utility of choosing a magnet technology that can produce a high value of  $B$  in order to produce a compact synchrotron light source with output of kilowatt

X rays. Good-quality normal conducting electromagnets constructed from copper windings and iron pole pieces are limited to about 1.5 T by the saturation properties of the iron. Superconducting magnets using niobium-titanium alloy superconductors operated at liquid helium temperature can produce accurate fields several times larger, and reduce the radius of the synchrotron source by an order of magnitude.

Among the industrial and medical applications of synchrotron radiation for which dedicated machines have been proposed are coronary angiography, protein crystallography, and X-ray lithography. Coronary angiography with synchrotron radiation uses an abrupt change in the absorption of X rays by iodine at 33.17 keV. By injecting an iodine-rich contrast agent, exposing the subject in two beams that straddle this energy, and subtracting the images generated, vascular structure can be clearly visualized. The diffraction of X rays has long been used to determine crystal structure. The superior brightness of synchrotron sources compared to alternative X-ray sources have made crystallographers major users of large synchrotron research centers. Despite continuing interest, as of this writing no dedicated compact synchrotron has been constructed for either angiography or crystallography. However, several have been made to enable the development of X-ray lithography.

X-ray lithography is a candidate to become the successor to optical lithography for the patterning of semiconductors in the highest-density integrated circuits. For 30 years, advances in optical lithography's ability to print fine features has led to geometrically increasing circuit densities and chip performance, a phenomenon characterized by Moore's Law, which observes that the number of transistors that can be put on a chip has been doubling about every 18 months. However, as the size of circuit features is reduced to values comparable to the wavelength of the ultraviolet radiation used to print the pattern (i.e., to 248 nm and below), diffraction effects limit the resolution, and the challenge of maintaining this trend becomes more difficult. X rays, with energy near 1 keV, wavelength near 1 nm, diffract much less, and have been used to fabricate transistors with dimensions of 100 nm and below. Lithography of complex, high-density circuits using X rays has progressed to the pilot line stage, using compact synchrotron X-ray sources developed for this purpose. A related technology using higher-energy X rays to fabricate micromachine parts also has much promise.

#### TECHNICAL DESCRIPTION OF A COMPACT SYNCHROTRON

Compact synchrotrons have the same major subsystems and are designed according to the same principles as larger synchrotrons, although innovation has occurred in the design of the superconducting magnets for compact light sources. Figure 2 shows a photograph of the Helios 1 compact synchrotron manufactured by Oxford Instruments. Other machines include Aurora by Sumitomo Heavy Industries, COSY by COSY Microtech, MELCO by Mitsubishi, and SuperALIS by NTT. The major subsystems on all these machines are:

- an RF accelerating structure,
- the magnet system,
- an injector and pulsed injection magnets on the ring,



**Figure 2.** The first Helios compact synchrotron, built by Oxford Instruments. The two superconducting magnets can be seen at either end of the structure. X-ray beams emerge from the ports on the sides of the magnets.

- a vacuum system,
- beam diagnostics and controls, and
- auxiliary power supplies and utilities.

The accelerating structure is a resonant radiofrequency cavity producing a longitudinal electric field imparting energy to the beam particles with each revolution. The energy gain serves both to raise the average beam energy during the acceleration cycle, and to compensate for energy losses to synchrotron radiation. Synchrotron operation is based on the principle of *phase stability*. Particles with small deviations from the central beam energy follow slightly different orbits and arrive at the cavity at slightly different times. They automatically receive more or less energy according to their phase. The result is effectively a restoring force causing the beam to coalesce into bunches, with individual particles executing longitudinal *synchrotron oscillations* about the center of the bunch. Phase stability puts a constraint on the circumference of the orbit, which must be an integer multiple of the radiofrequency wavelength. The choice of radiofrequency is largely dictated by convenient RF sources. Klystron-based sources at 500 MHz have been used, as have solid-state sources at 50 MHz.

The most common compact light source layout has been a racetrack using a pair of superconducting magnets each producing a vertical magnetic field to deflect the beam by 180°. This configuration allows other machine functions (e.g., the RF structure, optical elements, pulsed magnets for injection, and vacuum components) to be distributed in the two straight sections. It still takes advantage of the small bend radii that can be achieved with a superconducting magnet to produce a device that can be factory assembled and delivered to a user as a complete unit. The magnetic field is increased during the acceleration cycle because the beam maintains a constant orbit in a synchrotron, unlike the outward spiraling orbit of the beam in a cyclotron. In the case of the Helios system, the superconducting magnets produce a field of 4.5 T at full energy, which allows a 700 MeV beam to be bent into a semicircle of radius 51 cm. The entire device, exclusive of the injector and auxiliary systems, weighs about 20 tons and comfortably fits on the back of a truck.

In order to allow a beam particle with small deviations from the ideal orbit to survive for perhaps  $10^{11}$  revolutions without being lost, some form of focusing is required. All compact synchrotrons built save one have been *strong focusing, alternating gradient machines*. (The Auenra 1 compact light source built by Sumitomo Heavy Industries was a weak focusing machine.) Focusing in one of the transverse planes can be accomplished by a magnetic field gradient. However, if a magnetic field gradient focuses in the horizontal plane, it inevitably defocuses in the vertical plane, and vice versa. The alternating gradient concept uses the principle that pairs of lenses acting together, one focusing and one defocusing, have a net focusing effect. The gradients necessary for producing stable orbits in a compact synchrotron can be produced by independent quadrupole lenses, as in a separated function machine, or by gradients built into the deflecting field, as in a combined function machine. The Helios source shown uses high-precision placement of the superconducting wires is necessary to achieve the field quality required.

Most compact light sources have been designed with low-energy injectors. The purpose of the injector is to deliver to the ring an electron beam that is already relativistic and that is high enough in energy to limit the dynamic range required of the synchrotron's magnet system. Because the injection process is inefficient and produces radiation from particles that are not captured by the synchrotron, the energy of injection also determines the thickness of shielding required. Both linear accelerators and racetrack microtrons are in use as injectors for compact light sources. Rings designed for operation at 700 MeV or 1 GeV have successfully used injection energies from 50 MeV to 200 MeV. Pulsed magnets within the ring are required to deflect the incoming beam from the injector onto a stable orbit.

Ultra-high vacuum technology is required in a compact synchrotron because the lifetime of the beam can otherwise be limited by losses from collisions between the electrons and residual gas atoms, especially those liberated by collisions between synchrotron radiation and the vacuum chamber walls. The presence of a cryogenic system for the superconducting magnets makes cryopumping a natural choice. The Helios source has achieved beam lifetimes in excess of 50 h with 200 mA of circulating electrons. Because the injection and acceleration cycle require only about 10 min, a high utilization factor is possible.

## BIBLIOGRAPHY

1. J. Cornell, *Cyclotrons and Their Applications*, Singapore: World Scientific, 1996.
2. F. M. Russell, A fixed frequency, fixed field, high energy accelerator, *Nuclear Instrum. Methods*, 28: 229, 1963.
3. E. Wehrexer, *Compact Synchrotron Light Sources*, Singapore, World Scientific, 1996.
4. J. M. Warlaumont et al., *IBM J Res Dev*, 37 (3): 269-446, 1993.

DAVID E. ANDREWS  
Oxford Instruments, Inc.  
HENRY BLOSSBERG  
Michigan State University

## SUPERCONDUCTING FAULT CURRENT LIMITERS

The control of fault currents in electrical transmission and distribution networks is a major challenge. Short-circuit current levels are considerably larger than the rated current. All electrical equipment exposed to the short-circuit current must be designed to withstand in particular the mechanical forces under fault conditions, which are generally proportional to the square of the current. Circuit breakers that isolate the fault must be capable of absorbing the fault energy, as well as be proportional to the square of the fault current at the instant of breaker operation. Consider an infinite bus feeding a load through a transmission line and transformer with 8% combined impedance. In the event of a short circuit across the load, the short-circuit current is limited only by the line and transformer impedance. The steady-state short-circuit current will be 12.5 times larger than the rated current, and the initial transient current can reach a value about twice the peak steady-state fault current, resulting in short-circuit electromagnetic forces that are 625 times higher than the peak forces under rated conditions. To survive such a fault, all circuit components must be designed to withstand the peak forces.

A fault current limiter restricts the current fed to the fault to a lower value than with the limiter absent. Unlike mechanical breakers, where the mechanical inertia in the mechanism delays any effect on the fault current during the first few cycles, fault current limiters insert an impedance already in the first half cycle, thus limiting the peak current transient of the short-circuit event. Then all the network electrical equipment, such as lines, transformers, and circuit breakers, could be designed for smaller fault current amplitudes, which would result in a reduction of material and cost. In many transmission and distribution networks, load growth, distributed generation, and interconnections have produced increased fault levels that exceed the capability of the installed circuit breaker base to interrupt prospective fault currents. In some situations, the fault levels exceed the available ratings for circuit breaker replacement or upgrading, and economically or operationally attractive interconnections cannot be made. With the introduction of a fault current limiter, a system could use breakers with a smaller interrupting rating than the system fault level. Thus electrical grids can be interconnected to produce stiffer, more reliable systems without upgrading existing breakers to higher interruption ratings.

No fault current limiter technology is widely accepted or used today. Superconducting fault current limiters (SFCL) form one class of fault limiting technology under active development. After several technical generations of development, niobium-based low temperature superconductor (LTS) technology is mature enough for consideration by the electric utility industry. However, market penetration of LTS devices has been impeded by high cost and the use of liquid helium (LHe) cooling. The 1986 discovery of a new class of ceramic high temperature superconductors (HTS) (1) and the subsequent active development of HTS raised hopes for inexpensive superconductor (SC) applications due to cheaper HTS conductor and significantly reduced refrigeration requirements. The discovery stimulated reevaluation of all previously proposed SC applications, including several SFCL embodiments, and prompted many innovative SFCL concepts and designs.

Many SFCL programs are active, ranging from the proof-of-concept stage all the way to precommercial prototype

development. These programs are pursuing several distinct types of SFCLs, and it is not yet clear which approaches will ultimately be both technically and economically viable (2). At present there is no off-the-shelf SFCL product available in the marketplace, and even the most advanced of the several SFCL commercialization programs are not expected to be ready to offer utility-qualified products for some years to come. Here we outline the key issues of SFCL technology and application to give a basis for evaluating present and future developments.

Almost all SFCL development is targeted to power frequency grid applications. Conceptual studies of utility generation and transmission systems with a large proportion of SC generators, transformers, and transmission cable have looked to SFCLs to limit the higher fault levels inherent in the lower fault impedances of the SC system components. SFCLs would also be useful in high voltage direct current (HVDC) transmissions, where currently the challenges of dc fault current interruption have not been solved. More specialized applications can be considered in high-power radio-frequency systems, magnet energization circuits associated with high-energy physics, and aerospace and shipboard systems. This article will focus on the electric utility, power-frequency type of application. Several surveys have been carried out of the applicability of SFCL technology for electric utility needs (3,4), and some studies have projected SFCL costs (5).

The most important characteristics for the user is the limiting current level. As discussed later, for most utility grid applications, a fault level reduction of 10% to 50% is needed to coordinate with existing protection, whereas for industrial end users more aggressive reduction may be appropriate. Other parameters are the trigger level for the initiation of limiting and the reset time for the limiter to return after limiting operation.

Most SFCLs may be classified into (1) quenching SFCLs, which rely on the superconducting-normal phase transition to switch from a low to a high impedance; (2) desaturating SFCLs where ferromagnetic material switches between saturated and unsaturated states and the SC biases the quiescent operating point to the low-impedance unsaturated state; and (3) semiconductor SFCLs where the semiconducting elements switch to a high-impedance operating point from a low-impedance point defined by bias applied from the SC. Superconductivity is essential only to the quenching SFCL, but it is useful in desaturating and semiconductor fault limiters to reduce the bus system losses and improve system compactness.

**ELECTRIC UTILITY GRID APPLICATION**

To be useful, the SFCL must operate quickly enough to limit the first half-cycle of the fault so that the interrupter peak current duty is reduced. This imposes a fast SFCL triggering requirement. Most concepts are self-triggering, with the instantaneous fault current driving the SFCL transition to high impedance. Otherwise, external triggering driven by state-of-the-art relaying (detection time  $< \frac{1}{2}$  cycle) is used. A tunable self-triggering level is a very attractive SFCL feature. After SFCL triggering, a series power circuit breaker interrupts the fault at one of the line current zeroes, with an operating delay from  $1\frac{1}{2}$  to 5 cycles. The SFCL must endure the limited fault current until the maximum backup breaker opening time, as determined by the utility's protection coordination. Triggering selectivity should

also be considered to ensure that the SFCL does not trigger on transformer inrush and motor-starting transients.

We can distinguish between a *trimming* SFCL (where the effective SFCL impedance  $Z_L < Z_s$ , the system source impedance), which reduces the prospective fault current by 0-50%, and a *dominant* SFCL ( $Z_L > Z_s$ ), which achieves reductions significantly over 50%. In existing networks that have incrementally grown to high fault levels, utilities require only trimming limiters so that the existing selective relay equipment can be retained. In fact, dominant limiters would require major reconfiguration of system protection coordination, which is always unattractive. Limiting the short-circuit current to values three to five times the rated current is a typical requirement in many cases. In some specific applications, a dominant SFCL, such as to protect a complete new installation for an industrial customer, may be used.

The possible locations for a fault current limiter in a substation are shown in Fig. 1. Two transformers feed two separate buses, providing power to several feeders. Interconnecting the two buses by closing the bus tie breaker will improve load voltage regulation and overall system reliability, at the cost of increased fault levels, because both transformers can now feed into a fault. Fault current limiters are installed in series with all breakers. In the case of a bus fault (shown as FB), the fault current limiter in series with the B transformer secondary side breaker limits the short-circuit current. In the case where the bus tie breaker is closed, the fault current limiter in the bus tie reduces the short-circuit current contribution from the A transformer. Should a feeder experience a fault (shown as FF), the limiter in the feeder circuit reduces the short-circuit current. If the A transformer was a later installation and the parallel connection of both buses is desirable, the installation of the fault current limiter in the bus tie connection could enable continued safe operation of the system with the original breakers.

For an SFCL device to be accepted in an electric utility system, it must meet general safety, reliability, availability, basic

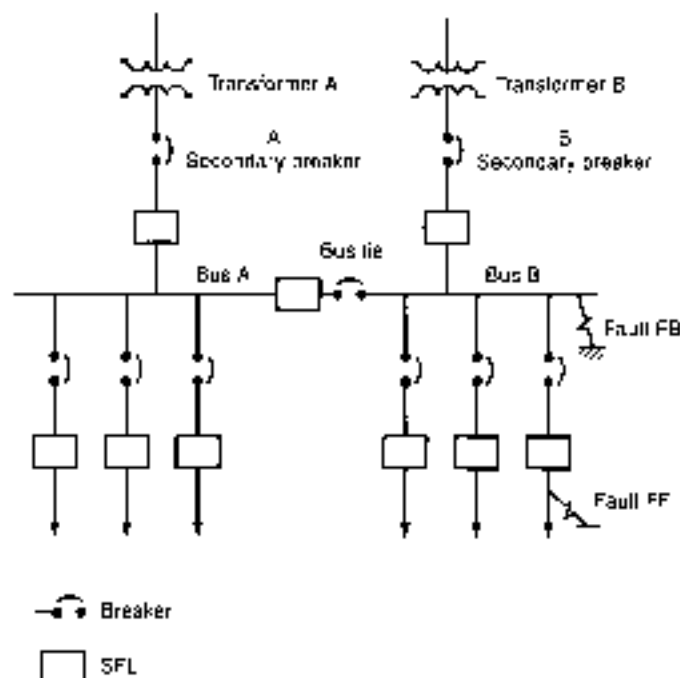


Figure 1. Superconducting fault limiter application in a substation.

insulation impulse level (BIL), and overload requirements and ratings comparable to other utility components like transformers or circuit breakers. Currently there are no standards for SFCL application to utility systems. Pertinent sections of standards for other system components can be adapted to meet the fault current limiter design requirements. The efficiency should be above 99%. The SFCL must operate for reclosure duty and must be fail safe. Under all operating conditions, stray field levels accessible to personnel should be no higher than levels generated by other equipment. The exposed field levels must be limited to avoid personnel harm and projectable effects on ferromagnetic tools (and also typical debris such as disassembled nuts and bolts).

In addition, the utility user must be reassured about new issues raised by the introduction of SC technology. Operational experience with utility system integration of SC is very limited. A 30 MJ/10 MW superconducting magnetic energy storage (SMES) system was installed in 1963–1984. Failures in the LHe refrigeration took the system down, but the SC coil did not fail. The operation of a small number of commercial LTS-LHe 1–3 MJ “micro-SMES” units in 480 V uninterruptible power supply systems in industrial power system environments since 1993 has demonstrated reliability matching battery systems. In addition, reliable cryogenic refrigerators have been developed for commercial application in medical magnetic resonance imaging (MRI). All of the above experience points to the *technical feasibility* for an LTS-LHe SC system to achieve the same reliability as a conventional system in an electric utility environment, but such reliability has not yet been demonstrated. The ongoing testing of a 1.2 MVA HTS-liquid nitrogen (LN<sub>2</sub>) SFCL system at a Swiss hydroelectric power station (6), followed by the planned 1998 demonstration of a 21 MVA HTS SFCL system at a distribution substation (2), will provide much-needed utility environment operational experience and reliability data, particularly for HTS-LN<sub>2</sub> systems.

The cryogenic system of the SFCL may present the utility with a new unfamiliar consumable, the liquid cryogen. It is very unlikely that LHe will be acceptable, because of its cost and specialized handling procedures. Liquid nitrogen (LN<sub>2</sub>) might be accepted because it is much cheaper, widely available, and already in utility use in some transmission cable applications, where it is used to freeze the oil dielectric in the course of maintenance procedures. Closed cryogenic systems are the most attractive, where limited amounts of cryogen boiloff are reliquified, or where no liquid cryogen is used at all.

## QUENCHING TYPES

The size of the SC element, which changes from a very low impedance state to a finite resistance to quench, is based on the material properties of critical current  $I_c$  and normal resistivity  $\rho_n$ , and the system requirements of crest trigger current  $I_n$  and normal state SFCL resistance  $R$ , to give a conductor cross sectional area of  $I_n/\bar{J}$ , length  $R/\rho_n/\bar{J}$ , and volume  $R I_n^2/\bar{J}^2 \rho_n$ . In general, high  $\bar{J}^2/\rho_n$  is desirable to minimize superconductor requirement.

The fault transient is divided into three time regimes: a pre-limit regime (period I) where  $I < I_n$  and the limiter is not yet active, a resistance growth regime (period II) where the SFCL is in a partially resistive state below the critical temperature  $T_c$  and the limiter effective resistance  $R_e = V_L/I < V_L$  is limiter volt-

age is rising from 0 to  $I_n$ , and a full-resistance regime (period III) where the SFCL temperature is above  $T_c$  and all superconductivity has quenched. Period I behavior depends on  $V_s$ ,  $L$ , and  $I_n$  alone.

The electrical performance in period III is analyzed in terms of an  $RL$  circuit transient, and the average temperature transient follows from the resulting energy absorption  $R I^2 dt$  divided by the SFCL mean heat capacity  $C$  (including all material coupled by thermal diffusion to the SC element over the limiting time). The behavior in period II is more complex and is best solved numerically (7). It is useful to describe the SFCL as weak or strong, based on whether  $R$  is less or more than  $V_s/I_n$ . With a strong limiter, the limiter voltage will overcome the source voltage during period II at which time the fault current will peak. It is still possible that there will be a subsequent higher current peak due to the power-frequency variation of  $V_s$ . It is also useful to describe the SFCL resistance growth as fast or slow, depending on whether the limiting current at the end of period II is more or less than the quasisubequilibrium fault current  $I_q = V_s/R$ . The fast weak SFCL will not force a current peak in period II, because always  $V_L < V_s$ . The slow weak SFCL will force a current peak in period II because period II extends so long that the limiter current grows until the voltage across  $L$  reverses and limiter  $dI/dt$  changes sign. It should be noted that the relative strength and speed of the SFCL depend on the power system parameters as well as the SFCL parameters  $R$ ,  $C$ ,  $I_n$ , and the difference between critical and quiescent temperatures.

Four simplified cases are considered in Fig. 2 with low (weak) and high (strong)  $R$  and slow and fast  $dI/dt$ . The fast strong limiter gives the lowest peak let-through current and  $dI/dt$  to interruption, but it also will produce SFCL terminal overvoltage. Weakening a fast limiter will reduce the overvoltage at the cost of higher let-through current and  $dI/dt$ . Slowing down a fast strong limiter will increase the let-through current and  $dI/dt$  and may increase the overvoltage. Clearly SFCL strength and speed are both desirable for the primary function of current limiting, but they will produce overvoltage. If only weak limiting is desired, the limiter overvoltage can be controlled by speed adjustment, and in fact all overvoltages can be suppressed by keeping speed below a particular threshold.

Period III continues until either the primary or backup circuit breaker interrupts the fault. Here the normal state SC element carries a large current and dissipates high power, which dominates total SFCL heating. For example, a resistive quenching SFCL that trips the fault level of a 345 kV transmission system (with  $X/R$  ratio 15) by 10% from 70 kA to 63 kA will dissipate 4.8 GW per quenching SFCL phase during period III. The main challenge in quenching SFCLs is SC energy management under fault conditions.

Quenching SFCLs are in-line (line current flows in the SC element) or transformer-coupled. In all in-line quenching SFCLs, SC heating (and thus the amount of SC and cryogenic system size and capacity) is reduced by a parallel fault current bypass resistor (R), inductor (L) or varistor (VD) with lower effective impedance than  $R$ . The SFCL-limiting impedance is of course reduced, but the symmetrical SC current is also reduced. The bypass element is large with significant (but not prohibitive) cost. Both SC and conventional (copper conductor) bypass inductors have been proposed. There is little to recommend the SC variant because it would need to be rated to the fault current level for both currents and forces. A conventional reactor



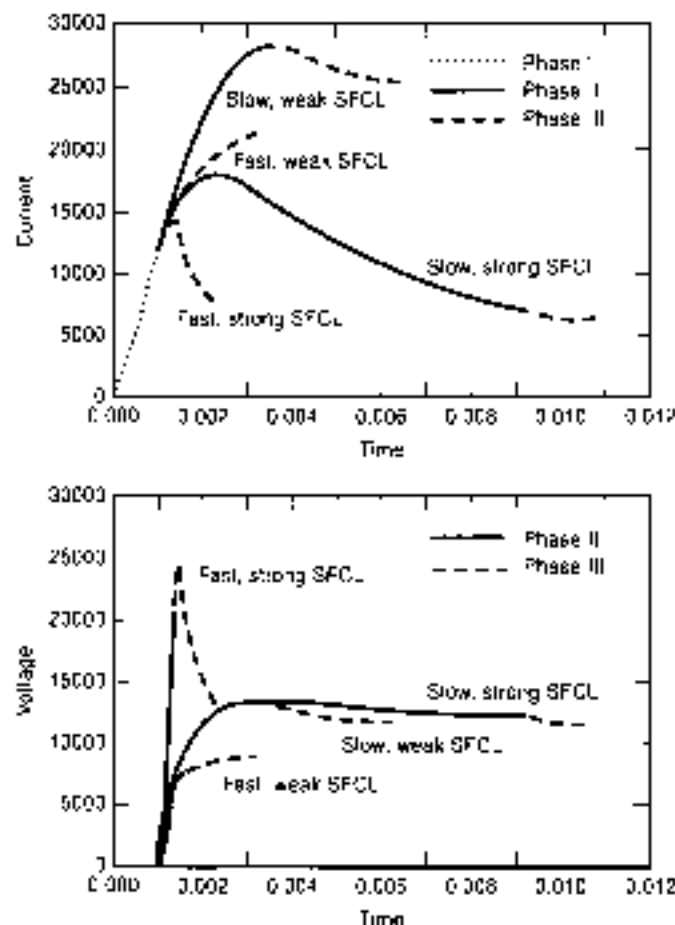


Figure 2. Four types of quenching SFCL performance depending on limiter strength and speed. Direct source voltage is applied; hence, power frequency effects are absent.

could be quite modestly sized to achieve the necessary transient duty, and a bypass resistor would also be reasonably sized. In weak current limiting applications, when designing for the same peak SC element temperature limit, significantly more SC material will be required for the resistive bypass design. A bypass varistor would have the additional benefit of limiting the SFCL transient overvoltage, but it has a high-energy capability requirement.

All in-line SFCLs require cryogenic current leads to connect the line current to the cryogenic system. These leads must be thermally rated for the SC element fault transient current and will become the dominant heat leak into the cryogenic system. Respectably rated distribution-level in-line SFCLs have been built with both inductive (91 (6.6 kV/1.5 kA) and resistive (8) bypass (7.2 kV/1 kA). Both used relatively short lengths (on the order of 100 m) of expensive power frequency capable ultra-fine filament NbTi superconductor with high-resistivity cupronickel stabilizer matrix, operating with LHe cryogen at 4 K. Important design issues include ac derating of multistranded conductor from single-strand dc values of  $I_c$ , unutilization of the conductors to avoid spontaneous quench due to conductor movement, and use of cryocoolers to produce zero boil-off systems. Heat losses from induced currents in the metallic cryostat and ac losses in nonstranded current leads and leads to SC wire joints were a problem. It is expected that limiting will deposit several megajoules in the cryostat, leading to sig-

nificant LHe vaporization, which must be vented or captured and reliquefied. Open-cycle operation (venting the boiloff) will drive the LHe replenishment requirement, whereas the reliable reliquefaction required for a closed system is technically challenging.

HTS materials have been applied to in-line SFCLs, but at a more modest scale than the NbTi devices because of the difficulty in producing transposed multifilamentary conductors for low ac loss. An advantage of HTS materials is the relatively high  $\rho_c$ , which can exceed  $10 \mu\Omega\text{m}$ . Modeling studies (10) suggest that  $J_c$  of at least  $10 \text{ } 100 \text{ MA/m}^2$  ( $1 \text{ } 10 \text{ kA/cm}^2$ ) is required. The high  $\rho_c$  drives designs to relatively low SC volumes and high peak temperatures. Extra heat capacity may be added (11) by using thin (0.2–2  $\mu\text{m}$ ) HTS films deposited onto sapphire or YSZ (yttria-stabilized zirconia) substrates, which provide much-needed heat capacity. It remains to be seen whether the preferred in-line HTS element geometry will be wound multifilament or thin film.

In the *shielding* SFCL, the SC is coupled to the line current by transformer action (Fig. 3(a)) and forms a shorted winding. A preferred geometry is a single turn of HTS material; Bi-2212 material is commercially available in tubular sizes up to 40 cm diameter (12). The thermal transient can be analyzed by the previously derived equations after the device circuit is transformed by the turns ratio to a T equivalent circuit. The insertion impedance will be the leakage inductance between the line and SC windings (13), and the limiting impedance will be the magnetizing inductance in parallel with the SC winding normal  $R$ . Desirable features of the shielding SFCL include the isolation of line current and potential from the SC element and no requirement for cryostat current leads. Several groups (14–16) are actively pursuing this promising approach and 6 months of operating experience protecting a hydroelectric plant auxiliary transformer circuit has been reported for a 1.2 MVA 10.5 kV Bi-2212 LN<sub>2</sub>-cooled shielding SFCL (6). Design challenges include the reduction of cryostat circulating currents by the use of nonmetallic cryostat, thus minimizing circulating currents in the metalized multilayer insulation; the mechanical support of transient bursting or compressive forces on the HTS shield, the amelioration of transient thermal inhomogeneity; and optimization of recool time. Most designs utilize nonlaminated ferromagnetic cores. Aircore devices will have larger leakage impedance. Control of shielding SFCL trigger level has been proposed by using a tertiary winding (16).

A quenching-type SFCL operation may be self-initiated by the fault current as discussed previously, or some quenching mechanism (17) may be externally triggered by a fault detection relay. Digital microprocessor-based relaying is now capable of giving subcycle fault detection, so that detector delay is no longer a barrier. Several trigger methods have been proposed, including discharging a capacitor to give a magnetic field pulse from an auxiliary coil, discharging a capacitor to add an extra overcurrent component, triggering heaters on the SC, or sending a warm cryogen burst into the cryostat. Currently, most SFCL development is aimed at self-triggering using fault overcurrent alone. However, in later generations, commercial SFCLs will be differentiated by the triggering system to maintain trigger level control independent of SC property variations and cryogenic operating point variability and to give flexibility in setting the SFCL trigger level.

Under fault limiting, some parts of the SC element will enter the lossy period II regime before others as a result of local

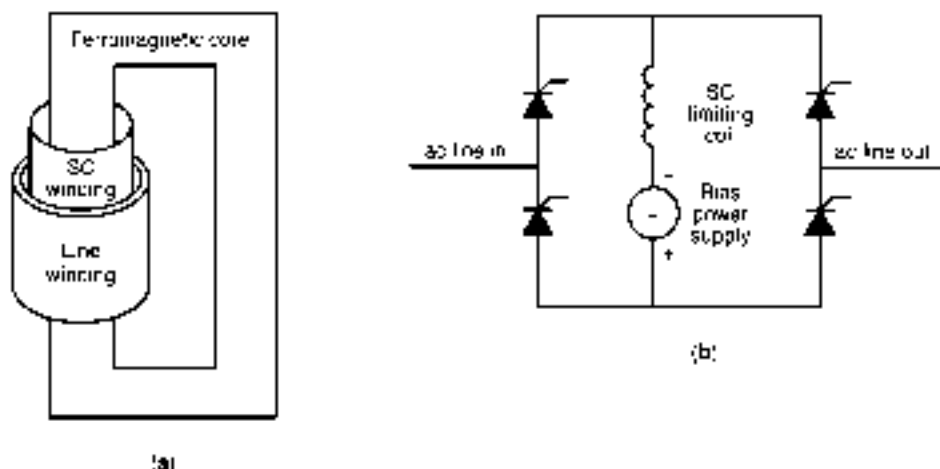


Figure 3. Principles of shielding and semiconductor SFCLs: (a) shielding SFCL, (b) semiconductor bridge SFCL.

variations in material quality, temperature, and magnetic field. The concern is whether such locations will bear a disproportionately large part of the complete limiting burden, in terms of temperature, electric field, or mechanical stress. This has been observed in both analysis and experiment (7,18), and significant thermal damage has been reported (14). For HTS systems, quench propagation is so slow (centimeters per second) as to be ineffective during power-cycle time periods, whereas LTS systems' faster propagation (meters per second to kilometers per second) may overcome inhomogeneous initiation. This inhomogeneity is most critical when the peak fault current is close to the trigger current, leading to a prolonged period II with incomplete fault current commutation to the bypass element (if present). Thus every quenching SFCL will be vulnerable when the fault occurs at the system phase corresponding to the lowest prospective fault current and should be designed to be capable of withstanding inhomogeneous initiation until the interrupter opens at a current zero.

After operation, a quenching SFCL must cool to its operating temperature before becoming available again. The cryogenic system temperature profile is inverted, with a "hot" SC and "cool" thermal shield. A high heat capacity thermal bulkhead to level the cool system thermal load is attractive. No matter what the cool system, the SC element reset time is seconds or longer, a major shortcoming for reclose-type fault relaying systems. Multiple parallel SFCL elements are a common solution, so that a cold element is available immediately after SFCL operation; the drawback is increased system complexity and cost from duplication of the SC element, thermal decoupling between elements, and a switching subsystem.

## FERROMAGNETIC TYPES

An SFCL using the principle of ferromagnetic desaturation was built in the early 1990s (19). A coil wound on a closed iron core will have a very high ac impedance when unsaturated. If, however, the iron core is saturated, the same coil will exhibit a very low impedance. Two cores are used per phase, each with an ac winding carrying the line current and also an SC dc bias winding. The dc bias windings are connected to saturate the two cores in opposite directions. The ac windings are series connected so that instantaneous line current adds to one core's bias MMF and subtracts from the MMF of the other core. Under

normal operation, both cores are so heavily saturated by the dc bias current that the MMF contribution of the ac does not bring either core out of saturation, and the impedance of the core is small. Under short-circuit conditions, however, each half wave of the higher ac fault current brings each core out of saturation in turn, thus inserting a high impedance into the circuit and limiting the current. At the fault current peak, the cores' resaturation (in the opposite polarity to the quiescent state) must be limited. In a three-phase system, six cores are necessary, two for each phase, but a single SC dc bias winding (and crystal enclosure) passing through all six core windows can be used. A single phase, 3 kV, 550 A prototype using an LTS coil was built and tested with peak short-circuit currents of close to 15 kA. Although the device performed as predicted, no commercial units were ever built because of the large size of a transmission system sized unit. The total core volume required corresponds to a transformer rated at twice the fault volt-amperes. The desaturating SFCL does not necessarily quench during limiting, so that it has no reset time delay, and the limiting level can be varied by adjusting the dc bias current.

The impedance of a series line reactor can also be changed by moving armature methods that vary the effective gap of the reactor magnetic circuit. One approach under construction (20) uses a mechanically balanced disk (hypercooled to 50–70 K for enhanced magnetic and conductivity properties) to rotate magnetic and nonmagnetic sectors in and out of the gap, in times below  $\frac{1}{4}$  cycle. The rotation is induced by inverter-fed HTS coils, controlled by line overcurrent relaying. Using SC in the rotation control coils significantly reduces device quiescent losses.

## SEMICONDUCTOR TYPES

The nonlinear resistance of semiconductor junctions is used in the bridge-type fault current limiter, which combines power electronics and an SC coil (21). Figure 3(b) shows the circuit of a single phase. A bias power supply provides a current in all four semiconductors. As long as the load current is less than the bias current, all four semiconductors are forward biased, and the ac load current flows unimpeded by the bridge circuit, assuming negligible losses in the semiconductors. In the quiescent condition, each thyristor conducts half the bias current superimposed with half of the line current. When a short circuit

occurs, one pair of semiconductors is turned on, and the other pair is turned off in each half cycle, automatically inserting the bridge inductor into the circuit. The inductor limits the fault current. Because the inductor carries a bias current under normal condition, use of an SC coil reduces the overall system losses. The bridge-type fault current limiter has several attractive features, such as automatic insertion of the current-limiting reactor, reduction of the first half-cycle short-circuit current, precise control of the amplitude of the short-circuit current, complete current interruption in less than a cycle if desired, operation with multiple fast reclosures, and high efficiency. In addition, it is conceivable to use the controlled thyristor bridge for other transmission or distribution network functions such as var control. For distribution and transmission system semiconductor SFCLs, the equivalent thyristors indicated in Fig. 3 by an series strings of thyristors, similar to those used in static var compensators or electronic transfer switches. A 2.4 kV, 2.3 kA single-phase fault current limiter was successfully tested in 1995 and a 15 kV, 500 A, three-phase unit is being designed to be tested in 1998 [2].

#### OTHER TYPES

SFCL designs where the limitation arises from magnetic circuit unbalance under fault conditions have been proposed—for example, with a three-phase SC winding on a common core [22]. The positive and negative sequence inductances equal the device leakage inductance, whereas the zero-sequence inductance equals three times the self-inductance. Single line-to-ground faults, which are the most common, are then limited by the magnetics alone, without quench or thermal recovery delay, whereas the rarer two- and three-phase faults will be limited by in-line SC winding quench. Thus, the SC winding thermal performance must be designed as for all in-line quenching SFCLs. An HVDC transmission SFCL [23] has balanced positive and negative line currents canceling in the magnetic circuit, whereas ground faults will unbalance the SFCL. Under these conditions, the SFCL inductance limits the fault current  $dI/dt$  until the core saturates.

#### SUPERCONDUCTOR REQUIREMENTS

Either niobium-based LTS or ceramic HTS materials may be considered. LTS conductor technology is relatively mature, with kilampere-rated commercial conductor optimized for either ac or dc operation available in long lengths from several vendors. Operating  $J_c$ 's are in the gigampere per square meter range and operating temperatures are up to 6 K (NbTi) or 14 K (Nb<sub>3</sub>Sn). To produce the required normal-state resistance in quenching SFCLs, the LTS conductor must have high values of effective  $\rho_n$ , which is governed by the stabilizer resistivity. The highest values (in available conductors) of about 30  $\mu\Omega/\text{m}$  (300 K) are achieved with CuNi alloy. Elementary power frequency conductor strands have 10–100 A  $J_c$  and pass through two levels of cabling and twisting to achieve the required kilampere-class  $J_c$ . At each level of cabling, the strands are grouped around a copper core, which enhances quench propagation. Current sharing is a problem that is further accentuated when kilampere-class conductors must be paralleled to operate at 10–100 kA-class transmission network-rated currents. Such an ac conductor is more expensive than a dc conductor. Desaturating and semicon-

ductor SFCL types do not need the same high ac performance and use more economical dc LTS conductors. The major drawback of LTS conductors is the refrigeration and cooling required to maintain the required operating temperature.

Bulk, film, or wire/tape HTS conductors can be used in SFCLs. Bulk rings or tubular cylinders, particularly useful for shielding SFCLs, have been manufactured from Bi-2212 [12], Bi-2223 [24], and YBCO [9]. All the materials have  $T_c$  above 85 K, effective  $J_c$  in the 0.1–0.5  $\text{GA}/\text{m}^2$  range at 77 K, and  $\rho_n$  around 10  $\mu\Omega/\text{m}$ . Even though bulk Bi-2212 parameters are overall no better than the competing materials, Bi-2212 has the most promise because there are practical manufacturing techniques to produce technically useful cylinders with up to 8 mm wall thickness and 40 cm diameter. Bulk forms are also used as current leads operating between  $T_c$  and  $T_n$ . HTS materials show significantly higher  $J_c$ 's in thin films. For example, 200 nm-thick YBCO on a CeO<sub>2</sub>-buffered sapphire substrate exhibited  $J_c = 30 \text{ GA}/\text{m}^2$  and was investigated for quenching SFCL designs [11]. The substrate heat capacity and heat transfer to LN<sub>2</sub> cryogen limits the temperature rise. Such high  $J_c$  requires using a stabilizer parallel layer to protect against catastrophic local overheating, so that the high intrinsic  $\rho_n$  is not used. YBCO thin film panels up to 20 × 20 cm have been tested, which could be used in arrays to produce quenching SFCLs. Thick Bi-2212 films are also useful; an array of MgO substrate cylinders (45 cm diameter, 12 cm long), each supporting an 0.5 mm Bi-2212 layer, has been used in a 6.6 kV, 400 A shielding SFCL [25]. Thick film  $J_c$ 's are similar to bulk LTS material.

The thrust of HTS wire and tape development is to develop stabilized conductor for magnet applications. Bi-2223 conductor, produced by the powder-in-tube (PIT) process, is currently closest to commercialization but limited to low operating magnetic field levels. The silver stabilizer of PIT conductor (with  $\rho$  around 4 n $\Omega/\text{m}$  at 100 K) bypasses the much higher intrinsic Bi-2223  $\rho_n$ . Multifilamentary twisted conductors are under development to produce lower ac losses. The manufacture of wire from other HTS materials such as Bi-2212 and YBCO is also under active development, and it is premature to predict which material will be most applicable to SFCLs.

The main attraction of HTS materials over LTS for SFCLs is higher operating temperatures and lower refrigeration costs. Present HTS  $T_c$ 's point to useful operating temperatures in the 30–60 K range. As in other HTS applications, tradeoffs must be made between the advantages of using cheap LN<sub>2</sub> cryogen (77 K operation), larger amounts of HTS, more exotic cryogens (e.g., liquid neon, boiling at 27 K), or none coupled with better use of the HTS (lower operating temperatures).

#### CRYOGENICS

Cryostat thermal design is of paramount importance because the required cooling power is an important part of the operating cost. The quiescent cryostat heat budget consists of SC ac losses, ambient-to-cryogenic current lead heat leak, dielectric losses, and other conduction and radiation to the cryostat. Even though the design and selection of superconductor for low losses have been already discussed, and cryostat and current leads design is well understood, dielectric ac losses present a less-appreciated heating mechanism. The heat dissipation for insulation deployed in high ac electric field regions should be determined from cryogenic loss tangent. Transmission and

distribution-rated in-line SFCLs using HTS may have quiescent heat loads 400 W and up as a result of ac loss and current lead penetrations. In all cases, a closed system needs refrigeration capacity to extract the significant heat leak. When refrigeration power is interrupted, the cryostat temperature will rise at a rate governed by ac losses, resistive losses, heat leak, and heat capacity, and ultimately the SC system will lose superconductivity. Auxiliary power interruptions are quite likely in an electric utility setting, and an SFCL must have some minimal ride-through capability. Electric utilities will find the cooldown time to operating temperature an important parameter. Liquid cryogen-based cooldowns may take as little as a few hours, whereas refrigerator cooldowns may take many hours to days.

SFCL cryostats may use liquid or gaseous cryogen for heat transfer and storage or may operate under vacuum. Cryogen consumption must be kept low to avoid replenishment servicing any more often than annually, so refrigeration must have the capacity to meet the quiescent heat leak. In the typical power and temperature range for LTS systems, two-stage (Gifford-McMahon (GM) cycle cryocoolers) with possibly a Joule-Thompson (JT) valve are typical, whereas for HTS systems, one- or two-stage GM or Stirling cycle cryocoolers are candidates. The energy efficiencies of these refrigerators are well below the Carnot cycle limit, and typically a 7 kW helium compressor will be required for 2–3 W cooling at 4 K or 100–200 W at 80 K. The maintenance interval of cryocoolers depends on internal seal wear, oil mist freeze-out, and JT valve clogging. Use of a good oil-free high-pressure ambient temperature helium compressor gives maintenance intervals longer than one year.

## CONCLUSION

SFCL is an emerging technology that is being driven by the confluence of superconductor capability developments and electric utility need. Fault current limiters would enable utilities to reap the economic benefits of operating transmission lines at the highest possible power transfer levels and the operational benefits of transmission system interconnection. High-temperature superconductor developments may soon become commercially available to enable production of some of the many SFCL approaches under development at present. The utility requirements of equipment robustness and reliability present a challenge for SFCLs which can be technically met. A number of utility demonstration projects of SFCLs are either underway or close to being underway, which will provide much-needed operational, robustness, and reliability data. A yet unresolved issue is ultimate SFCL cost resulting from the inherent costs of high-technology cryogenic and superconducting systems. There is scope for significant inventions and improvements in fundamental SFCL concepts and their implementation and integration into electric transmission and distribution systems and components.

## BIBLIOGRAPHY

- J. G. Bednorz and K. A. Müller, Possible high- $T_c$  superconductivity in the  $Ba_{1-x}Pb_xCuO$  system, *Z. Phys. B*, **84**: 189–193, 1986.
- F. Leung, Surge protection for power grids, *IEEE Spectrum*, **34**(7): 26–30, 1997.
- R. F. Giese and M. Runde, Assessment study of superconducting fault-current limiters operating at 77 K, *IEEE Trans. Power Del.*, **8**: 1139–1147, 1993.
- G. C. Dijkstra et al., Superconducting technology for current limiters and switchgear, *CYBERE*, 1990.
- I. Salomon et al., Comparison of superconducting fault current limiter concepts in electric utility applications, *IEEE Trans. Appl. Supercond.*, **5**: 1079–1082, 1995.
- W. Paul et al., Test of 1.2 MVA high- $T_c$  superconducting fault current limiter, *Applied Superconductivity 1997, Proc. EUCAS 1997*, Third European Conference on Applied Superconductivity, The Netherlands, June 30–July 2, 1997, **2**: IOT, Bristol UK, 1997.
- M. Landmayer, High temperature superconductors as current limiters—An alternative to contacts and arcs in circuit breakers?, *Proc. 39th IEEE Winter Conference on Electrical Contacts*, Pittsburgh 1989, pp. 1–10.
- T. Verhage et al., Experimental 7.2 kV rms-1 kA rms-3 kA peak current limiter system, *IEEE Trans. Appl. Supercond.*, **3**: 574–577, 1993.
- T. Hiza et al., Development of a new 6.6kV/1500A-class superconducting fault current limiter for electric power systems, *IEEE Trans. Power Deliv.*, **8**: 192–193, 1993.
- M. Landmayer and M. Schubert, Resistive fault current limiters with HTSC—Measurements and simulation, *IEEE Trans. Appl. Supercond.*, **8**: 884–888, 1993.
- B. Gronoff et al., Resistive current limiters with YBCO films, *IEEE Trans. Appl. Supercond.*, **7**: 828–831, 1997.
- J. Buck, S. Elschner, and P. E. Herrmann, MCP BSCCO 2212 tubes for power applications up to 10 kA, *IEEE Trans. Appl. Supercond.*, **6**: 1400–1412, 1996.
- K. Kajikawa et al., Design and current-limiting simulation of magnetic-shield type superconducting fault current limiter with high  $T_c$  superconductors, *IEEE Trans. Magn.*, **32**: 2667–2670, 1996.
- J. R. Cave et al., Testing and modelling of inductive superconducting fault current limiters, *IEEE Trans. Appl. Supercond.*, **7**: 632–635, 1997.
- L. S. Fleishman et al., Design consideration for an inductive high- $T_c$  superconducting fault current limiter, *IEEE Trans. Appl. Supercond.*, **3**: 570–573, 1993.
- M. Joo and T. K. Ko, The analysis of fault currents according to core saturation and fault angles in an inductive high- $T_c$  superconducting fault current limiter, *IEEE Trans. Appl. Supercond.*, **8**: 62–67, 1996.
- H. J. Boonig et al., Anisotropic high temperature superconductors as variable resistors and switches, *IEEE Trans. Appl. Supercond.*, **5**: 1040–1043, 1995.
- V. Meerovich et al., Development of high- $T_c$  superconducting inductive current limiter for power systems, *Cryogenics*, **34** (ICEC supplement): 757–760, 1994.
- R. P. Raja, K. C. Partan, and T. C. Bartram, A current limiting device using superconducting dc bias, *IEEE Trans. Power Appar. Syst.*, **PAS-101**: 3173–3177, 1982.
- S. B. Kuznetsov, Superconducting fault current limiters, *Superconductor Industry*, **10**(1): 29–30, 1997.
- H. J. Boonig and D. A. Paice, Fault-current limiter using a superconducting coil, *IEEE Trans. Magn.*, **19**: 1951–1953, 1983.
- S. Shimizu et al., Equivalent circuit and leakage reactances of superconducting 3-phase fault current limiter, *IEEE Trans. Appl. Supercond.*, **3**: 578–581, 1993.
- T. Jahigahka and M. Suzuki, Fundamental test of new dc superconducting fault current limiters, *IEEE Trans. Magn.*, **MAG-27**: 2941–2944, 1991.
- M. Ichikawa and M. Okuzaki, A magnetic shielding type supercon-

**Table 1. Basic Characteristics of the Most Used HTS Thin Films and Dielectric Substrates for Microwave Applications**

| HTS Films   | YBCO   | TBCCO   |
|---|--|---|
| Critical Temperature ( $T_c$ )                                      | 90 K   | 105 K   |
| Surface resistance ( $R_s$ ) at 77 K and 10 GHz ( $f^2$ dependence) | 0.2–0.5 m $\Omega$                                       | 0.2–0.5 m $\Omega$                                  |
| Film thickness  | 400–800 nm   | 800–1000 nm   |
| Critical dc current density ( $J_c$ ) at 77 K                       | 10 <sup>6</sup> A/cm <sup>2</sup>                        | 10 <sup>6</sup> A/cm <sup>2</sup>                   |
| Penetration depth (nm)  | 200 nm   | 200 nm  |
| Substrates  | LaAlO <sub>3</sub> (LaAO)                                | MgO   |
| Relative dielectric constant ( $\epsilon_r$ ) at 77 K               | 23.4   | 9.7   |
| Dissipation factor ( $\tan \delta$ ) at 77 K                        | <10 <sup>-5</sup>  | <10 <sup>-6</sup>                                   |
| Typical dimensions  | 5, 7.5, and 10 cm diameter;<br>250 and 500 $\mu$ m thick | 5 and 7.5 cm diameter;<br>250 and 500 $\mu$ m thick |

Both types of HTS Films can be grown in either substrate.

The most important characteristics of these two HTS materials and their substrates are given in Table 1. The properties listed are those a microwave designer would want to know of engaged in an HTS device design project.

#### Device Processing

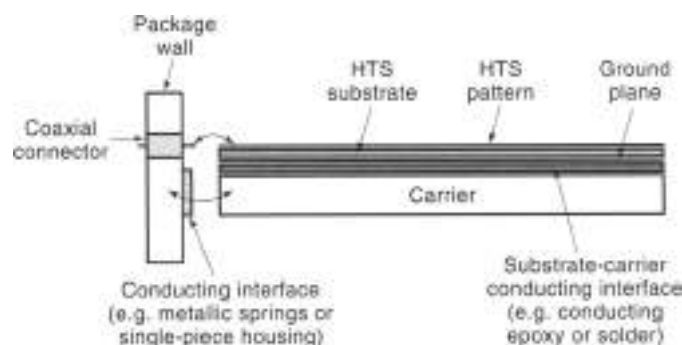
Fabrication of microwave devices using either YBCO and TBCCO follows relatively straightforward photolithographic techniques. Patterning of the superconducting layer is typically accomplished by Ar-ion milling. The processing must include the deposition and patterning of low-resistivity contacts for interfacing with other devices or instrumentation. These are typically made by depositing a thin (200 nm to 300 nm) layer of gold or silver followed by an annealing step at 500°C to 600°C (7). Interfacing with other devices via coaxial connectors or directly to other substrates, superconducting or otherwise, can be accomplished using gold wire or ribbon attached to the low-resistivity contacts by ultrasonic thermal-compression bonding or gap-welding (ribbons). Fabrication details of filters and delay lines made at Northrop Grumman can be found in (8). Other institutions follow similar procedures.

#### Microwave Packaging of High-Temperature Superconducting Devices

Proper packaging of cryogenic microwave devices is critical to the success of the technology. In general, planar microwave devices are made up of a dielectric or semiconducting substrate, typically mounted on a metallic package or carrier by means of conducting epoxy or soldering. The mounted device then interfaces with other devices or measurement instrumentation most commonly via coaxial connections. In this case, the metallic carrier forms part of the ground terminal by connecting to both the outer jacket of the coaxial interconnection and the ground plane defined on the substrate. This means that microwave currents flow into the planar device through the center conductor of the coaxial interface and back out through the ground plane on the substrate, the metallic carrier, and its connection to the coaxial connector outer jacket. This current flow into the packaging

assembly must occur without appreciable impedance mismatch or discontinuity. Figure 2 illustrates these points schematically with an example for a microstrip device.

The criticality of these connections between substrate, carrier, and connectors is exacerbated when cryogenic cooling of the microwave package is required because a large thermal mismatch between the various components may cause cracking of the substrate or degradation of the interface between substrate and carrier or connector. Furthermore, the quality of the interface between the cryoelectronic substrate and the outside world must be preserved through many temperature cycles between ambient and cryogenic temperature to allow for repeated testing of the device and an operational environment, which may require that the device warm up to room temperature while not in use. Suitable microwave package designs must therefore include one or more of the following elements: (1) the use of thermally matched materials, (2) adhesives that remain sufficiently pliable at cryogenic temperatures, and (3) configurations that allow the various parts of the package to contract and expand freely while maintaining good electrical contact.



**Figure 2.** Microwave packaging of HTS devices is challenging because mechanical and electrical integrity must be maintained when the device is cycled from ambient to cryogenic temperatures. Special attention must be paid to the ground-current return path so as not to introduce parasitic reactances that could severely affect the performance.

ducting fault current limiter using a Bi-2212 thick film cylinder. *IEEE Trans. Appl. Supercond.* 5: 1067-1070, 1995.

- 25 H. Kaji and M. Ichikawa, Performance of a high-Tc superconducting fault current limiter. *IEEE Trans. Appl. Supercond.* 7: 993-996, 1997.

LEMMIT SALARCO  
General Electric Company  
HEINRICH J. HOFMIG  
Los Alamos National Laboratory

## SUPERCONDUCTING FILTERS AND PASSIVE COMPONENTS

The advent of high-temperature superconducting (HTS) materials has enabled a number of applications in passive microwave electronics. Superconductors exhibit very low losses at microwave frequencies and, although finite, at the practical operating temperature of 77 K (boiling point of nitrogen), these losses are more than two orders of magnitude lower than normal conductors at frequencies of 10 GHz and below. This has allowed the possibility for high performance planar microwave components since high-quality epitaxial films can be deposited on both sides of low microwave loss, single-crystal substrates allowing the fabrication of components in planar configurations such as microstrip, stripline, and coplanar waveguide. These configurations are widely used throughout the microwave community in a variety of technologies ranging from GaAs microwave monolithic integrated circuits (MMIC) to integrated circuits using ceramic and laminated substrates. Figure 1 shows a schematic cross-section of all three most common planar microwave structures. The use of HTS in these circuit configurations results in passive devices such as resonators, filters, and delay lines with performance far superior to conventional planar technology and with the attractive feature that many well-established design techniques can be used for HTS circuits as well.

The discussion in this article will focus on HTS microwave technology with the understanding that conventional low-temperature superconductors (LTS), for example, Nb or NbN, can also be used in the same fashion. Practical LTS materials

operate typically at 4.2 K, the boiling temperature of liquid He. Furthermore, it should be kept in mind that a larger variety of substrates can be used in LTS technology because it does not require single-crystal epitaxial films. For example, Nb microwave and digital circuits have been demonstrated on Si and single-crystal sapphire (1).

## HIGH-TEMPERATURE SUPERCONDUCTING FILM PROCESSING TECHNOLOGY

The basic elements of HTS thin-film technology from the point of view of microwave applications will be discussed here briefly. Only epitaxial thin films will be addressed because, to date, significantly better properties can be obtained from epitaxial material than from polycrystalline, nonepitaxial techniques. Most, if not all, of the developments covered in this article, however, are valid for the case of polycrystalline materials. The main advantage of these materials is cost and coating of nonplanar surfaces like the inside of a cylinder to form a high-Q cavity resonator (2).

Among the various high-temperature superconductors there are two that have achieved a level of maturity and acceptance in the industry: (1)  $\text{YBa}_2\text{Cu}_3\text{O}_7$ , usually referred to as YBCO, and (2)  $\text{Tl}_2\text{Ba}_2\text{CaCu}_2\text{O}_8$  or TBCCO. High-quality epitaxial thin films of these materials can be deposited on both sides of a variety of low microwave loss, single-crystal substrates by physical vapor deposition techniques, mainly sputtering (3) and laser ablation (4) and also by metal-organic chemical vapor deposition (MOCVD) (5) and coevaporation (6). The two most common substrates to date are  $\text{LaAlO}_3$  (LAO) and  $\text{MgO}$ . Circuits on sapphire have been demonstrated but the crystal lattice mismatch between HTS and sapphire is large enough to restrict the film thickness to below desirable values of at least 500 nm. Both LAO and  $\text{MgO}$  substrates are available commercially from a variety of suppliers around the world and can be readily obtained in circular wafers up to 7.6 cm in diameter and 250  $\mu\text{m}$  to 500  $\mu\text{m}$  in thickness. At 77 K and for frequencies between 1 GHz and 10 GHz the loss factor ( $\tan \delta$ ) of LAO and  $\text{MgO}$  is less than  $10^{-6}$ . This is 10 to 100 times smaller than most practical microwave substrates and is compatible with the low conductor loss of HTS.

YBCO is typically grown in situ as an epitaxial single-crystal film with thickness ranging from 400 nm to 600 nm. This means that the desired crystalline phase of the material is formed as the film grows because the growth conditions can be adjusted to obtain such results. In contrast, an amorphous film of TBCCO is deposited first and then the film is subjected to a postdeposition annealing treatment to form the right crystalline phase. One advantage of the in situ deposition of YBCO films is the ability to deposit other crystalline layers such as insulators or normal conductors, needed for the fabrication of Josephson junctions. HTS microwave technology can leverage the research on Josephson junctions by making use of insulating layers in other ways such as lumped element capacitors, for example. TBCCO films have the advantage of a higher critical temperature (Table 1) and are typically grown with thickness close to 1  $\mu\text{m}$ . These are advantages because practical devices must be made with thicknesses two or three times greater than the London penetration depth (200 nm to 300 nm) at 77 K. Tl(Cu)O films thus offer a greater operating margin than YBCO films.

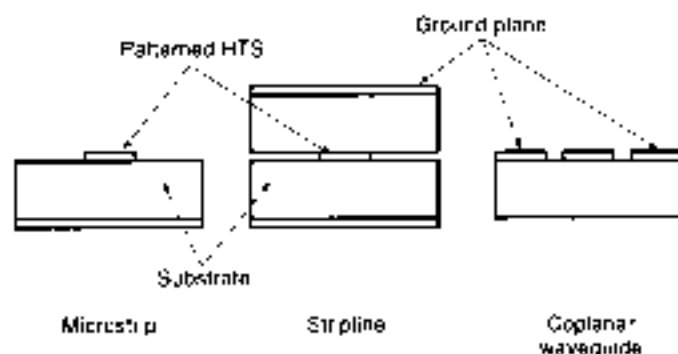
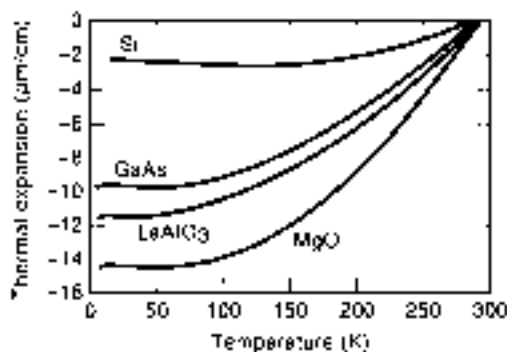


Figure 1. HTS technology lends itself to the fabrication of microwave planar devices such as those schematically depicted here in cross-section. One advantage of coplanar waveguide (and its variant slot line, not shown) is that only one patterned HTS film is needed. One disadvantage of stripline is that it is difficult to package. HTS devices have been demonstrated in all three structures.

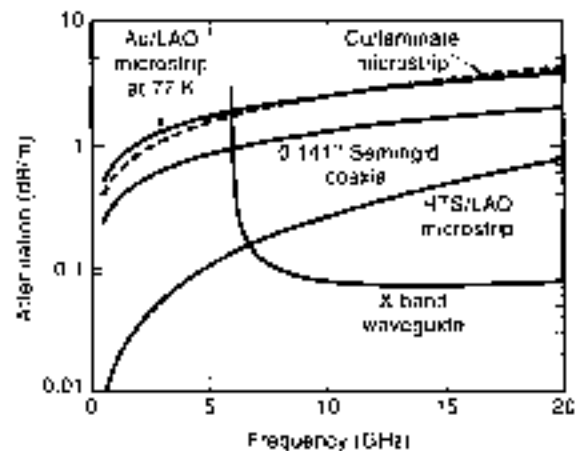


**Figure 3.** Measured relative thermal contraction of several materials of interest in HTS technology, including Si and GaAs. These are similar to those that should be used in the selection of carrier materials and the packaging of hybrid HTS-semiconductor components. If a good thermal match between parts that must remain in intimate mechanical and electrical contact cannot be obtained, sufficiently pliable conductive adhesives may be adequate for some applications, particularly if the mismatched parts are small.

When considering materials that are thermally matched to HTS substrates the important parameter is the total contraction between room temperature and the operating temperature, for example, 77 K, and not the thermal coefficient of expansion. Figure 3 shows the relative thermal contraction for several materials, including LAO and MgO, in  $\mu\text{m}/\text{cm}$  (i.e., at a given temperature, a 1 cm-long piece of material shrinks so many  $\mu\text{m}$  from its length at 300 K). As can be seen, the slope of this curve, that is, the thermal coefficient of expansion, varies greatly with temperature over the range of interest, making it a practically useless parameter for the selection of appropriate materials. Although there are some differences in the rate of contraction as the materials shrink from room temperature, the key parameter is the total contraction at 77 K. For example, Nb and LAO are fairly well matched at 77 K, even though their rate of contraction as a function of temperature is slightly different. This is borne out by extensive experimentation (6). Other substrate/carrier material pairs have been successfully used as well.

#### PASSIVE SUPERCONDUCTING MICROWAVE DEVICE FUNDAMENTALS

The key reason for developing a microwave HTS technology is the exploitation of the low loss afforded by HTS compared with conventional metals like gold and copper. High-performance, low-loss devices using conventional materials can generally be made at the expense of high volume, usually in the form of hollow or partially dielectric-filled waveguide components. The potential of HTS is to enable components with the same or better performance in a much reduced volume which must include the cryocooler. Figure 4 shows a comparison between calculated losses in several common types of transmission line, including HTS microstrip on LAO, for the parameters listed in Table 1. For calibration, included in this comparison is Au microstrip at 77 K, also on LAO. For the microstrip cases only the conductor and dielectric losses were calculated; radiation losses or coupling to spurious surface modes were ignored. Also included are the losses in X-Band waveguide, which are lower than HTS for the parameters chosen, although HTS microstrip offers the



**Figure 4.** Calculated attenuation comparison for various transmission line types including HTS and gold microstrip on 500  $\mu\text{m}$ -thick LAO substrates. The line impedance for each line was 50  $\Omega$ . All microstrip substrates were assumed to be 600  $\mu\text{m}$  thick. Notice that X-band waveguide has lower loss than HTS microstrip for the parameters chosen. However, HTS microstrip has broader bandwidth and the potential for smaller volume because it facilitates the integration of several microwave components.

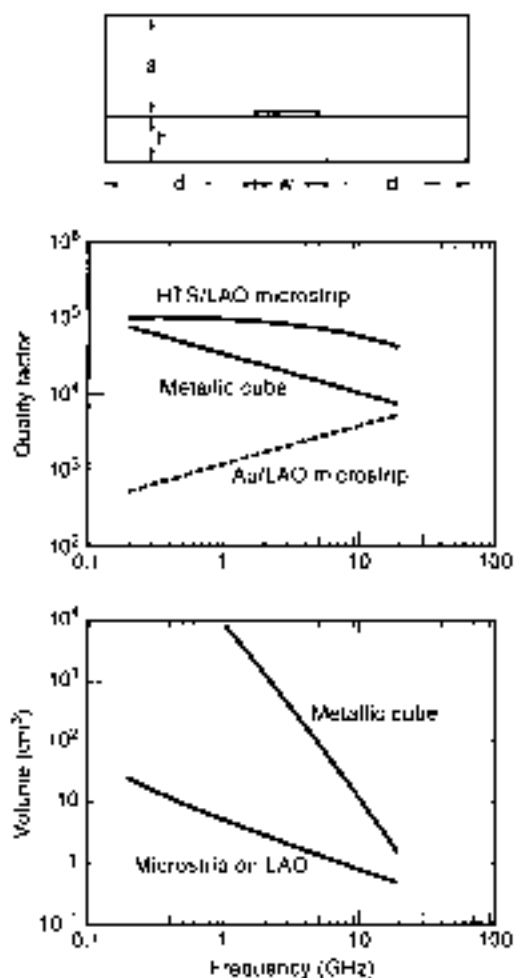
advantage of much wider bandwidth and ready integration of several components into a small volume. Notice that the useful frequency range for X-band waveguide is typically only 8 GHz to 12 GHz. Integration capability combined with low loss are key elements of HTS technology because most systems insertion opportunities will arise for applications offering significant size advantages with respect to conventional approaches.

A different and also useful way of getting insight into the advantages of HTS planar circuits is from the point of view of resonant structures, which form the basic building block of passband filters. This can best be discussed in terms of quality factor  $Q$ , which is the ratio between stored and dissipated energy.

$$Q = \frac{\text{Stored energy}}{\text{Dissipated energy}} \quad (1)$$

For empty electromagnetic cavity type resonators, this is, in essence, a figure of merit measuring the degree of compromise between volume (stored energy) and surface area (microwave losses on the conducting surfaces). In general, both stored and dissipated energy depend on the dielectric constant and the geometric configuration used, the dissipated energy depending also on the surface resistance of the superconducting surfaces and the losses in the dielectric. In practice, there could be other types of losses such as radiation, which, for simplicity, will be neglected in this discussion. In principle, the performance of any passive device can be projected from the  $Q$  of the type of structure used to make up the device. For example, a resonator made up of a section of microstrip line can be calculated from well-known expressions (9). The insertion loss of a filter can, in turn, be estimated from the  $Q$  of the resonators that make up the filter (10).

Figure 5 is a plot comparing the  $Q$  and volume, as a function of frequency, of resonators made up of a microstrip line section and an empty metallic tube, respectively. Since the dimensions of the resonator are specified at each frequency, the volume



**Figure 5.** Calculated comparison of quality factor ( $Q$ ) and volume for two types of resonators:  $50\ \Omega$  HTS microstrip on  $500\ \mu\text{m}$ -thick LAO and a metallic cube. The  $Q$  of a gold-on-LAO microstrip at  $77\ \text{K}$  was also calculated as a reference. The metallic cube is representative of a simple cavity resonator and was chosen because its  $Q$  is easy to calculate [12]. The volume for the microstrip was chosen as a device with the cross-section shown in the figure with  $w/h = 10$ ,  $d/h = 20$ , and a length of  $1.2 \times 4h$ . This assumed cross section is independent of wavelength and loses its meaning at the higher frequencies plotted, where the substrate would, in practice, be thinner, making the microstrip always smaller than the waveguide.

calculated is that of the smallest cube capable of resonating at a given frequency. The microstrip HTS resonator volume was calculated assuming it is in an enclosure with cross-section as shown in the figure, where the walls and the lid are sufficiently far away from the superconducting strip that their contribution to the loss is negligible. The  $Q$  of HTS microstrip, although higher, is within the same order of magnitude as the  $Q$  of the cube resonator. However, the estimated volume can be two or more orders of magnitude smaller, especially at the lower microwave frequencies.

Although to first order the main feature distinguishing HTS from conventional planar microwave passive devices is low loss, other important differences exist and are discussed in the following sections.

**Surface Impedance and Penetration Depth**

The surface or internal impedance of a conductor is the characteristic impedance seen by a plane wave incident

perpendicularly upon a planar (super)conducting surface. For both normal (e.g., copper, gold) conductors and superconductors, the surface impedance per unit length and width is given by [11,12]:

$$Z_s = R_s + jX_s = \sqrt{\frac{j\omega\mu}{\sigma}} \tag{12}$$

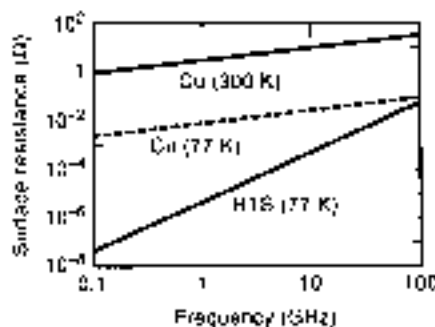
where  $\omega$  is the angular frequency,  $\mu$  is the permeability, and  $\sigma$  is the conductivity.  $\sigma$  is real for normal conductors but is complex for superconductors. In both cases the RF fields decay exponentially inside the material, defining a field penetration depth. In the case of superconductors, however, this parameter is independent of frequency and is orders of magnitude smaller than the normal conductor penetration depth (usually referred to as the skin depth). The reasons are derived from the perfect diamagnetism of superconductors, the so-called Meissner effect, and are explained by the Gorter-Casimir and London two-fluid model of superconductivity [11,13].

Table 2 summarizes the differences between normal and superconductors from the point of view of their microwave surface impedance. Notice that the surface resistance of superconductors,  $R_s$ , has a frequency-squared ( $f^2$ ) dependence. In contrast, normal conductors depend on the square root of frequency ( $\sqrt{f}$ ). Figure 6 shows the difference between copper at  $300\ \text{K}$  and  $77\ \text{K}$ , and HTS at  $77\ \text{K}$ . This must be taken into consideration, especially when designing wide-band components. The figure also highlights the large difference between copper and HTS at frequencies below  $1\ \text{GHz}$ .

As with planar microwave devices using normal conductors, best performance control is obtained when the geometric inductance of the current dominates the kinetic (or internal) inductance of the superconductor. That is, from a practical point of view, the thickness of the superconductor must be at least two to three times larger than the penetration depth at the temperature of operation. The penetration depth is a strong function of temperature and for HTS it is given approximately by [11,13]:

$$\lambda_L(T) = \frac{\lambda_L(0)}{\sqrt{1 - \left(\frac{T}{T_c}\right)^2}} \tag{13}$$

where  $\lambda_L(0)$ , the penetration depth at  $0\ \text{K}$ , is a fundamental parameter of the material,  $T$  is the temperature and  $T_c$  is the critical temperature. For YBCO,  $\lambda_L(0) \approx 150\ \text{nm}$ , which results in  $\lambda_L(77\ \text{K}) \approx 214\ \text{nm}$ . The HTS film must be at least  $500\ \text{nm}$



**Figure 6.** Surface resistance of HTS at  $77\ \text{K}$  and copper at  $77\ \text{K}$  and  $300\ \text{K}$  as a function of frequency. The surface resistance of copper scales with frequency as  $f^{1/2}$ ; for HTS it scales as  $f^2$ .



**Table 3. Microwave Surface Impedance Comparison Between Normal Conductors and Superconductors**

| Normal Conductors                                      | Superconductors   |
|--|---|
| $\sigma$ - Real  | $\sigma$ - Complex  |
| Penetration (skin) depth:                              | Penetration depth:  |
| $\delta = \sqrt{\frac{1}{\sigma f \mu \epsilon}}$      | $\lambda_L(T) = \frac{\lambda_L(0)}{\sqrt{1 - \left(\frac{T}{T_c}\right)^2}}$ |
|  | $\lambda_L(0) \approx 150 \text{ nm for YBCO}$                                |
| $R_s = X_s = \frac{1}{\sigma \delta} \approx \sqrt{f}$ | $R_s \propto f$   |
|  | $X_s \approx 2\pi f \mu \lambda_L$  |
| Copper at 300 K:                                       | YBCO at 77 K:   |
| $\delta = \frac{2.1}{\sqrt{f}} \mu\text{m}$            | $\lambda_L(77) \approx 0.2 \mu\text{m}$                                       |
| $R_s = 8.24 \sqrt{f} \text{ m}\Omega$                  | $X_s = 5f^2 \mu\Omega$  |
| Copper at 77 K:  |   |
| $\delta = \frac{0.9}{\sqrt{f}} \mu\text{m}$            |   |
| $R_s = 3.4 \sqrt{f} \text{ m}\Omega$                   |   |

to 600 nm thick for operation at 77 K, in order for the kinetic inductance effects to be negligible with respect to the total inductance of the circuit.

For practical microwave design purposes, this allows treating the superconductor as a normal conductor with a surface resistance that can be obtained from measured values and a frequency-squared scale factor. It has become customary for workers in the field to normalize the surface resistance to 10 GHz and 77 K, even though measured data may have been taken at a different frequency. Devices where the kinetic inductance is allowed to dominate have been demonstrated [14]. However, they are lossy, difficult to fabricate, and quite dependent on temperature because of the strong temperature dependence of the penetration depth.

#### LaAlO<sub>3</sub> (LAO) Substrate Properties

The relative dielectric constant of LAO, one of the preferred HTS substrates, is  $\epsilon_r = 23.4$  at 77 K. This is a higher value than most common microwave substrates whose dielectric constants usually do not exceed 10 ( $\epsilon_r = 9.7$  for MgO), another preferred HTS microwave substrate. The significance of this is that common planar component design techniques are based on empirical circuit models whose validity may not extend to the relatively high dielectric constant of LAO. For example, the design of microstrip parallel-coupled-line filters involves the use of quarter-wave coupled-section models that generally are not valid for  $\epsilon_r > 18$ . This requires that more sophisticated design techniques be developed, making use of electromagnetic field solvers, for example, or empirically extending the range of existing models.

LAO has a cubic crystal structure above about 450°C. Below that temperature it transitions to a rhombohedral structure, which is only a slight distortion from the cubic structure consisting of a very small stretching of the cubic unit cell along its diagonal. In order to release stress, the crystal will form

twin structures, symmetrically related regions oriented in different directions. Noncubic crystals are anisotropic and, as a result of this twinning, the LAO substrate is made up of slightly anisotropic regions randomly distributed throughout the substrate [11]. The net result is an average dielectric constant with a uniformity of approximately 1%. This means a 2% variation in the resonant frequency of a planar resonator and there are many filter applications, for example, where this is not acceptable. In contrast, MgO, which is cubic, has a uniformity of 0.1% [15].

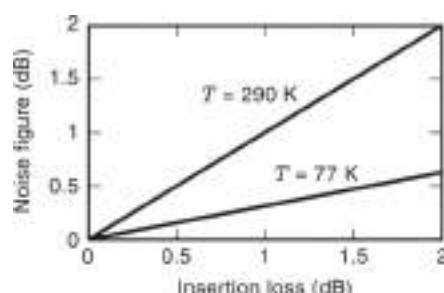
#### Dynamic Range Considerations: Noise Figure

An important consideration for any electronic device is its dynamic range, or range of signal power levels over which the device will operate properly. In the case of passive HTS devices, they are expected to be linear over a certain dynamic range, limited below by noise and above by the onset of nonlinear behavior.

Starting at the lower end, the noise generated in a passive device will generally be of a thermal nature. A measure of how much noise any device generates is given by the noise figure [16], which is, by definition, related to the excess noise generated in the device when a matched resistor at 290 K (ambient temperature) is placed at the input. Thus the noise figure would be equal to 1 (or, equivalently, 0 dB) if the device were perfectly noiseless or if it were an ideally lossless passive device. The accepted noise figure definition as a function of device temperature is [16]:

$$F_{\text{dB}} = 10 \cdot \log \left[ 1 + (L - 1) \frac{T}{290} \right] \quad (4)$$

Here,  $L$  is the insertion loss of the device as a number greater than or equal to 1 (i.e.,  $10 \log L \geq 0$  dB) and  $T$  is the temperature in degrees Kelvin. For a passive, lossy device at 290 K,



**Figure 7.** Noise figure of a cryogenically cooled passive device as a function of insertion loss. Notice that the noise figure equals the loss at 290 K (by definition), but it is lower for devices operating at a lower temperature.

the noise figure turns out to be equal to its insertion loss, a rule that system designers commonly use when dealing with passive components such as filters or lengths of transmission line. HTS devices, however, because they operate at cryogenic temperatures (77 K, typically), will have a lower noise figure, according to the accepted definition (16). Figure 7 shows this expression graphically as a function of the insertion loss of the device for 77 K and 290 K (ambient temperature). Thus, in considering the dynamic range of HTS devices, the lower end of the range will tend to be lower than for conventional devices, not only because of their inherent low loss, but also because they operate at cryogenic temperatures. Measurements reported in the literature (17) confirm, to first order at least, that the noise in HTS passive devices is indeed thermal in nature.

#### Dynamic Range Considerations: Nonlinearity and Power Handling

The dynamic range of HTS passive devices is limited above by nonlinearities in the superconductor. This is in contrast to conventional technology, for which this upper limit could be orders of magnitude higher, generally limited by such phenomena as the voltage breakdown of air or the dielectric used, or melting of the metallic pattern due to high currents. HTS, on the other hand, is fundamentally limited by the critical magnetic field, above which the material loses its superconducting properties. Nonlinear behavior occurs as the RF magnetic field approaches its critical value. Furthermore, in practical devices the microwave fields will generally tend to be nonuniformly distributed so that critical values are exceeded first in selected areas of the device. For example, a microstrip line has much higher current density near the edges of the line than along the center. Thus, as the signal power level is increased, the current density at the line edges will approach critical state first, generating nonlinearities and increased losses, degrading the performance of the device.

When a device is nonlinear it produces intermodulation distortion; that is, two signals of different frequencies applied to the device will generate mixing products. In general, the largest mixing products are those of the third order. The upper end of the dynamic range is then reached when the power level of the applied signals is such that the third-order products rise above the noise floor and can be mistaken and processed by the system as real signals. The nonlinear behavior in a device is characterized by the third-order intercept point (TOI) (18).

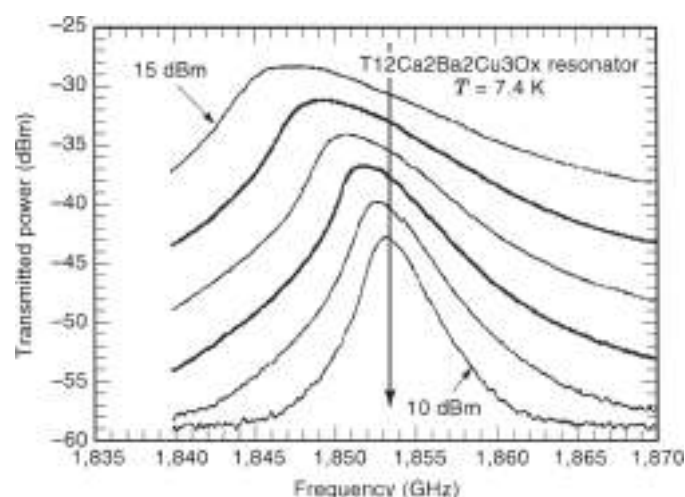
A system dynamic range can, in turn, be determined from that of its components. An important example is that of a microwave receiver front-end, usually consisting of a low-noise amplifier (LNA) placed after the antenna, which is then followed by one or more downconversion stages. The dynamic range of the receiver is greatly determined by the noise figure, gain, and TOI of the LNA, with the components that follow having much less influence. Many applications demand a preselector filter between the antenna and the LNA to reject strong interfering signals that could generate unwanted mixing products due to the nonlinearity of the LNA (17). The preselector filter must not significantly degrade the receiver dynamic range and so it must have low insertion loss (i.e., low noise figure) and a TOI sufficiently higher than the LNA's with respect to both in-band signals and the rejected out-of-band interfering signals. This is an important example because HTS filter technology is a strong candidate for this type of preselector in some applications like wireless communications base station receivers (19).

The linearity condition, which may be a relatively minor issue with conventional filter technology as, in contrast, very key in HTS filter technology. Specific developments in this area will be discussed in the section on HTS Filter Technology.

A fundamental characterization of the nonlinear behavior and power handling in HTS materials is through the surface impedance and its dependence on the RF magnetic field  $H_{RF}$  (20–25).

$$Z_s(H_{RF}) = R_s(H_{RF}) - jX_s(H_{RF}) \quad (15)$$

The essence of the nonlinear dependence of the surface impedance on signal power level or, equivalently,  $H_{RF}$ , is readily understood by observing the response of a microstrip resonator, shown in Figure 8. As the input power is increased the resonator  $Q$  degrades ( $R_s$  dependence on  $H_{RF}$ ) and the resonance shifts to lower frequencies ( $X_s$  dependence on  $H_{RF}$ ). Several



**Figure 8.** Effect of increasing the input power level on a superconducting microstrip resonator. This measurement (courtesy of Dr. M. Golubovsky, Hebrew University of Jerusalem) captures the essence of the nonlinear RF power dependence of the surface impedance  $Z_s = R_s - jX_s$ . As power level increases so does  $R_s$ , and the resonance  $Q$  degrades. On the other hand, the effect of increased power level on  $X_s$  manifests itself on a shift of the resonance toward lower frequencies (21).

regimes have been identified in the study of nonlinear phenomena in HTS (22). A linear region at sufficiently low power levels, a weakly nonlinear region where nonlinear behavior is dominated by grain-boundary weak links (Josephson-junction-like defects in the crystalline make-up of the HTS film) and a strongly nonlinear region dominated by hysteretic vortex penetration. Above this regime breakdown of the superconducting state occurs, with the surface resistance increasing abruptly due to heating and the formation of normal-state domains (23).

If the magnetic field exceeds its critical value the material becomes a normal conductor and dissipates heat which must be removed by the cryocooling system and can even damage the device. The device ceases to operate as a superconducting device and, if no damage has occurred, must recover after the high-power source has been removed. The related topic of intentionally provoking a superconducting-to-normal transition as a switching mechanism has been studied extensively (26–28).

Systematic studies of intermodulation distortion in HTS samples and devices have been reported in the literature (21,29–32). TOI values in excess of +70 dBm at 1.3 GHz and 80 K have been obtained (29) for HTS planar transmission lines a few millimeters long. This is well above most semiconductor low noise amplifiers, for example. However, it must be kept in mind that intermodulation distortion is a function of the stored energy in the device, that is, the group delay. This is important when considering the performance of HTS passband filters because filters have a delay characteristic that is lower near the center of the passband and higher toward the edges, depending on the filter order and type of response. Thus, in a practical situation, if a passband filter is intended to protect the system from a relatively high-power interfering signal, this signal may lay on the filter skirts at a given rejection level in a region of relatively high delay. The intermodulation of this interferer with a desired signal in the middle of the passband (relatively low delay) will produce spurs that define the dynamic range of the filter from that specific system's perspective. To give a quantitative example, such dynamic range might be specified as a third-order intermodulation spur level of –90 dBm for an interfering signal of +40 dBm maximum power at  $\pm 15$  MHz from a desired signal at passband center that has a maximum power level of +5 dBm. In a case like this it would be difficult to talk about TOI, which is usually defined as resulting from the intermodulation of two signals of the same power level. Such a definition would apply to a spur-free dynamic range specification for the straightforward case of two in-band signals undergoing the same group delay.

## HIGH-TEMPERATURE SUPERCONDUCTING FILTERS

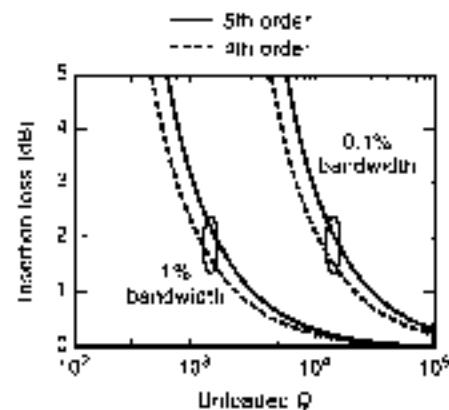
One of the most important applications of HTS microwave technology are high-performance passband filters because they can be made in planar configurations. Filters are often the dominant contributor to system volume, in particular when banks of low-loss filters are required. As discussed earlier, high- $Q$  structures can be obtained at the expense of high volume. HTS planar configurations like microstrip or coplanar waveguide have  $Q$  comparable to cavities at a much smaller volume (see Fig. 5) and so HTS is an attractive approach to reducing the volume of high-performance filters.

A straightforward way of thinking of bandpass filters is as coupled resonators. The performance of a resonator is charac-

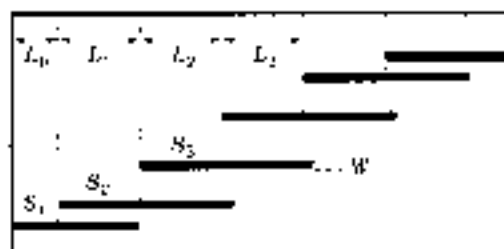
terized by its quality factor  $Q$ , defined in Eq. (1). When the resonator forms part of an electrical circuit, the circuit delivers and takes back energy from the resonator, affecting its characteristics. The unloaded  $Q$  of a resonator,  $Q_u$ , is its intrinsic quality factor, without the effects of an external circuit. A first-order filter consists of a single resonator. Its bandwidth can be adjusted by the degree of coupling into the resonator by the external circuit. In a lightly coupled resonator, little disturbance is introduced by the input and output circuits and its resonant conditions and bandwidth are close to those of an ideal, unloaded resonator. When coupling into the resonator is strong, the disturbance is large and the  $Q$  is now dominated by both the resonator and the external circuit, making the total or loaded  $Q$  lower than the ideal, unloaded  $Q$ , and therefore the filter bandwidth is now wider. For a higher-order filter made up of various resonators coupled together, the idea is the same. Narrow-band filters require that the resonators be loosely coupled to each other and the minimum bandwidth is limited by the unloaded  $Q$  of the resonators. Wider bandwidth filters will have tighter coupling among resonators. Clearly, then, narrow band filters are a desirable application for HTS because its inherent high  $Q$  enables narrow bandpass filters with low loss and small volume. This was illustrated in Fig. 5, which helps to understand the significance of the loss–volume trade-off from the point of view of using HTS and metallic cavity resonators to make filters. There is also a trade off between bandwidth, filter order (number of resonators), and insertion loss. The following is an approximate expression for the mid-band insertion loss of a filter (10), in dB, which reflects this trade-off:

$$L_{\text{dB}} \cong \frac{1.34}{B} \cdot \sum_{k=1}^n \frac{g_k}{Q_{uk}} \quad (6)$$

Here,  $n$  is the filter order,  $g_k$  are the normalized series inductance and shunt capacitance values of the low-pass prototype filter (10),  $B$  is the filter bandwidth as a fraction of the center frequency, and  $Q_{uk}$  is the unloaded  $Q$  of the  $k$ th resonator. For the purposes of estimation, it is reasonable to assume that all the resonators in the filter will have the same  $Q_u$ . Figure 9



**Figure 9.** Estimated insertion loss of fourth- and fifth-order Chebyshev passband filters of 1% and 0.1% fractional bandwidths as a function of the unloaded  $Q$  of the resonators that make up the filter. It was assumed that all the resonators have the same  $Q$ . The chart shows the increase in insertion loss caused by increasing the filter order by one and by reducing the fractional bandwidth by a factor of ten.



**Figure 10.** Parallel coupled microstrip or stripline filter topology. The lines are  $\lambda/2$  resonators coupled as  $\lambda/4$  backward-coupled sections. This is a well-known configuration suitable for fractional bandwidths of less than 1%, early demonstrations of HTS filters made use of it. However, control of the weak coupling required for bandwidths less than 1%, which the low loss of HTS allows, is very difficult and results in designs that are very sensitive to material and geometrical tolerances.

illustrates the trade-off between insertion loss, bandwidth, and filter order as a function of resonator  $Q_w$ . It shows how expression (6) can be used to estimate the potential of a certain filter technology, in this case HTS, and understand its limitations. The insertion loss was estimated for Chebyshev-type filters (10) of the fourth and fifth orders, respectively, and for 1% and 0.1% fractional bandwidths. The purpose of this chart is to point out the difference in loss caused by increasing the filter order by one and by increasing the fractional bandwidth by a factor of ten. Figure 9 complements Fig. 5 by helping to make a connection between the insertion loss of a filter of a given order and bandwidth and a specific structure and its volume. The information provided by these two figures can readily be extended to cover other structures and technologies.

### Design Considerations

As discussed above, some of the most important applications of superconductors are in narrow passband filters because they can be realized in planar technology, which lends itself to small structures that can be readily integrated with other filters and circuitry. It was also discussed that the coupling between the resonators that make up a narrow-band filter needs to be weak. The mechanism for implementing weak coupling between resonators must allow for its control and predictability so that robust filter designs which are relatively intolerant of external spurious coupling mechanisms can be implemented. An example illustrating this point can be found in the parallel coupled line filter topology. Figure 10 shows this topology, which is well known as being suited for microstrip filters with relative bandwidths below 1% (10). An analysis based on Figs. 5 and 9, however, shows that if HTS is used then bandwidths below 1%

are possible from the loss standpoint. Indeed, this structure was used by several research groups to make initial HTS filter demonstrations with 1% to 2% fractional bandwidths (33,44). Table 3 shows the couplings required to achieve a 1.25% bandwidth, fourth-order Chebyshev filter with 0.1 dB ripple at 4 GHz (8), as well as the distance between resonators (see Fig. 10). This distance was calculated using commercial software based on coupled microstrip line circuit models and a simple look-up table technique (6) generated using a two-dimensional electromagnetic field solver as a tool to refine the results of the circuit-model-based software. The effectiveness of this technique was demonstrated experimentally (8). Notice in Table 3 that two of the three required couplings are less than -30 dB and that the error in estimating the coupled-line distances given by the circuit-model software increases as the coupling gets weaker. For narrower bandwidths, weaker couplings are needed, which would result in larger separation between resonators and hence increased difficulty in accurately predicting and controlling the required coupling. Bandwidths of less than 1% with this filter topology could probably be achieved with great difficulty and very low yield because of the practical issues associated with controlling the weak couplings required (35).

Workers in this field have realized this fundamental problem and have identified structures which allow significantly better control of weak interresonator coupling. Recognizing that in microstrip backward-coupled resonators, such as those in Fig. 10, the problem is compounded by the presence of spurious forward coupling, researchers have demonstrated good coupling control using forward-coupling alone in microstrip (36) and backward coupling alone in stripline (37). Also, the use of planar lumped elements (38), and inductive coupling in coplanar line (39,40) have been demonstrated successfully. Today, HTS filters are being made with bandwidths of 1% or less by dedicated commercial companies for the base-station wireless market as pre-production prototypes (19,41).

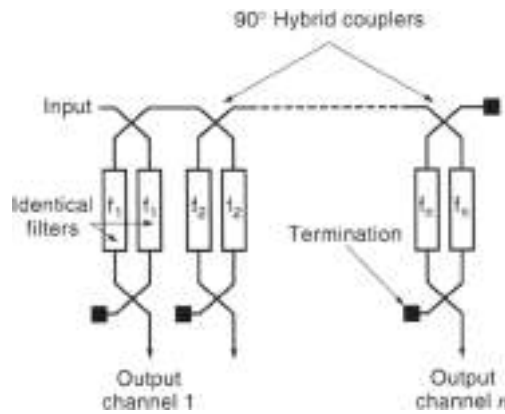
### Complex Structures

The potential for filters with performance similar to bulkier waveguide components but at significantly smaller sizes can be most readily fulfilled for the case of banks of filters, whether switched or multiplexed. Because HTS technology is planar, a relatively high level of integration is possible so that, as opposed to waveguide or dielectric resonator filter technology, a bank of  $N$  filters each occupying a volume  $V$  occupies a volume  $\approx N \times V$ , where  $V$  is significantly smaller than for conventional technologies for the same performance. HTS filter banks have been demonstrated by several groups in the form of bandpass multiplexers (8,42-44) or banks of switched band-reject filters (45).

**Table 3. Comparison Between Conventional and Look-Up Table Approaches (see Fig. 10)**

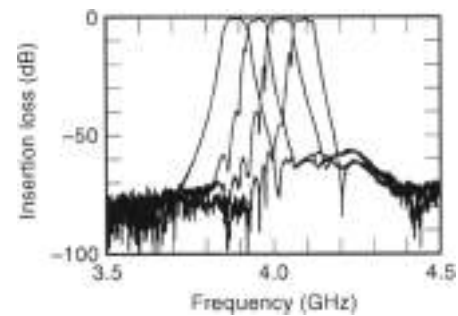
| Parameter | Required Coupling (dB) | Conventional (mm) | Look-Up Table (mm) |
|-----------|------------------------|-------------------|--------------------|
| $S_1$     | -17.6                  | 0.575             | 0.530              |
| $S_2$     | -35.8                  | 2.367             | 1.831              |
| $S_3$     | -37.8                  | 2.772             | 2.161              |

$$f_c = 4 \text{ GHz}; W = 0.176 \text{ mm}; l_1 = l_2 = l_3 = 1.788 \text{ mm}$$



**Figure 11.** Multiplexer architecture used to demonstrate a four-channel HTS microstrip device. This configuration has the advantage of allowing as many channels as the bandwidth covered by the 90° hybrid coupler. Each filter is terminated in 50  $\Omega$  and is essentially isolated from the others. Other schemes require that the impedance termination in each filter be adjusted to account for the presence of all the filters in the multiplexer, practically limiting the maximum number of channels to two or three.

**Example: Four-Channel Pass-Band Multiplexer.** Figure 11 shows a diagram of a multiplexer architecture (8,42). It can accommodate as many channels as the bandwidth of the 90° hybrid coupler covers. Input microwave energy is equally split at the first coupler. If the frequency is within the passband of the two identical filters connected to the coupler, it passes through the filters and adds in phase at the output port of the output hybrid coupler for Channel 1. If the frequency is not within the passband of the Channel 1 filters, the signal is reflected back to the input coupler where it recombines such that it is out-of-phase at the input port and in-phase at the input of the second channel hybrid coupler. The process then repeats itself until the signal exits the device through the appropriate channel port. Figure 12 presents details of one implementation of this device (8,46) showing one input, four outputs, and a through port

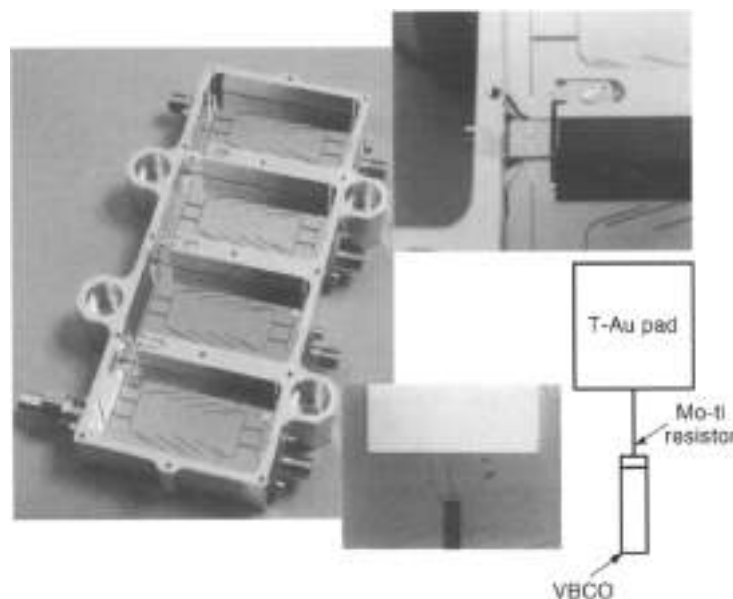


**Figure 13.** Measured response of the four-channel HTS multiplexer. Further details can be found in Refs. 8 and 46.

terminated in an external coaxial load. Additional channels could be connected to this port provided they are still within the bandwidth (about 10%) of the hybrid coupler used in this demonstration. Figure 12 also shows a detail of the assembly, which includes the internal HTS interconnections between filters and the integrated thin-film resistive terminations at the out-of-phase part of the output hybrid in each channel.

The HTS material used for this work was YBCO thin film deposited on 500  $\mu\text{m}$ -thick LAO substrates. The package included an aluminum frame holding the external coaxial connectors and aluminum carriers onto which the LAO substrate pieces were mounted. Niobium is a good thermal expansion match to LAO, so the electrical and mechanical integrity of the device was preserved when cycling from ambient temperature to near 77 K. The substrate-carrier assemblies were mounted on the aluminum frame using a beryllium-copper spring arrangement. Figure 13 shows the measured performance. The low-frequency skirts of Channels 2, 3, and 4 show some level of interaction between channels that can be eliminated using a wider guard-band between channels. Refs. 8 and 46 include a full discussion of the design, fabrication, and measurements on this device.

This unit was one of a series of demonstration devices delivered to the US Navy's High-Temperature Superconductivity



**Figure 12.** Photo montage of a four-channel YBCO-on-LAO microstrip multiplexer demonstrated under the US Navy's High-Temperature Superconductivity Space Experiment II (HTSSE-II). Details of the design, fabrication, and assembly of this device can be found in Refs. 8 and 46.

Experiment II program by several contractors for inclusion into the space package (47).

### High-Power-Handling Designs

The promising applications of HTS require that HTS filters handle sufficient signal power levels as to maintain linearity over a significant dynamic range. As explained earlier, current crowding along the edges of typical planar transmission lines (e.g., microstrip, stripline, and coplanar waveguide) ultimately limits the maximum power level that can be handled. Increasing the quality of the material and improving the design of filter structures has been a major endeavor at several institutions. Improved filter designs are based on planar structures which avoid the effects of significant current crowding at the edges, as is the case of low-impedance microstrip lines (48). Most significant is the work employing planar resonator structures based on the circular  $TM_{10}$  mode (29,49,50). The most salient features of this approach are shown in Figs. 14(a) and (b), which show the electromagnetic fields and current profile in a microstrip and a disk resonator, respectively. In the latter the RF magnetic fields do not close above the substrate but within it, under the disk. Thus the current density does not peak at the edges of the resonator and its distribution is more uniform. The only possible drawback of this approach is that the fields are more confined to the disk resonator, and intercoupling between resonators to form a filter may be more difficult, perhaps requiring three-dimensional structures for proper control of the coupling. This would eliminate some of the planar integration advantages. Fully planar filters using this concept, however, have been successfully demonstrated (29).

### HIGH-TEMPERATURE SUPERCONDUCTING DELAY LINES

Work on superconducting delay lines started at Lincoln Laboratory well before the advent of high-temperature supercon-

ductivity, and concentrated mostly on linearly dispersive delay lines for analog signal processing. Linearly dispersive delay lines have delay characteristics which vary linearly with frequency over a certain operating bandwidth and can be used to perform pulse compression, a technique to process and detect small signals which may be below the receiver noise floor (1). The pioneering work at Lincoln Laboratory in this area using LTS and, more recently, HTS thin-film technologies has been extensively documented in the literature (1,51).

Nondispersive delay lines have a constant delay versus frequency characteristic and are typically used as analog memory elements that can store a signal for, say, up to a few hundred nanoseconds while the system is engaged in other processing steps. Work on HTS nondispersive delay lines has also been significant (52-55). Including two recent instantaneous frequency measurement subsystems based on banks of delay lines (62,55). Clearly, the advantages of superconductivity are that a long length of line can be fabricated in a small volume by defining a long, planar transmission line on a wafer. Ref. 54 compares conventional nondispersive delay lines, which require amplifiers between sections of transmission line (e.g., coaxial), with HTS delay lines using projections based on measurements made on relatively short (22 ns) delay lines. Key delay-line parameters are delay, bandwidth, insertion loss, and third-order intercept point. Conventional delay lines that must resort to amplification to boost the signal are limited in dynamic range by the amplifiers.

### CRYOGENIC PACKAGING

Key to the insertion of superconducting microwave circuits into electronic systems is the integration of the HTS components with a cryogenic refrigerator and its associated control electronics. Clearly, for HTS technology to be ultimately successful, the

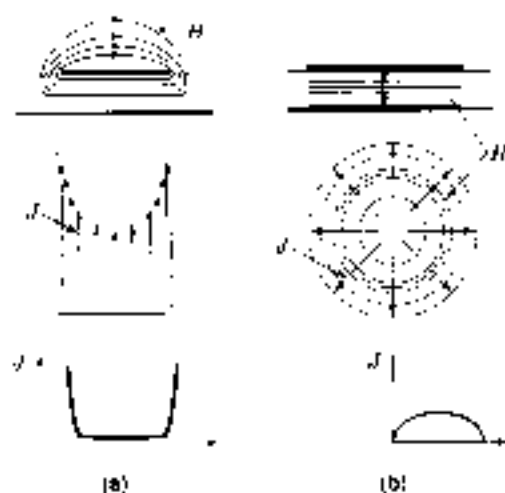


Figure 14. Diagram (courtesy of Dr. Z-Y Shen, E. I. du Pont de Nemours and Co.) (29), showing the magnetic field and current distribution in a  $TM_{10}$  microstrip resonator (a) and a  $TM_{10}$  printed disk resonator (b). In the disk the magnetic field lines are circular and remain on the plane of the disk, so the current is not highly nonuniform as is the case of the regular microstrip resonator. The advantage of the disk is that it can handle much higher power levels (29,49,50).

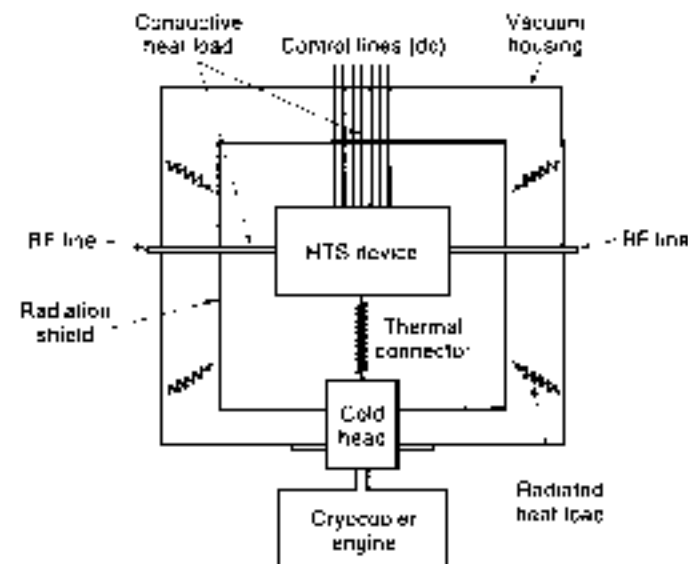


Figure 15. Schematic diagram of the cryogenic package for a hypothetical HTS device showing conducted heat inputs through input/output RF and dc control lines and the mechanical support of the cold head as well as the radiated heat input from the (warm) wall of the vacuum housing. The purpose of this diagram is to show the main elements that affect the design of the cryogenic package.

**Table 4. Sample System Requirements That Will Affect the Choice of Cooler and Cryogenic Packaging Approach**

| Requirement                             | Comments   |
|---|--|
| Size and weight                         | Stringent in almost all applications   |
| Cool-down time                          | Some applications may require very fast turn-in time (e.g., a few minutes). They would be a driver toward higher cooler power and lower HTS device thermal mass.   |
| Vibration                               | For example, a minute amount of mechanical disruption or a current caused by vibration from the cooler may generate a phase modulation that degrades the circuit performance.                                      |
| Power consumption and power supply type | E.g., 120 V ac   |
| Mode of operation                       | E.g., continuous, intermittent, short missions and then mostly idle, etc.  |
| Temperature stability and control       | While any fine temperature feedback control loop ( $\pm 0.01$ K) tends to be done using heaters and a temperature sensor, some applications may require a certain degree of cooling engine control ( $\pm 0.5$ K). |
| Unattended lifetime                     | Some applications (e.g., space) may require a lifetime on the order of 10 years or more.   |
| Vacuum lifetime                         | All-welded construction, use of getters in a clean, well-conditioned (baked) system.   |

user must be rendered able to ignore the fact that cryogenics are used at all, by providing long-lifetime cryocoolers and optimally small cryogenic packages with standard envelop characteristics and interfaces (e.g., 19-in rack mounts and back-plane blind-mate connectors).

Many important considerations enter into the design of a cryogenic package suitable for a microwave HTS subsystem. Figure 15 is a schematic representation of this package, showing its main elements and the various heat inputs that must be considered for an appropriate thermal design. Ref. 41 provides specific details on the cryogenic package for a communications filter subsystem.

The choice of a cryocooler will depend on the system and the cooling requirements. An airborne military application may require the use of a small Stirling-cycle cooler because of volume restrictions. On the other hand, a communications ground station in a remote location that needs to operate unattended for a long time may require a larger, more reliable refrigerator of the Gifford-McMahon type. Cooling requirements are imposed by the component or subsystem to be cooled and will determine the amount of cooling power required at the operating temperature. Typical sample system and cooling requirements and some comments as to their significance are given in Tables 4 and 5, respectively.

Cryocoolers likely to be used in microwave HTS technology will typically have from 1 W to 5 W of cooling capacity. A primary concern systems designers have is the reliability of cryogenic refrigerators, which varies greatly depending on their



**Figure 16.** Photograph of a HTS filter assembly for commercial wireless applications (courtesy of Superconductor Technology, Inc.).

type and size. Leveraging developments in other fields, such as infrared detectors, the reliability of small, military tactical cryocoolers has steadily increased in recent years, with some manufacturers claiming up to 20,000 h of mean-time-to-failure (MTTF). On the other hand, larger laboratory or industrial units and specialized coolers for aerospace applications operate for 5 years to 10 years and require minimal servicing. Table 6 lists some of the cryocooler types of interest. The intent here is not to be all-inclusive but to provide a basic reference to the type of coolers most likely to be employed in HTS microwave technology. Reference 50 is a good source of the latest developments in cryocooler technology. Figure 16 is a photograph of a

**Table 5. Cooling Requirements That Will Influence the Cooling Power (Heat Load) Required for a Given Application**

| Requirement                              | Comments   |
|--|--|
| Power dissipated in the device           | A filter with a 0.5 dB insertion loss that must pass a 50 W signal will dissipate 2 W of heat that must be removed by the cryocooler. Also, semiconductor devices such as low-noise amplifiers, which improve in noise and gain performance when cooled, always dissipate a certain amount of heat which must be taken into consideration.   |
| Number of microwave and dc control leads | These are the electrical interface between the cryocooled device and the outside world. For example, a filter might require two microwave leads (input and output) and two pairs of dc control lines for the heat sensor and a small heater to keep the temperature constant. These conductors represent a heat loss that the cooler must overcome because they connect the outside ambient temperature with the cold device. While the dc control lines are typically made of thin low thermal conductivity, high-resistivity wire (e.g., gauge 32 manganin), the microwave leads must achieve a compromise between insertion and thermal loss. |
| Surface area                             | Radiation loss is another form of heat loss that the cooler must overcome and therefore must be minimized. The total surface area and their infrared radiation emissivity are important design parameters. Low-emissivity radiation shields are typically used between the warm vacuum vessel wall and the cold device.  |
| Thermal mass                             | For those applications that have a cool-down time requirement, the thermal mass of the device to be cooled is important and will be affected by the microwave packaging material and its shape.  |

Table 6. Some Cryogenic Refrigerator Types Likely to Be Used in HTS Technology

| Cooler Type       | Heat-Lift Range Available at 80 K | Comments  |
|-------------------|-----------------------------------|---|
| Split Stirling    | 0.5–3 W                           | Available from many manufacturers; most prepared in the tactical military infrared detector industry. Has a cold head separated from a compressor by a metallic transfer line up to 15 cm long.                   |
| Integral Stirling | 0.5–5 W                           | Also used in infrared detectors; at least one version is being used in an HTS development prototype. The compressor and cold finger are integrated into one unit.   |
| Gifford-McMahon   | 2–2000 W                          | Widely used in the support of vacuum systems for semiconductor industry; highly reliable and versatile. The compressor and cold head are separate units connected by fluid lines that can be several meters long. |
| Throttle-cycle    | ~1 W                              | Reliable and low cost. The compressor and cold head are separate units connected by fluid lines that can be several meters long.  |
| Joule-Thomson     | 0.5–2 W                           | Generally used as an open-cycle cooling system for short tactical missile IR detector applications.   |
| Pulse tube        | 0.5–2 W                           | Emerging technology; low cold-head vibration and long life-time potential.  |

commercial HTS filter subsystem, showing the cryocooler and associated electronics in their open package.

## CONCLUSION

High-temperature superconductor microwave technology offers unique advantages derived from the low microwave loss of HTS materials and the inherent low thermal noise in cryogenically cooled components. The main applications to date are related to increased microwave receiver sensitivity, and this is most likely to have an impact on wireless military and commercial communications systems. The reason is that receiver sensitivity and dynamic range must be preserved in the presence of a large number of spurious signals which, if unfiltered, degrade receiver performance. Generation of clean transmitted signals requires filtering in the transmitter and this, coupled with the need to reject unwanted high-power signals at the receiver, has spurred work on high-power handling in HTS filters. Great interest in the United States and abroad exists in the wireless commercial communications market and several companies are testing base-station receiver front-ends consisting of cryogenically cooled filter-LNA subassemblies.

HTS microwave filters are therefore a promising technology, especially at frequencies below 3 GHz where the loss in conventional microwave materials force high performance filters to be very large in order to achieve the required low insertion losses and selectivity. Leveraging developments in infrared imaging detector technology and perhaps new developments of cooled semiconductor components for fast computer workstations, cryocooler technology is progressing to the point where long lifetimes and small-size, low weight coolers are now widely available.

## BIBLIOGRAPHY

1. R. S. Withers and R. W. Halton, Superconductive analog signal processing devices, *Proc. IEEE*, **77**: 1247–1263, 1989. This paper contains many references to further work by the authors.
2. P. A. Smith et al., YBCO thick films for high Q resonators, *IEEE Trans. Appl. Supercond.*, **7**: 1763–1766, 1997.
3. D. W. Foner et al., Large area YBaCuO films for high power microwave applications, *IEEE Trans. Appl. Supercond.*, **5**: 1581–1586, 1995.
4. A. Pique et al., Microwave compatible YBaCuO<sub>2-x</sub> films on ferrimagnetic garnet substrates, *Appl. Phys. Lett.*, **67**: 1778–1780, 1995.
5. E. J. Smith, J. Musolf, and E. Soares, Composition controlled MO/VO as a route to high Q HTS thin film devices, *Proc. 3rd Int'l IEEE/MRS Int'l Workshop Supercond. Suitable Mater. For HTS Appl. Towards Next Decade*, 1997, p. 145.
6. B. Uitz et al., Deposition of YBCO and NbCO films on areas of 9 inches in diameter, *IEEE Trans. Appl. Supercond.*, **7**: 1272–1277, 1997.
7. J. W. Ekin, A. J. Pansau, and B. A. Blankenship, Method for making low-resistivity contacts to high T<sub>c</sub> superconductors, *Appl. Phys. Lett.*, **62**: 331–333, 1993.
8. S. H. Jilisa et al., High-temperature superconducting sparse-cashed multiplexers and delay lines, *IEEE Trans. Microwave Theory Tech.*, **44**: 1229–1239, 1996.
9. R. K. Hoffmann, *Handbook of Microwave Integrated Circuits*, Norwood, MA: Artech House, 1987.
10. G. L. Matthaei, L. Young, and P. M. T. Jansen, *Microstrip Filters: Impedance Matching Networks, and Coupling Structures*, Dedham, MA: Artech House, 1990.
11. Z.-Y. Shien, *High-Temperature Superconducting Microwave Circuits*, Norwood, MA: Artech House, 1994.
12. S. Ramo, J. R. Whinnery, and T. Van Duzer, *Fields and Waves in Communication Electronics*, New York: Wiley, 1965.
13. I. Vendik and O. Vendik, in E. Kullberg (ed.), *High Temperature Superconductor Devices for Microwave Signal Processing*, St. Petersburg, Russia: Solodyn, 1997, 3 parts.
14. K. E. Carroll, J. M. Ford, and E. J. Cuknoskas, Superconducting kinetic-inductance microwave filters, *IEEE Trans. Appl. Supercond.*, **3**: 9–16, 1993.
15. Unpublished data obtained jointly by Northrop Grumman, NASA Lewis Research Center and Superconductor Technology, Inc.
16. W. W. Mumford and F. H. Scheibe, *Wide Performance Factors in Communication Systems*, Dedham, MA: Horizon House-Microwave, 1969. UMI Out of Print; Books on Demand.
17. S. H. Talisa et al., Dynamic range considerations for high-temperature superconducting filter applications to receiver front-ends, *IEEE MTT-S Int. Microw. Symp. Dig.*, 1994, pp. 897–1000.
18. J. B. Y. Tan, *Microwave Beamers with Electronic Warfare Applications*, Malabar, FL: Krieger, 1992.
19. G. Koepf, Superconductors improve coverage in wireless networks, *Microw. RF*, **37**: 40–63–74, 1998.
20. D. E. Oates et al., Nonlinear surface resistance in YBaCuO<sub>2-x</sub> thin films, *IEEE Trans. Appl. Supercond.*, **3**: 1114–1119, 1993.
21. P. Walker et al., Nonlinear effects in high-temperature



- superconductors: 3rd order intercept from harmonic generation, *IEEE Trans Appl Supercond*, **5**: 1666-1670, 1995.
22. M. Golusovsky, Physical mechanisms causing nonlinear microwave losses in high-T<sub>c</sub> superconductors, *8th Workshop RF Supercond*, Albano Terme, Italy, 1997, Invited Paper.
23. S. Sridhar, Non-linear microwave response of superconductors and a response of the critical state, *Appl Phys Lett*, **66**: 1054-1056, 1994.
24. J. H. Oates et al., A nonlinear transmission line model for superconducting stripline resonators, *IEEE Trans Appl Supercond*, **7**: 17-22, 1993.
25. D. E. Oates et al., Microwave power dependence of YBa<sub>2</sub>Cu<sub>3</sub>O<sub>7-x</sub> thin-film Josephson edge junctions, *Appl Phys Lett*, **68**: 705-707, 1996.
26. J. Vendik and O. Vendik, *High Temperature Superconducting Devices for Microwave Signal Processing*, St. Petersburg, Russia: Skladen, 1997, Part 2, Chap. 7.
27. B. S. Karasik et al., Subnanosecond switching of YBaCuO films between superconducting and normal states induced by current pulse, *J Appl Phys*, **77**: 4064-4070, 1995.
28. J. Vendik et al., The superconducting microwave devices based on S-N transition in HTS films, *27th Eur Microw Conf Proc*, 1997, pp. 909-914.
29. Z.-Y. Shen et al., Power handling capability improvement of high temperature superconducting microwave circuits, *IEEE Trans Appl Supercond*, **7**: 2446-2450, 1997.
30. Z. Ma et al., RF power dependence study of large area YBCO thin films, *IEEE Trans Appl Supercond*, **7**: 1911-1916, 1997.
31. T. Dahmi and D. J. Scalapino, Theory of intermodulation in a superconducting microstrip resonator, *J Appl Phys*, **81**: 2002-2008, 1997.
32. O. G. Vendik, J. B. Vendik, and T. B. Samulova, Nonlinearity of superconducting transmission line and microstrip resonator, *IEEE Trans Microw Theory Tech*, **45**: 173-178, 1997.
33. S. H. Talisa et al., Low- and high-temperature superconducting microwave filters, *IEEE Trans Microw Theory Tech*, **39**: 1444-1454, 1991.
34. W. G. Lyons et al., High-T<sub>c</sub> superconductive microwave filters, *IEEE Trans Magn*, **27**: 2537-2539, 1991.
35. W. G. Lyons and L. H. Lee, Accuracy issues and design techniques for superconducting microwave filters, *Comput Aided Des Supercond Microw Components Workshop, IEEE Int. Microw Symp*, 1994. The authors are with Lincoln Laboratory, Massachusetts Institute of Technology.
36. D. Zhang et al., Compact forward-coupled superconducting microstrip filters for cellular communications, *IEEE Trans Appl Supercond*, **6**: 2656-2659, 1995.
37. G. L. Matthaei and G. L. Ely-Shipton, Novel staggered resonator array superconducting 2.3-GHz bandpass filter, *IEEE Trans Microw Theory Tech*, **41**: 2345-2352, 1993.
38. D. G. Swanson, R. Forse, and H. J. L. Nelson, A 10 GHz thin film lumped element high temperature superconductor filter, *IEEE MTT-S Int. Microw Symp Dig*, 1992, pp. 1191-1195.
39. D. G. Swanson and R. Forse, An HTS end coupled CPW filter at 25 GHz, *IEEE MTT-S Int. Microw Symp Dig*, 1994, pp. 199-202.
40. A. Veng and W. Jeda, An HTS narrow bandwidth coplanar shunt inductively coupled microwave bandpass filter on LaAlO<sub>3</sub>, *IEEE Trans Microw Theory Tech*, **45**: 483-487, 1997.
41. M. J. Scheeren et al., Filter subsystems for wireless communications, *IEEE Trans Appl Supercond*, **7**: 3744-3749, 1997.
42. R. R. Mousam et al., Design considerations of superconductive input multiplexers for satellite applications, *IEEE Trans Microw Theory Tech*, **44**: 1213-1228, 1996.
43. C. Raucher, J. M. Paul, and G. B. Tao, Cryogenic microwave channelized receivers, *IEEE Trans Microw Theory Tech*, **44**: 1240-1247, 1996.
44. S. J. Pieczuszu et al., Low loss multiplexers with planar dual mode HTS resonators, *IEEE Trans Microw Theory Tech*, **44**: 1248-1257, 1996.
45. N. O. Fanzo et al., An optically switched bank of HTS bandpass filters, *IEEE MTT-S Int. Microw Symp Dig*, 1994, pp. 195-198.
46. S. H. Talisa et al., High temperature superconducting four-channel filterbanks, *IEEE Trans Appl Supercond*, **6**: 2079-2082, 1995.
47. T. G. Kawerki et al., The high temperature superconductivity space experiment (HTSSE-II) design, *IEEE Trans Microw Theory Tech*, **44**: 1198-1212, 1996.
48. G. C. Liang et al., High-temperature superconducting microstrip filters with high power handling capability, *IEEE MTT-S Int. Microw Symp Dig*, 1995, pp. 191-194.
49. H. Chaloupka et al., Superconducting planar disk resonators and filters with high power handling capability, *Electron Lett*, **32**: 1736-1738, 1996.
50. S. Kolesov et al., Planar HTS structures for high power applications in communication systems, *J Supercond*, **10**: 179-187, 1997.
51. W. G. Lyons et al., High temperature superconductive wideband compressive receivers, *IEEE Trans Microw Theory Tech*, **44**: 1258-1276, 1996.
52. G. C. Liang et al., Space-qualified superconductive digital instantaneous frequency-measurement subsystem, *IEEE Trans Microw Theory Tech*, **44**: 1289-1299, 1996.
53. N. Funzi et al., Development of high temperature superconducting 100 nanosecond delay line, *SPIE Proc*, **2168**: 143-151, 1994.
54. S. H. Talisa et al., High-temperature superconducting wide band delay lines, *IEEE Trans Appl Supercond*, **5**: 2291-2294, 1995.
55. M. Biehl et al., A 4 bit instantaneous frequency meter at 10 GHz with coplanar YBCO delay lines, *IEEE Trans Appl Supercond*, **6**: 2236-2242, 1995.
56. E. G. Ross, Jr. (ed.), *Cryocoolers: A Publication of the International Cryocooler Conference*, New York: Plenum, 1997.

SALVADOR H. TALISA  
Nucleon Corporation

## SUPERCONDUCTING LEVITATION

The levitation of a permanent magnet over a superconductor is one of the basic tests of superconductivity, and it is a sight that has inspired the investigation of applications not possible with any other technology [1-5]. In this article, I briefly review the fundamental physics of superconductor levitation, discuss basic levitational phenomena and the features of superconductor levitation pertinent to bearings, and mention some possible applications of superconductor levitation, with emphasis on high efficiency flywheel energy storage.

### STABLE LEVITATION

In its simplest form, a superconducting levitational system comprises a permanent magnet levitated in a stable position over a superconductor. This stability is in stark contrast to most magnetic systems, which are statically unstable. Earnshaw showed that there is no stable, static 3-D arrangement of a collection of poles (electric, magnetic, or gravitational) whose magnitudes do not change and which interact via a  $1/r^2$  force

law (6). Braunbek extended the result to show that no stable static configuration exists when paramagnetic or ferromagnetic material is included in the arrangement (7). These results collectively are often referred to as Earnshaw's theorem. Earnshaw's theorem is grasped intuitively by most people when they release a permanent magnet next to the ferromagnetic door of their refrigerator. The magnet moves to stick to the door, or it falls on the floor. It does not hover in space near the point where it was released.

Diamagnetic materials, such as superconductors, are not governed by Earnshaw's theorem, and they enable the possibility of creating stable levitation systems. One of the fundamental properties of superconductors is their tendency to exclude magnetic flux from their interiors. This exclusion of magnetic flux (the Meissner effect), makes them behave like strong diamagnets. Accordingly, a superconductor with a permanent magnet positioned close above it, as shown in Fig. 1(a), develops a shielding current, which excludes flux so that the actual magnet "sees" its mirror image.

More specifically, if the magnetization of a permanent magnet is in the vertical direction, with its north pole down, the image is also vertically magnetized, but with its north pole up,

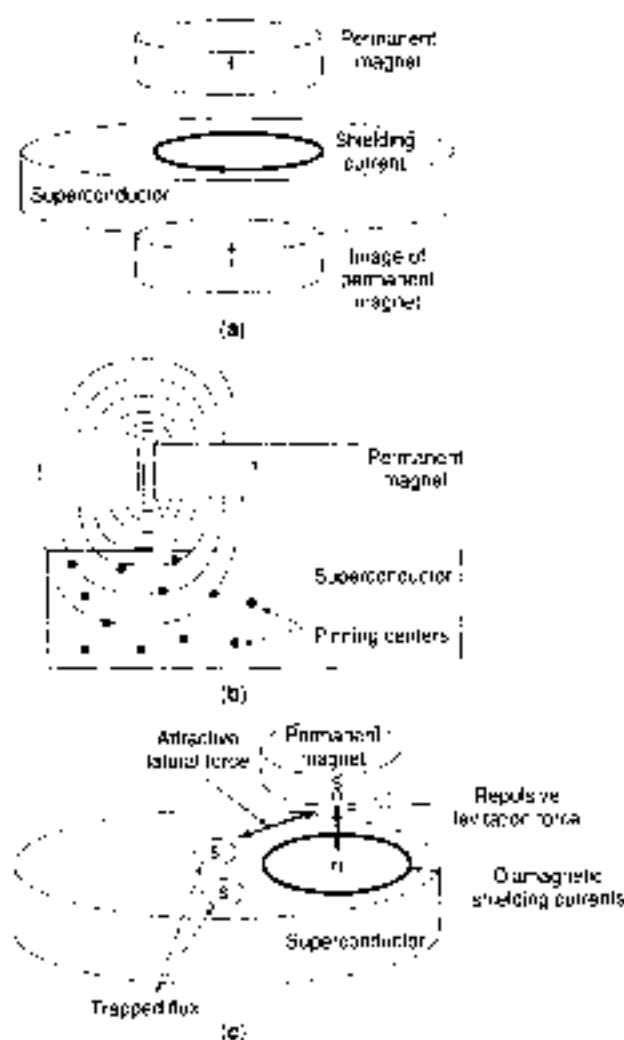


Figure 1. Schematic diagrams of levitation basics: (a) diamagnetic response; (b) flux pinning; (c) flux trapping.

exerting a repulsive force on the real magnet. The closer the magnet moves to the superconductor, the stronger the repulsive force. The farther away the magnet moves, the weaker the force. In fact, if the magnet flips over so that the opposite (south) pole faces the superconductor, the screening currents in the superconductor will reverse so as to reverse the magnetization of the image in the superconductor, maintaining the repulsive interaction. This setup yields levitation stable in the vertical direction. Horizontal stability is obtained if the superconductor is given a concave shape, so that vertical superconducting walls are formed around the magnet, as first demonstrated by Arkadiev using lead, a Type I superconductor, in which magnetic flux is totally excluded from the interior (8).

From a technological viewpoint, the most useful superconductors are usually Type-II superconductors, in which, above a first critical field,  $H_{c1}$ , it is energetically favorable for magnetic flux to enter the interior of the superconductor in discrete localized regions that become normal (i.e., not superconducting) with each region of flux surrounded by a small vortex of superconducting shielding current. In Type-II superconductors, the stability of the levitational phenomena resulting from the diamagnetic response is greatly enhanced by the additional phenomena resulting from flux pinning, shown in Fig. 1(b). A flux pinning center is a nonsuperconducting region, such as an inclusion, crack, or other crystalline defect. Because the superconducting region surrounding the nonsuperconducting center is strongly inclined to exclude magnetic flux, a magnetic flux line through the center often becomes trapped there. When a sufficient number of flux lines is trapped in the superconductor, the permanent magnet remains levitated in position, even over a flat surface. The flux lines between the permanent magnet and the superconductor act in an imperfect analogy to mechanical springs with attachments on the permanent magnet and in the superconductor. If the magnet moves a small distance laterally, so that the flux lines remain in their pinning centers, the flux lines bend and produce a laterally restoring shear force, according to the Maxwell electromagnetic stress tensor. If the magnet moves vertically or horizontally, the "springs" pull the magnet back to its equilibrium position. If the flux pinning is sufficiently strong, the magnet is stably suspended below the superconductor (9) or even along its side (10).

If the magnet moves far enough laterally that the flux lines move from their original pinning centers to new ones, then an additional stabilizing force, involving trapped flux, begins to act. Trapped flux consists in regions of induced magnetization in the superconductor of the same pole orientation as the levitated magnet and results from movement of flux lines from their pinning centers that decreases the local flux. As shown in Fig. 1(c), this results in an attractive interaction that reduces the levitation force but provides a lateral restoring force.

## SUPERCONDUCTOR LEVITATORS

Interest in the potential of superconductor levitation in various applications greatly increased with the discovery of superconductors whose critical temperatures (i.e., temperatures at which they transit from the normal state to the superconducting state) exceeded the boiling point of nitrogen. Although one could create a superconducting wire magnet for levitation, most of the present efforts involve the use of bulk superconductors or thin-film superconductors. Unlike superconducting wire

applications, in which the supercurrent must pass from grain to grain along quite a distance, the supercurrent for levitation applications needs to circulate only within individual grains.

The present material of choice for superconducting levitation is Y-Ba-Cu-O (YBCO), because it exhibits a high magnetic irreversibility field at liquid nitrogen temperatures and has the ability to grow large grains. In addition to the two temperature-dependent phase-transition fields,  $H_{c1}$  and  $H_{c2}$ , all superconductors have a magnetic irreversibility field,  $H_{irr}$ , that lies between  $H_{c1}$  and  $H_{c2}$ .  $H_{irr}$  is the field at which the magnetization  $M$  as a function of applied field  $H$  is no longer double-valued (11). For the low-temperature superconductors NbTi and Nb<sub>3</sub>Sn,  $H_{irr}$  is extremely close to  $H_{c2}$ , and there is no important distinction between them. At higher temperatures, thermal activation is much greater, which leads to easier flux motion near  $H_{c2}$  for HTSs.  $H_{irr}$  marks a phase-transition between the region where magnetic flux is solidly pinned in the superconductor and the region where flux may move. Sometimes the curve is said to denote the boundary between the region where flux is frozen and the region where flux is melted. Of all the known HTSs, YBCO has a relatively low critical temperature of 92 K, but it has the highest irreversibility curve at 77 K and lower temperatures. For stable levitation, it is important that the flux be frozen in the superconductor. Otherwise, the permanent magnet would slowly lose levitation height.

The magnetization of the superconductor is proportional to the product of the critical current density and the grain diameter. Large grain diameters are important to achieve sufficiently large magnetizations for useful levitation forces. In bulk materials, the grains grow to diameters of several centimeters when made by a melt-texturing process (12). In the present state of the art, the upper limit of the grain diameter produced by this process is about 10 cm. The ability to produce good-quality YBCO thin films is also limited to about this size.

## LEVITATIONAL PHENOMENOLOGY

If the permanent magnet is pulled hard enough to the side or vertically, it is possible to move one or more of the trapped flux lines into new pinning centers and so change the equilibrium position. Such a change results in a hysteretic effect in the levitational force and an associated energy loss if it occurs in a cyclic pattern. To explore the hysteretic effect, it is convenient to divide the behavior into two processes. The first is called field-cooled. It occurs when the superconductor is cooled below its critical temperature while there is a substantial magnetic field present, that is, the permanent magnet is close to the superconductor. Field cooling produces less repulsive levitation force but can be used to make an attractive-force bearing. The second is called zero-field-cooled. It occurs when the superconductor is cooled below its critical temperature in the absence of a magnetic field, that is, when the permanent magnet is far from the superconductor. Zero-field cooling results in the largest repulsive force but may be practically inconvenient because it requires the cooling of the superconductor prior to the assembly of the bearing.

The hysteretic nature of the levitational phenomenon for movements in the vertical direction is illustrated in Fig. 2. In this example, a cylindrical, vertically magnetized permanent magnet was kept with its bottom surface at a height of 10.0 mm above the top surface of a cylindrical YBCO superconductor

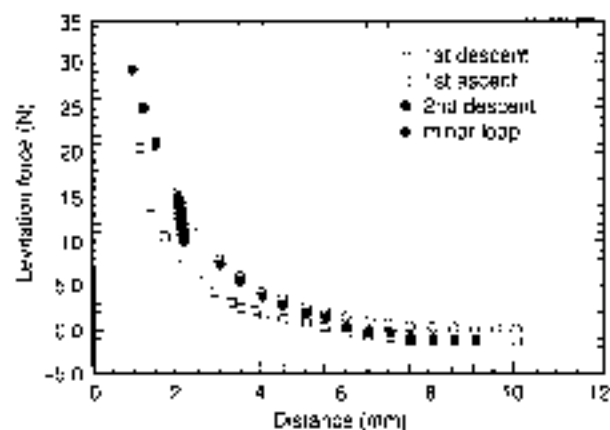


Figure 2. Levitation force versus distance between the permanent magnet and superconductor.

while the YBCO was cooled, essentially a zero-field-cooled condition. Then the magnet was slowly brought down to a position 1 mm above the superconductor, and the force was measured at various points along this first descent. Then the magnet was moved away from the superconductor, back to its original zero-field-cooled position. Then it followed a second descent during which a minor reversal of 0.4 mm was made at a distance of 2.0 mm. A second ascent was identical with the first, and a third descent was identical with the second, etc. As seen in Fig. 2, the force during the first descent is always larger than the force during the second. Upon reversal from 1 mm during the first ascent, the force drops very quickly and even becomes negative, indicating an attractive force. A combination of the first ascent and second descent forms a major hysteresis loop, and the area under the curve is equivalent to the hysteretic energy loss. The width of the minor loop is much smaller than that of the major loop. Thus, the superconductor acts as a nonlinear damper, and the damping coefficient increases with amplitude. From Fig. 2, it is also clear that the slope of the minor loop, which represents the magnetomechanical stiffness of the system, is considerably higher than that determined from the major loop.

A feature of the first descent is that over several millimeters above the superconductor surface, the force is exponential with distance, as shown in Fig. 3. In practice, it is difficult to measure the force immediately above the superconductor, partly because surfaces are not flat, but mainly because the surface is usually covered with liquid nitrogen. The exponential behavior shown in Fig. 3 allows extrapolating the force to the surface, and various superconductors can be compared this way by using a permanent magnet of known strength.

In the examples shown in Figs. 2 and 3, the superconductor was cooled while the permanent magnet was far from its surface, that is, zero-field-cooled. In the field-cooled case, the levitational force is approximately zero after the superconductor is cooled below its critical temperature. However, the magnetomechanical stiffness is approximately the same as for the zero-field-cooled case at the same separating distance. The stiffness is dependent on the amplitude of vibration (13,14). For small amplitudes, the stiffness is constant, but, after some critical amplitude that depends on height, the stiffness begins to decrease as the amplitude increases. The higher the current density of the superconductor, the higher the critical value for the onset of stiffness decrease.

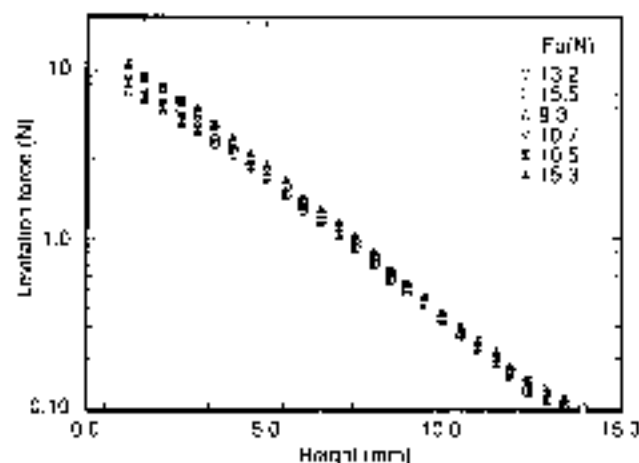


Figure 3. Levitation force versus distance on first descent between reference magnet and several superconductors, showing the levitation force extrapolated to zero height.

The levitational force that the superconductor provides is proportional to its average magnetization, which is proportional to the product of its grain diameter times its critical current density. The critical current density in typical melt-textured YBCO samples at 77 K is about 40 kA per  $\text{cm}^2$ , which, together with a diameter of several  $\mu\text{m}$ , allows the levitational pressures between the YBCO and an NdFeB permanent magnet to be as high as about 280 kPa (40 psia). Such pressures have been measured in the author's laboratory with the very best YBCO samples. However, this pressure occurs at zero separation distance. In a practical system with a finite separation, the pressure will be at least a factor of 2 to 3 times lower.

Thin-film YBCO often has critical current densities that exceed 1 MA per  $\text{cm}^2$ . Because the thickness of these films is only about 1 micron, they do not provide much levitational force. However, the stiffness of such films is often of the same order as that produced by bulk materials of greater thickness (15).

## THEORETICAL MODELS

Various theoretical models account for different aspects of superconductor behavior when a permanent magnet approaches. In many cases in which hysteretic behavior is not important, the superconductor is a pure diamagnet and the magnet is a set of magnetic dipoles (16). A "frozen" mirror image may be used in conjunction with the diamagnetic mirror image to describe the elastic properties and energy loss in a field-cooled system (17). The Bean model, which assumes that the current circulating in the superconductor is either at its critical value or zero, is often used in efforts to address the hysteretic behavior (16).

## SUPERCONDUCTING BEARINGS

If the azimuthal inhomogeneity of the magnetic field of the permanent magnet is high, for example, if the magnet is a cylinder with uniform magnetization throughout, the levitated magnet rotates freely above the superconductor. As long as the distribution of magnetic flux in the superconductor does not change, rotation encounters no resistance. If the magnet is spinning, the

hysteretic loss in the superconductor decreases the rotational rate.

In an electromechanical system, such as a magnetic bearing, the parameters of interest are the levitational force, stiffness, damping, and rotational loss. The 280 kPa levitational pressure is lower than that achievable in a conventional electromagnetic bearing ( $\sim 1$  MPa) and significantly lower than that typically achieved in mechanical roller bearings ( $\geq 10$  MPa). The amount of mass levitated directly depends on the number and size of permanent magnets and superconductors available. In the present early period of technological development for superconducting bearings, several laboratories have stably levitated masses greater than 100 kg.

In practical superconductor bearings, the low levitational pressure available in the interaction between the permanent magnet and the superconductor is often augmented by various hybrid schemes in which interactions between pairs of permanent magnets provide the bulk of the levitational force. These interactions are unstable, as Earnshaw's theorem predicts, but the inclusion of a properly designed superconducting component in the bearing is sufficient to stabilize the complete bearing. Augmentation takes the form of an Evershed-type design, in which a pair of permanent magnets is in attractive levitation, employs permanent magnets in repulsive levitation, or uses active magnetic bearings (5).

The hysteretic nature of a superconducting bearing also makes damping of translational motion amplitude-dependent. For low-amplitude vibrations, damping is small, but quickly increases as the vibrational amplitude increases. This hysteretic nature of the HTS bearing thus contributes to the robustness of the system. The hysteretic nature also results in a larger uncertainty of the equilibrium position of the rotor than is typical in most rotating machinery. This uncertainty requires larger running gaps between moving and stationary parts.

The ease with which a permanent magnet spins, when levitated over a superconductor, and the absence of contact between the surfaces, produce the illusion that the rotation is frictionless. In reality, small magnetic losses gradually slow the rotation. The losses are primarily the result of azimuthal inhomogeneities in the magnetization of the permanent magnet, which produce hysteretic loss in the superconductor. Typically, in permanent magnets with the best homogeneities, at a fixed radius above the rotating surface, the amplitude of the ac component of the magnetic field is of the order of 1% of the average field at that radius. Although small, this inhomogeneity is sufficient to cause a detectable decay in rotational rate when the magnet spins in vacuum.

A figure of merit for the rotational decay of a bearing is the coefficient of friction (COF), defined as the rotational drag force divided by the levitational force (weight of the levitated rotor). The COF for a mechanical roller bearing is of the order of  $10^{-4}$ . The COF for an active magnetic bearing is about  $10^{-4}$  when parasitic losses for the feedback circuits and power for the electromagnets are factored in. Measured COFs for simple superconductor bearings are as low as  $10^{-5}$ . The parasitic losses of a superconducting bearing are the power required to keep the superconductor cold. For refrigerators that operate at about 30% of Carnot efficiency (the theoretical maximum), about 14 W of electricity are required to remove 1 W of heat at liquid nitrogen temperatures. Thus, the equivalent COF for an HTS bearing is about  $2 \times 10^{-6}$ , about two orders of magnitude lower than the best alternative bearing.

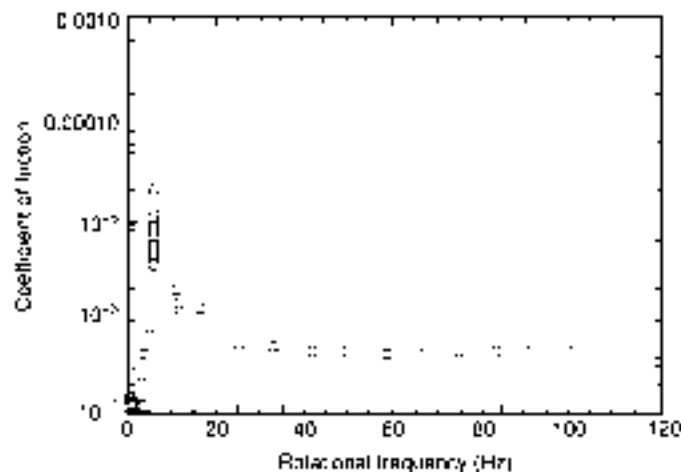


Figure 4. Coefficient of friction versus rotational frequency for a 25.4 mm dia., 9.8 mm high cylindrical permanent magnet levitated 10 mm above a YBCO cylinder.

Two magnetomechanical resonances occur in a magnetically suspended rotor with a polar moment of inertia greater than the transverse moment (i.e., a disk geometry): one is vertical and one is radial. In practice, the vertical resonance has a minimal effect on the COF in most superconducting bearing systems. The radial resonance occurs when the rotational frequency is close to that of the natural radial frequency of the rotor's vibration. If the rotor has a transverse moment of inertia greater than its polar moment, then there is an additional resonance, having the form of a conical vibration, that is, with the top of the rotor moving to one side while the bottom of the rotor moves to the opposite side. Because superconducting bearings have low stiffness, the resonances occur at low frequencies on the order of several hertz. This, together with the large clearances possible with superconducting bearings, leads to a robust bearing system.

Figure 4 shows the COF as a function of rotational frequency for a cylindrical permanent magnet levitated 10 mm above a single YBCO cylinder and spinning in a vacuum chamber. One may divide the behavior into three regions: below the resonance, the radial resonance, and above the resonance. Below the resonance ( $f < 3$  Hz), the losses are caused by the inhomogeneity of the permanent magnet's field. The resonance region (3–20 Hz), shown in Fig. 4, is relatively broad. In some systems, especially those with well-balanced rotors, the resonance is very narrow. Above the resonance ( $f > 20$  Hz), the losses are affected by an additional factor, which is caused by the rotation of the magnet about its center of mass rather than its center of geometry or center of magnetism. As shown in Fig. 4, above and below the resonance, the losses are mostly velocity-independent. However, detailed studies of losses in superconducting bearings show that small velocity-dependent effects are present which are intrinsic to the superconductors (19–21).

Because of size limitations encountered when high-performance bulk superconductors are produced, a large bearing system needs an array of superconductors. In the case of a single magnet levitated over an array of superconductors, an additional loss arises from magnetization of the individual superconductors upon levitation. The magnetization of the array leads to an  $\mu_0$  magnetic field seen by the rotating permanent

magnet and eddy current losses that depend on the electrical conductivity of the permanent magnet.

Some alternative bearing concepts that involve superconductors exhibit even lower COFs. The Evershed-type hybrid has a COF of just over  $10^{-4}$  (22). The velocity-dependent losses associated with superconducting arrays are also greatly reduced with this bearing design. A mixed- $\mu$  (where  $\mu$  is the magnetic permeability) bearing (23) has a COF of just over  $10^{-4}$ . In this bearing, a soft ferromagnetic cylinder ( $\mu > 1$ ) is levitated in attractive levitation between two permanent magnets and stabilized by a superconductor ( $\mu < 1$ ) placed between the rotor and each of the permanent magnets.

## APPLICATIONS

The availability of superconducting bearings that are so nearly friction-free naturally leads to their consideration for flywheel energy storage. Flywheels with conventional bearings typically experience high-speed idling (i.e., no power input or output) losses of the order of about 1% per hour. With superconducting bearings, it is believed that losses as little as 0.1% per hour are achievable. When coupled with efficient motors, generators and power electronics (capable of losses as low as 4% on input and output), the potential exists for constructing flywheels with diurnal storage efficiencies of 90%. Probably only one other technology is capable of achieving such high diurnal storage efficiencies: large superconducting magnetic energy storage, which employs superconducting coils hundreds of meters in diameter.

Electric utilities have a great need for inexpensive energy storage, such as flywheels, because their inexpensive baseload capacity is typically underutilized at night and they must use expensive generating sources to meet their peak loads during the day. A distributed network of diurnal-storage devices could also make use of underutilized capacity in transmission lines at night and add robustness to the electric grid. These factors are expected to become more important in the coming deregulation of the electric utility industry. Efficient energy storage would also be beneficial to renewable energy technologies, such as photovoltaics and wind turbines.

With modern graphite-fiber/epoxy materials, the inertial section of a flywheel rotates with rim speeds well in excess of 1000 m/s and achieves energy densities greater than those of advanced batteries. The kinetic energy in a large Frishonized flywheel with this rim speed is about 1 kWh, and a personalized flywheel could store 20–40 kWh. Although design concepts for flywheels that employ superconducting bearings with up to 10 MWh have been proposed, the most advanced experimental versions at present are in the 100 Wh to 1 kWh class.

Because superconducting levitation is versatile over a wide range of stiffness and damping, it has been suggested for numerous applications. Superconducting bearings, like magnetic bearings, do not require a lubricant, which could be a major advantage in harsh chemical or thermal environments. Superconducting bearings are particularly interesting for cryogenic turbopumps. The low friction of the superconducting bearing allows its use in high-precision gyroscopes and gravimeters. The hysteretic nature of superconducting levitation has suggested its use in docking vehicles in space. As one vehicle approaches another it would experience a decelerating repulsive force. After the relative velocities have disappeared at a small separation distance, the vehicles would experience an attractive

force if their distances tend to separate. The stable levitational force suggests application in magnetically levitated conveyor systems in clean-room environments where high purity requirements mandate no mechanical contact. Trapped-field HTSs have been suggested for constructing vehicle magnets to be used in electrodynamic levitation of high-speed trains (see MAGNETIC LEVITATION).

## BIBLIOGRAPHY

1. F. C. Moon, *Superconducting Levitation*, New York: Wiley, 1994.
2. J. R. Hull, guest editor, *Applied Superconductivity*, 2 (7-8): 1994. Contains several state-of-the-art papers on superconducting levitation and HTS flywheels.
3. T. D. Bessing and J. R. Hull, Magnetic levitation, *The Physics Teacher*, 29 (9): 552-552, 1991.
4. E. H. Brandt, Rigid levitation and suspension of high-temperature superconductors by magnets, *Am. J. Phys.*, 58: 43-49, 1990.
5. J. R. Hull, Flywheels on a roll, *IEEE Spectrum*, 20-25, July 1997.
6. S. Earnshaw, On the nature of molecular forces which regulate the constitution of luminous ether, *Trans. Cambridge Philos. Soc.*, 7: 98-112, 1842.
7. W. Braunrück, Freischwebende Körper im elektrischen und magnetischen Feld, *Zeit. Physik*, 112: 753-763, 1939.
8. V. Arkadiev, A floating magnet, *Nature*, 160: 333, 1947.
9. P. N. Peters et al., Observation of enhanced properties in samples of silver oxide doped YBa<sub>2</sub>Cu<sub>3</sub>O<sub>x</sub>, *Appl. Phys. Lett.* 62: 2066-2067, 1993.
10. W. G. Harner, A. M. Hermann, and Z. Z. Sheng, Levitation effects involving high T<sub>c</sub> thallium based superconductors, *Appl. Phys. Lett.* 53: 1119-1121, 1988.
11. T. P. Shahan, *Introduction to High-Temperature Superconductivity*, Chap. 14, New York: Plenum, 1994.
12. M. Murakami, *Melt-Processed High-Temperature Superconductors*, Singapore: World Scientific, 1992.
13. S. A. Nasinger, J. R. Hull, and T. M. Mulcahy, Amplitude-dependence of magnetic stiffness in bulk high-temperature superconductors, *Appl. Phys. Lett.* 57: 2342-2344, 1990.
14. J. R. Hull et al., Magnetic levitation and stiffness in melt-textured Y-Ba-Cu-O, *J. Appl. Phys.* 72: 2069-2071, 1992.
15. P. Schönhuber and F. C. Moon, Levitation forces, stiffness and force-creeep in YBCO high T<sub>c</sub> superconducting thin films, *Appl. Supercond.* 2: 523-534, 1994.
16. Z. J. Yang, Lifting forces acting on magnets placed above a superconducting plane, *J. Supercond.* 5 (3): 259-271, 1992.
17. A. A. Kordyuk, Magnetic levitation for hard superconductors, *J. Appl. Phys.* 68 (11): 610-612, 1998.
18. A. B. Riise, T. H. Johansen, and H. Bratsberg, The vertical magnetic force and stiffness between a cylindrical magnet and a high-T<sub>c</sub> superconductor, *Physica C*, 234: 105-114, 1994.
19. Z. J. Yang and J. R. Hull, Energy loss in superconducting bearing systems, *IEEE Trans. Appl. Supercond.* 7: 317-321, 1997.
20. A. A. Kordyuk and V. V. Semushikalenko, High-speed magnetic motor with HTS bearings for precision energy losses investigations, *IEEE Trans. Appl. Supercond.* 7: 926-931, 1997.
21. C. E. Rasman, J. J. Budnick, and B. R. Wernberger, Correlation of frictional losses of spinning levitated magnets with air-sealability in high-temperature superconductors, *Appl. Phys. Lett.* 70: 253-257, 1997.
22. J. R. Hull, T. M. Mulcahy, and J. F. Labarsch, Velocity dependence of rotational loss in Evenshed type superconducting bearings, *Appl. Phys. Lett.* 70: 653-657, 1997.
23. J. R. Hull et al., Low friction in mixed-mu superconducting bearings, *J. Appl. Phys.* 78: 6623-6638, 1995.

JOHN R. HULL  
Argonne National Laboratory

## SUPERCONDUCTING MAGNETS FOR FUSION REACTORS

The magnetic confinement of plasma is the most promising option to use controlled nuclear fusion as a power source for future generations. A number of different magnetic field configurations have been proposed to achieve plasma ignition, all requiring high field strength over a large volume. Most of the experimental machines use conventional, copper windings operated in pulsed mode, to investigate the plasma physics. The advanced plasma experiments, as well as the future fusion reactors, call for long confinement time and high magnetic field, which can be reasonably maintained only by superconducting coils.

Unlike other applications of superconductivity, for fusion magnets there is no "normal conducting" alternative: whenever a magnetic confinement fusion power plant will operate, it will have superconducting windings. For this reason, fusion magnets are an important, long-term factor in the market of superconducting technology. Today, for NbTi based conductors, fusion is a nonnegligible share of the market, with over 50 t of strand recently used for the LHD and about 40 t committed for W7-X. For Nb<sub>3</sub>Sn technology, two large devices, the T-15 tokamak and the ITER model coils, have used most of the conductor ever produced (about 25 t of strand), being the driving input for the development of high performance Nb<sub>3</sub>Sn strands.

The first use of superconducting coils in experimental fusion devices dates back to the mid-1970s. In the last twenty-five years, six sizable devices for magnetic plasma confinement have been built with superconducting coils (see Table I): T-7 and T-15 in the former Soviet Union, MFTF in the United States, TRIAM and LHD in Japan, and Tore Supra in France. In Germany, Wendelstein 7-X is under construction. Moreover, a number of developmental and prototype coils have been tested in the scope of large international collaborations (large coil task, demonstration poloidal coils, ITER model coils).

The operating requirement for fusion magnets may vary over a broad range, depending on the kind of confinement and the size of the device (i.e., for example, from medium-field, pure de mode in the helical coils of the stellarators, to the high-field, fast rate in the central solenoid of the tokamaks). There is no general recipe for the magnet design, but a few common issues can be identified. Long-term reliability calls for a conservative component design and generous operating margins. The maintenance by remote handling in a nuclear environment imposes strong restrictions to either repair or replacement of individual parts. Safety regulations are also a major issue for superconducting magnets in a fusion reactor: the design must account for any likely or less likely failure mode of the coil system and provide that it will not turn into a nuclear-grade accident. Last but not least, the cost of the magnets, which is a large fraction of the reactor cost, must be contained to be commercially competitive with other power sources.

Only low-temperature superconductors have been considered to date for use in fusion magnets at field amplitudes up

Table 1. Summary of Superconducting Magnet Systems for Fusion Devices

|                        | Stored Weight (t) | Conductor Cooling              | Stored Energy (MJ) | Peak Field (T) | Operating Current (kA) |
|------------------------|-------------------|--------------------------------|--------------------|----------------|------------------------|
| Tokamak T-7            | 1                 | NbTi/FF                        | 30                 | 5              | 6                      |
| Tokamak T-15           | 25                | Nb <sub>3</sub> Sn/FF          | 730                | 9.5            | 6.6                    |
| MPFF (all coils)       | 74                | Nb <sub>3</sub> Sn (NbTi)/pool | 1 000              | 9-12.75        | 1.5-5.9                |
| TRIAM                  | 2                 | Nb <sub>3</sub> Sn/pool        | 76                 | 11             | 6.2                    |
| Tore Supra             | 43                | NbTi/pool 1.8 K                | 600                | 9              | 1.4                    |
| LHD Helical (2 coils)  | 16                | NbTi/pool 4.5/1.8 K            | 900 (1 650)        | 6.8-9.2        | 19-17.3                |
| LHD Poloidal (6 coils) | 43                | NbTi/FF                        | 1 200              | 5-6.5          | 20.8-31.25             |
| Wendelstein 7-X        | 37                | NbTi/FF                        | 600                | 6              | 16                     |

FF: forced flow

\* Operation at superfluid helium is passed at a later stage

Design values achieved on single coil test

to 13 T. A substantially higher field, which would make attractive the use of high-temperature superconductors, is not likely to be proposed as the electromagnetic loads, roughly proportional to the product of field, current and radius, already set a practical limit for structural materials. It may sound surprising that the actual superconducting material cross-section is mostly smaller than 5% of the overall coil cross-section. The choice between NbTi and Nb<sub>3</sub>Sn conductors is dictated by the operating field. The upper critical field of NbTi conductors is  $\approx 10$  T at 4.5 K and  $\approx 13$  T at 1.8 K. According to the design current density and the temperature margin, the operating field is set at least 3 T to 4 T below the upper critical field. In the conservatively designed fusion magnets, the peak field for NbTi conductors is up to  $\approx 9$  T for coils cooled by a superfluid helium bath (e.g., Tore Supra and LHD helical coils) and up to  $\approx 6$  T for supercritical helium forced flow (e.g., W7-X and LHD poloidal coils). At a higher operating field, the choice of Nb<sub>3</sub>Sn conductors is mandatory to obtain adequate temperature margins and high current density. The increasing confidence in Nb<sub>3</sub>Sn technology, as well as its slowly decreasing cost, tends to move down the field threshold for the NbTi versus Nb<sub>3</sub>Sn. Conductors based on Nb<sub>3</sub>Al are in a developmental stage and may become an alternative to Nb<sub>3</sub>Sn for selected high-field magnets (e.g., the D-shaped toroidal field coils, because of the better tolerance to bending strain).

The winding packs may be either potted in epoxy resin or laid out as a spaced matrix of noninsulated conductors in a liquid helium bath. This last option offers the advantages of constant operating temperature and potential high stability due to the bath-cooled conductor surface. The drawbacks are the poor stiffness of the winding and the limited operating voltages (the insulation relies on the helium as dielectric). The potted coils with forced-flow conductors have superior mechanical performance and may operate at higher voltage, as a rule of thumb, they become a mandatory option for stored energy in excess of 1 GJ to 2 GJ. The cable-in-conduit conductors became, in the last decade, the most popular option for forced-flow conductors because of the potential low  $\mu$  loss and the good heat exchange due to the large wet surface. To withstand the mechanical and electromagnetic loads, the coils are fitted in thick-walled steel pipes, either welded or bolted. More structural material may be added, if necessary, both in the conductor cross-section and in winding substructures, for example, plates and cowound strips.

The magnetic stored energy is very large, up to 130 GJ for the proposed magnet system of ITER. In case of a quench (local

transition from superconducting to normal state), the stored energy must be dumped into an outer resistor to avoid an overheating and damage of the winding. A large operating current is needed to reduce the number of turns, that is, the winding inductance, and extract quickly the stored energy at a moderately high voltage (up to 10 kV to 20 kV). The operating current density in the superconducting cross-section (NbTi or Nb<sub>3</sub>Sn filaments),  $J_{op}$ , is selected according to the specific design criteria to be a fraction of the critical current density,  $J_c$ , at the highest operating field. Typically,  $J_{op}$  is in the range of 200 A/mm<sup>2</sup> to 700 A/mm<sup>2</sup>, and  $J_{op}^2/J_c = 0.3$  A/mm<sup>2</sup> to 0.6 A/mm<sup>2</sup>. The current density over the coil cross-section is over one order of magnitude smaller.

In the non-steady-state tokamak machines, the normal operating cycles and the occasional plasma disruption set additional, challenging requirements, in terms of mechanical fatigue of the structural materials and pulsed field loads on the superconductors.

## REVIEW OF SUPERCONDUCTING COILS FOR FUSION

### Mirror Devices

In the last decade interest in the magnetic confinement in linear machine (mirror fusion) has strongly declined, but they were very popular in the 1970s. At the Kurchatov Institute in Moscow, a plasma trap named LIN-5 was built in 1970 by a split solenoid system with 0.2 m inner diameter and 5.8 T peak field at 1 kA. In 1975, for the LIN-5B machine, a 5 tonne, bath-cooled baseball type coil was wound with 6 km of square, monolithic NbTi conductor, 6.2 × 6.2 mm<sup>2</sup> (2), at an operating current of 2 kA (75% of the short sample current), the peak field at the conductor is 5.6 T. The basic coil and conductor design is very similar to the US Baseball II-T winding (see below).

In the United States, a superconducting magnetic mirror apparatus (SUMMA) consisting of four coils with 0.9 m outer diameter, 8.6 T peak field, and 16 MJ stored energy was built by NASA in the early 1970s. At Livermore, the first superconducting baseball coil (Baseball II-T) was wound in 1971 with a square monolithic NbTi conductor, 6.35 × 6.35 mm<sup>2</sup>; the peak field in the pool boiling cooled winding is 7.5 T at 2.4 kA, with a stored energy of 12 MJ.

The Mirror Fusion Test Facility (MFTF) (see Fig. 1) was assembled at Lawrence Livermore National Laboratory in 1985 (3). It is the largest set of superconducting magnets for fusion,

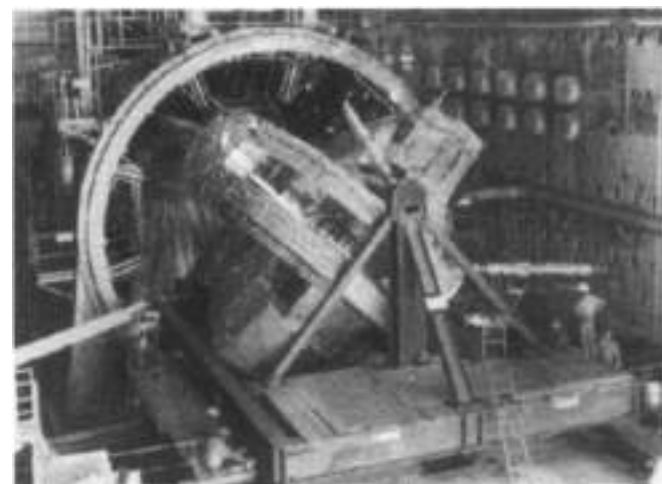


Figure 1. The Yin-Yang coils being assembled at one end the mirror fusion test facility (courtesy of C. H. Henning, Lawrence Livermore National Laboratory).

with a total mass is 1200 t and about 75 t of strand. It consists of 8 C-type coils, 12 low-field, and 4 high-field solenoids, including the A2 coils with a  $\text{Nb}_3\text{Sn}$  insert (see Fig. 2). All magnets are pool cooled at 4.5 K by natural convection, with two-phase coolant outlet 1:5% gas). The coils are all wound in the coil case, which acts as a cryostat, with ground insulation applied before winding. In the Yin-Yang coils, the winding form is fitted into the thick case by a copper bladder filled with urethane. The turn insulation is provided by G10 spacers, the layer insulation by perforated G-10 plates. The conductors for the solenoid (two types of NbTi and one  $\text{Nb}_3\text{Sn}$ ) consist of a thick multifilament composite soldered to the copper stabilizer (see Fig. 3). The joints are made by cold-welding of the composite. For the Yin-Yang coils, the conductor is a square NbTi monolith with a perforated Cu strip wrapped around and soldered to increase the wetted surface. The conductor for the pancake-wound  $\text{Nb}_3\text{Sn}$  insert, A21, is a flat multifilamentary composite soldered in the Cu housing after heat treatment (react and wind): the outer copper surface is oxidized to prevent solder wetting. All the conductors are designed to be cryostable. The fraction of  $I_{op}/I_c$  is always smaller than 2/3. After cool-down and successful commissioning of the magnet system, the project was discontinued in 1985. In 1990, the A2 coils were extracted and reassembled in the FENIX conductor test facility, which operated for three years at field levels as high as 13 T.

### Tokamaks

The toroidal arrangement of plasma is the most promising confinement geometry, with the largest number of experimental

devices, including the pinch and reverse pinch, the stellarator group of machines, and the tokamaks, where a toroidal plasma current is initiated and sustained by a pulsed ohmic heating coil (central solenoid). Most of the superconducting magnets for fusion belong to the tokamak family, including four plasma experiments (T-7, T-15, TRIAM, Tore Supra) and a number of sizable technology demonstration devices (LCT, TESPE, DPC, Polo, ITER Model Coils).

The very first superconducting tokamak, named T-7, was built at the Kurchatov Institute, Moscow in 1975–1976 and first cooled down in 1977 (see Fig. 4). The toroidal winding system (4) consists of 48 circular double pancakes, in aluminum case, with 60 turns in each coil. The average coil diameter is 1 m, the overall coil mass is 12 t and the stored energy 20 MJ at the nominal operation point. The torus was preassembled into eight segments, individually tested before final assembly. The forced flow conductor (see Fig. 5), is made from a strip of nine copper pipes, 2 mm inner diameter, with 16 multifilamentary and 32 single-core NbTi strands sitting in the grooves between the pipes and bonded to them by electroplating a 0.6 mm copper layer up to the final size of  $28 \times 4.5$  mm. The cooling is by two-phase helium at 4.5 K, with all the pancakes connected in parallel. A 15-mm-thick copper shell surrounds each coil and acts as an eddy currents shield for the poloidal field variations. The strands are not transposed and the conductor suffered from severe flux jumps, triggering quenches, during ramp up and ramp down. However, 80% of the design current (6 kA at 5 T peak field) was achieved. Operation of T-7 was discontinued at the Kurchatov Institute in 1987. Later, T-7 was transferred to the Chinese Institute of Plasma Physics, where it has been operating since 1996.

The toroidal field coils of the T-15 tokamak, first operated in 1988 at the Kurchatov Institute in Moscow, are the largest worldwide application of  $\text{Nb}_3\text{Sn}$  conductors (see Fig. 6) (5). The 24 circular coils, with average diameter 2.4 m, consist each of 12 single pancakes, coiled in parallel, with the He inlet at the inner radius joints. The turn insulation is obtained by wet winding to balance the uneven conductor contour. Two stacks of six pancakes are vacuum impregnated in two-halves steel cases, eventually bolted together. The react and wind method was applied. The forced-flow conductor (see Fig. 5), is a flat cable of 11 nonstabilized  $\text{Nb}_3\text{Sn}$  strands, bonded after heat treatment to two copper pipes by an electroplated Cu layer, 1.2 mm thick. Over 100 km of conductor have been manufactured, in units of 200 m. Each coil was tested before assembly and achieved the specification, although a steady-state voltage was observed, by far larger than the joint voltage, in the range of 2.5 mV to 10 mV/coil. The design current in the tokamak is 5.6 kA at 9.3 T peak field. The limited size of the cryopant (allowing a mass flow rate of only 0.38 g/s conductor) and the large radiation loss limited the operating temperature to the range of 9 K to 16 K. The highest operation point was 3.9 kA at 6.5 T, 7 K to

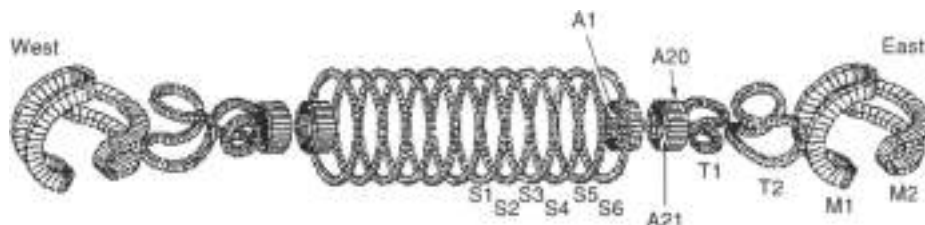


Figure 2. The Axial coil configuration for the MFTR, with the Yin-Yang coils at the ends.



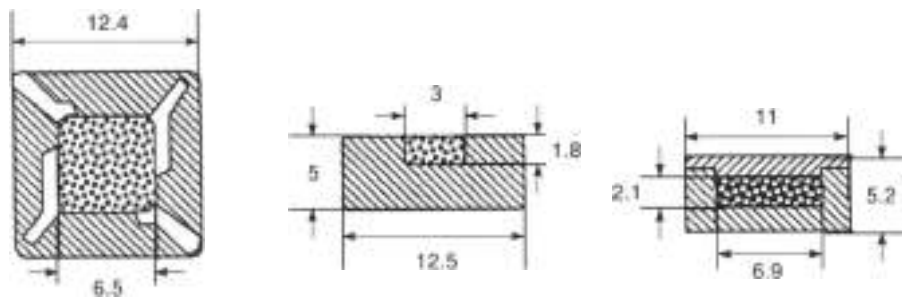


Figure 3. Conductors for the MFTF magnets: from left to right: NbTi square conductor with enhanced wet surface for the M, T A1, and A20 coils, NbTi conductor for the S solenoids, react and wind Nb<sub>3</sub>Sn conductor for the A21 insert. (See Fig. 2 for coil identification.)

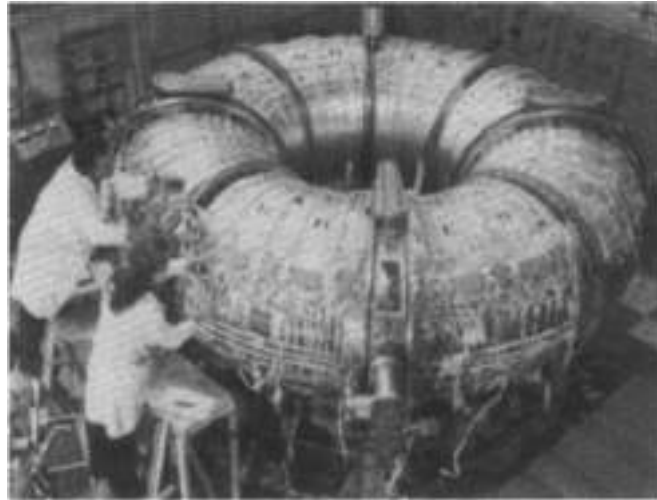


Figure 4. The 24 double pancakes of TF in the final assembly (courtesy of V. Keilin, Kurchatov Institute).



Figure 6. The 24 Nb<sub>3</sub>Sn coils of T-15 assembled with the 12 horizontal parts (courtesy of V. Keilin, Kurchatov Institute).

8 K, in agreement with the single-coil test and in excess of the original strand specification. The field transients due to plasma disruption, up to 40 T/s, were withstood without quench, with an increase of the outlet temperature by 0.25 K. The design ground voltage is 1.5 kV. For fast discharge, 250 V was applied at the terminals, with a time constant of 104 s.

The TRIAM device at the University of Kyushu, Fukuoka (Japan), is a compact, high-field tokamak, first operated in 1986. The superconducting 16 D-shaped toroidal field coils, with 7.3 m average perimeter, are cooled by a pressurized liquid helium bath at 4.5 K. The poloidal field coils, wound with

normal conductor, are placed inside the TF coils. The magnet cold mass is 30 t. The toroidal field conductor (see Fig. 7), consists of a large Nb<sub>3</sub>Sn bronze composite (10.5 × 3.3 mm for high grade, over a half million filaments) soldered after heat treatment in a copper housing with roughened side surfaces to improve the heat exchange. Beside the copper housing (RRR = 90), a Cu clad high-purity Al profile (RRR = 3000) is used as a stabilizer. Each coil is a stack of six double pancakes, laminated in a steel case. Three conductor grades, with the same width and decreasing height, are used with soldered joints for the double pancake. The conductor is designed to be cryostable and can withstand an energy input up to 7.8 J/cm<sup>2</sup> at the operating conditions; in case of plasma disruption, a normal zone

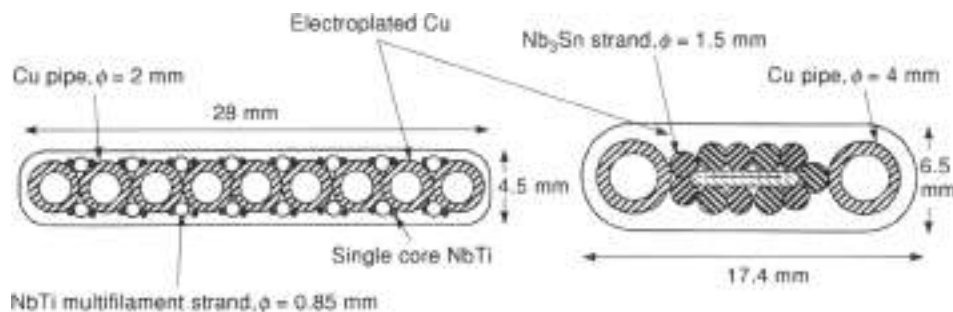


Figure 5. The forced flow conductors for T-7 (left) and T-15 tokamaks, bonded by copper electroplating.

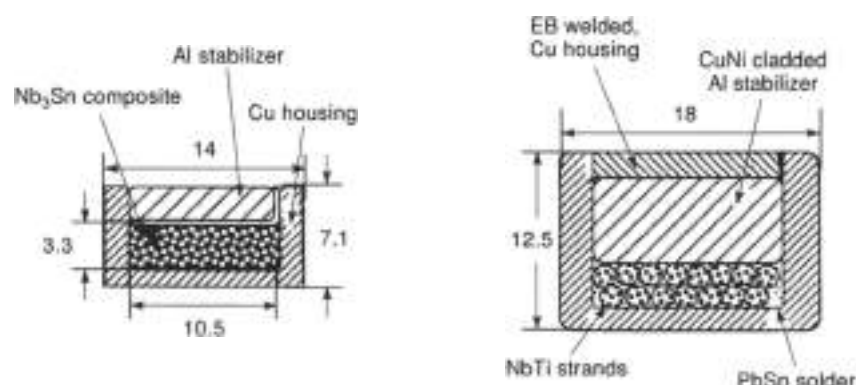


Figure 7. Soldered monolithic conductors, stabilized with high-purity aluminum profiles, for TILAM (left) and the helical coils of LHD (right).

may locally occur but is recovered within 0.8 s. At 6.2 kA, 11 T, the ratio  $I_{\text{op}}/I_c$  is 0.6. No quench event has been reported after three years of operation, including plasma disruptions (6).

The tokamak Tore Supra has been assembled at Cadarache, France, in 1987 (7). The poloidal coils are wound from copper conductors. The 18 pool-cooled circular TF coils, wound from NbTi superconductor operating at superfluid helium, are the largest magnet mass, 160 ton, cooled at 1.8 K. Each coil is made out of 26 double pancake with an average diameter of 2.5 m (see Fig. 8). A cowound prepreg tape, 0.15 mm thick, is used for the turn insulation. The pancake spacers, 2.2 mm, are built by a perforated prepreg thin plate with glued glass-epoxy bottoms. The ground insulation is obtained by overlapped prepreg plates. A 2-mm-thick steel case is shrink-fitted to the winding and contains the atmospheric, 1.8 K He bath. A thick steel case,

with thermal insulation, is shrink-fitted and cooled at 4.2 K. The conductor is a rectangular NbTi/Cu/CuNi multifilament composite, 2.8 × 5.6 mm, wound on the short edge to minimize the ac loss from the poloidal field variation. The temperature margin is ~2.5 K, with  $T_c = 4.25$  K at 1400 A, 9 T peak field. Little copper cross section is used in the conductor for stability; the He bath enthalpy up to the  $\lambda$  point is available, due to the very high thermal conductivity of He II. The heat exchanger is placed underneath the coil case, open only at the bottom. In case of quench, a He gas pressure builds on the top of the case/cryostat, and siphons the whole He volume within 3 s through the bottom opening, providing a very fast quench propagation and limiting the hot spot temperature below 80 K. In 1988, about six months after first operation, an interpancake short occurred at one coil during a fast discharge, with 1.5 kV across the coil and ~60 V across pancakes, well below the expected Paschen minimum for helium. The damaged coil was later replaced with a spare coil and the dump voltage was decreased to 500 V in order to limit the pancake voltage to ~20 V, which was experimentally assessed as the safe threshold to avoid interpancake discharge. The poloidal field variations and the plasma disruption result in a temperature increase in the He II bath as small as 0.01 K.

Besides the four above described tokamaks, a number of prototype superconducting coils have been built under national or international auspices. The demonstration coils provide valuable opportunities to learn about magnet and conductor technology. In such projects, the pressure for a conservative design is less strong and the performance margins can be better explored than in a plasma experimental device. In the IAEA Large Coil Task at the Oak Ridge National Laboratory (8), six large D-shaped magnets, 3 × 3.5 m bore, have been built to the same common specification using substantially different design approaches (see Table 2). The coils, assembled as a tokamak (see Fig. 9), operated (1984–1985) at the same design point (8 T peak field) with margins ranging from 120% to 140%. For the first time, a cable-in-conduit Nb<sub>3</sub>Sn conductor (react and wind coil manufacture) was used in a large-scale application and, despite the broad resistive transition observed in selected coil sections (similar to T-15 behavior), the coil reached 8 T with  $T_c = 8$  K. The cryostable conductors for the bath-cooled coils (GE, GE, JA) could be easily graded (both layer and pancake windings) and, using soldered copper profiles as stabilizers, achieved impressive results in terms of effective use of strand: from 1.4 t of strand in the GE conductor to 8.2 t in the forced-flow conductors of EU (NbTi) and WH (Nb<sub>3</sub>Sn), cooled by supercritical

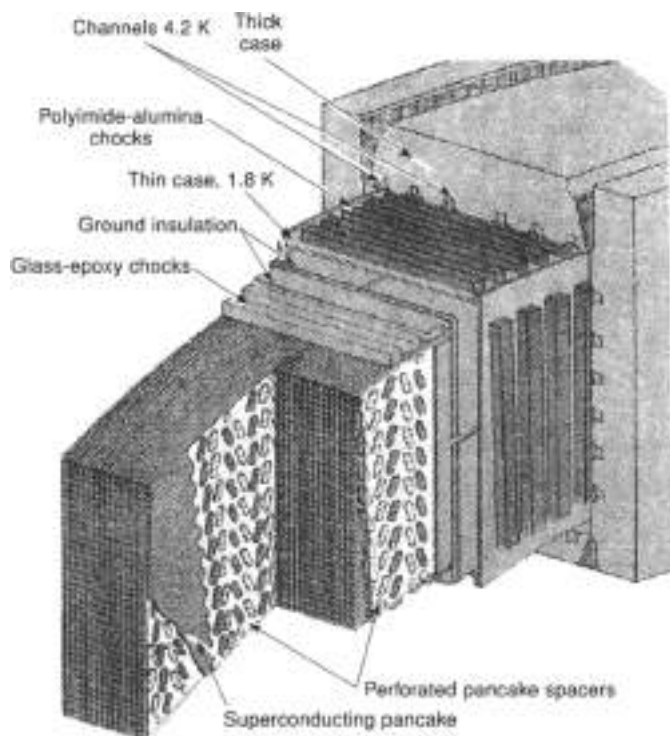


Figure 8. Winding pack layout for the pool-cooled, NbTi (inside) field coils of Tore Supra courtesy of B. Touck, Tore Supra.

Table 2. Summary of the LCT Coil Characteristics

|                | GD                              | GE                              | WH                                      | CH                        | EL                                    | JA                               |
|----------------|---------------------------------|---------------------------------|---|---------------------------|---------------------------------------|----------------------------------|
| Winding type   | 14 layers<br>3 grades           | 7 double pan-<br>cake, 3 grades | 4-in-band, 12<br>double pan-<br>cakes   | 23 pancakes               | 2-in-band, 7 dou-<br>ble pancakes     | 20 double pan-<br>cake, 2 grades |
| Cooling method | Pool boiling                    | Pool boiling                    | Forced flow                             | Forced flow               | Forced flow                           | Pool boiling                     |
| Conductor      | Soldered flat<br>cable, no edge | Divided flat cable              | Nb <sub>3</sub> Sn cable-in-<br>conduit | Square, soldered<br>cable | Divided, flat<br>cable-in-<br>conduit | Soldered flat<br>cable, no edge  |
| Non-Cu $J_c$   | 596 A/mm <sup>2</sup>           | 525 A/mm <sup>2</sup>           | 586 A/mm <sup>2</sup>                   | 302 A/mm <sup>2</sup>     | 393 A/mm <sup>2</sup>                 | 327 A/mm <sup>2</sup>            |
| Winding $J_c$  | 27.4 A/mm <sup>2</sup>          | 24.7 A/mm <sup>2</sup>          | 20.1 A/mm <sup>2</sup>                  | 30.3 A/mm <sup>2</sup>    | 25.7 A/mm <sup>2</sup>                | 26.6 A/mm <sup>2</sup>           |
| He inventory   | 1320 l                          | 1755 l                          | 440 l                                   | 110 l                     | 687 l                                 | 1425 l                           |
| Total weight   | 42.9 t                          | 36.6 t                          | 33.7 t                                  | 41.7 t                    | 39 t                                  | 38 t                             |
| SC strand      | 2 t                             | 1.4 t                           | 3.2 t                                   | 3.5 t                     | 3.2 t                                 | 2.6 t                            |
| Test voltage   | 9 kV                            | 2.5 kV                          | 9.2 kV                                  | 10 kV                     | 12 kV                                 | 3 kV                             |

helium at 3.6 K, 10 bar to 16 bar. However, two out of three pool cooled coils could not be dumped to the design voltage of 1 kV. The nuclear heat load was simulated by heaters on the inner radius. The poloidal field coil variations were reproduced by a pulsed coil traveling inside the torus: the pulsed field test, with  $\Delta B = 0.1$  T,  $\Delta \dot{B} = 0.14$  T/s,  $t_r = 1$  s, could be completed only for three coils (JA, CH, EL). Over 90% of the stored energy could be dumped in the external resistors, except for the WH coil, with short circuited radial plates.

At the same time of the LCT project, two small-size experiments with D-shaped coils were carried out at the Forschungszentrum Karlsruhe (Germany) and at Toshiba (Japan). The six toroidal coils of TESPE at Karlsruhe (9) have a 0.5 × 0.6 m bore and are pool cooled at 4.2 K (8 t total cold mass). The coils are wound as double pancakes, shrink-fitted in steel housings insulated by glass epoxy laminate. The steel case is electron-beam welded and serves both as liquid helium container and mechanical reinforcement. The conductor is a soldered flat cable of 24 multifilament NbTi strands  $\phi = 1.45$  mm, operating at 7 kA with a peak field of 7 T. The TESPE torus was first operated in 1984 with a test program focused

on mechanical load and high-voltage safety issues. The double pancake built by Toshiba (10) in 1983 had a Nb<sub>3</sub>Sn cable-in-conduit conductor, 18.3 × 15.7 mm, with 426 strands, encased into a 1-mm-thick 316L steel jacket. The D-shaped coil, 1.1 × 0.9 m, was wound after heat treatment and tested at 10 kA, in a peak background field of 10 T, provided by a small split coil. The test included hydraulic friction factor and stability.

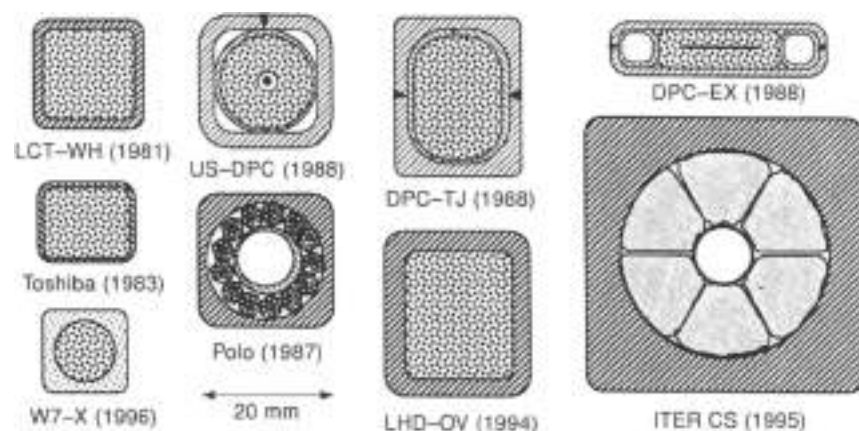
The most recent development project for tokamak's toroidal field coil is the ITER TF Model Coil, to be tested at the Forschungszentrum Karlsruhe by the end of 1999 in the background field of the EU-LAT coil (11). The race-track-shaped coil, with 2.5 × 1.4 bore, has an overall weight of 31 t and a stored energy of 60 MJ at 70 kA (compared to ≈400 t and ≈4 GJ in an individual full size ITER coil). The peak field is about 9 T, compared to 12.5 T in ITER. The main goal of the ITER TF Model Coil is to demonstrate the winding technique, which uses precision-machined steel radial plates where the pancake wound Nb<sub>3</sub>Sn cable-in-conduit conductor is encased after the heat treatment (react and transfer method).

The Demonstration Poloidal Coil (DPC) project at JARRI, Naka (Japan), aimed at comparing conductor design options for pulsed-field coils. Three Nb<sub>3</sub>Sn react and wind winding models with  $\phi_w = 1.5$  m, DPC-EX (12), US-DPC (13) and DPC-TJ (14) were tested in the DPC facility in pulsed mode up to about 7 T to 8 T, sandwiched between two pulsed NbTi solenoids (DPC-U), connected in series. The conductors are all forced-flow cable-in-conduit (see Fig. 10). The jacket material is Incoloy for the US-DPC (used for the first time). The DPC-TJ had a double jacket, the outer one is 3-D machined without bending and fitted by spot-welding to the conductor after the heat treatment. The strand surface is bare copper for the DPC-TJ, and the coupling loss is 1000 times larger compared to the US-DPC and DPC-EX conductors, with Cr plated strands (2 mol). Ramp rate limitation was observed in the US-DPC, probably due to transposition errors in the cable. The NbTi strands in the cable-in-conduit of the background coils were insulated to avoid interstrand coupling loss: the conductor turned to be unstable due to the inability to redistribute effectively the current among the strands.

The Polu coil, a NbTi circular winding with  $\phi = 3$  m, has been tested in 1994 at the Forschungszentrum Karlsruhe (Germany) (15). The cable-in-conduit conductor (see Fig. 10), has two separate hydraulic circuits: stagnant, supercritical He at 4 bar in the annular cable region and forced flow, 2 g/s, two-phase He at 4.5 K in the central pipe. This design allows a homogeneous temperature along the conductor with a small pressure drop,



Figure 9. The six D-shaped LCT coils assembled as a torus in the vacuum tank (courtesy of M. Lubell, Oak Ridge National Laboratory).



**Figure 10.** Selection of cable-in-conduit superconductors, drawn to the same scale. The strands of Pk-a, LHD-OV and W7-X are NbTi, all the other are Nb<sub>3</sub>Sn strands. The jacket material is steel except for ITER (Incoloy) and W7-X (aluminum alloy).

The strand is a NbTi/Cu/CuNi composite and the subcables have CuNi or insulating barriers, resulting in very low coupling loss,  $\tau = 210$  ns. Four stainless-steel corner profiles are laser-welded around the cable. Phase-resolved partial discharge was first used at 4 K to assess the integrity of the glass-epoxy insulation. A midpoint electrical connection in the winding enables to create very-high-field transients in a half coil by a fast discharge of the other half coil. The coil has been tested up to 15 kA, 3.6 T. A degradation of  $I_c$  by 30% in dc operation has been observed compared to the strand performance. Polo does not have ramp rate limitation, the stability criterion being the only limiting criterion. Very-fast field transient, up to 1000 T/s are withstood without quench. High-voltage operation, up to 23 kV, has been demonstrated.

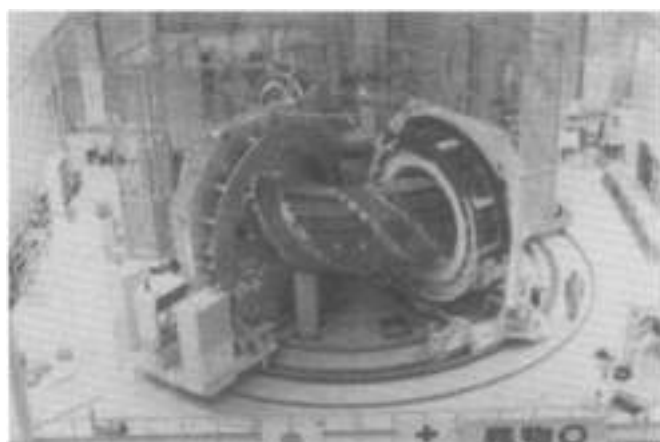
The most recent pulsed-coil development for tokamak is the ITER CS Model Coil (16), a layer-wound, two-in-hand solenoid to be operated at 45 kA, 13 T with 0.4 T/s field rate, scheduled to be tested at Naka (Japan) in 1999. The stored energy is 64 MJ, compared with 13 GJ in the full size central solenoid. The conductor is a Nb<sub>3</sub>Sn cable-in-conduit with a thick-walled Incoloy 905 jacket, manufactured as extruded and drawn-down bars, assembled by butt-welding into a oversize pipe, up to 250 m long: the strand bundle is pulled into the jacket and eventually rolled to the final size. The conductor is insulated by prepreg glass fabric with interleaved kapton foil, applied after the heat treatment by controlled unspinning of the individual layers.

### Stellarators

The plasma confinement can be achieved in stellarators by a number of winding configurations. The two large projects using superconducting coils are the Large Helical Device (LHD) at Toku, Japan, and the Wendelstein VII-X (W7-X), to be assembled at Greifswald, Germany. The superconducting magnet system of LHD operates in dc mode and consists of a two helical coils (17) wound around the toroidal vacuum vessel and three sets of circular poloidal coils (16). The helical coils are pool-cooled, initial operation will be at 4.2 K, 6.9 T, 13 kA and, at a later stage, at 1.8 K, 9.2 T, 17.3 kA. A precision tool with 13 numerically controlled driving axes has been used to wind in situ the two helical coils (see Fig. 11). The turn and layer insulation is provided by glass epoxy spacers graded across the winding pack to provide the best mechanical support in the stressed area and the largest wet conductor surface at the peak field (the highest field and the highest mechanical stress are not

at the same winding location). The conductor (see Fig. 7), is a NbTi flat cable soldered to a CuNi-cladded Al stabilizer into a copper housing, eventually sealed by two electron beam welds. The conductor is designed to be cryostable, with  $I_{op}/I_c = 0.55$ . The forced-flow conductor for the poloidal coils (see Fig. 10), is a NbTi cable-in-conduit with 486 strands ( $\phi = 0.76$  mm or 0.89 mm), 38% void fraction. To improve the current sharing among strands, the surface is not coated, with a coupling loss constant of 300 ms. The temperature margin is 1.2 K to 1.6 K,  $I_{op}/I_c = 0.33$ . The joints between the conductor sections are realized by filament joining, resulting in 0.14 mΩ resistance at full current. The OV coil,  $\phi = 11.5$  m (see Fig. 12), has been wound on the site, with prepeg turn insulation. The first operation of the LHD was successfully started in March of 1998.

The magnet system of the W7-X torus consists of 50 nonplanar and 20 planar coils assembled in five modular segments (19). A first nonplanar model coil was completed in 1997 and tested in 1995. The completion of the machining is scheduled by the year 2002. The forced-flow conductor to be used for all the coils (see Fig. 10), is a NbTi cable-in-conduit with 243 strands,  $\phi = 0.57$  mm, 37% void fraction. The square jacket is made of a hardenable Al alloy, extruded around the cable: it is soft after extrusion and during the winding process. After hardening at 170 C, it provides the required stiffness to the winding park. The temperature margin is 1 K and  $I_{op}/I_c = 0.5$ .



**Figure 11.** Winding tool with 13 numerically controlled axes for the helical coils of the LHD (courtesy of K. Takahata, NIFS).



**Figure 12.** The OV poloidal field coil of the JET courtesy of K. Takahashi (NIFS).

## CRUCIAL ISSUES FOR THE DESIGN OF FUSION MAGNETS

### Electrical insulation

The very high level of neutron and gamma radiation at the vacuum vessel of a fusion reactor must be screened to limit the nuclear heat load and the radiation damage at the winding components. The size of the shield (up to 1 m thick in the ITER project) may have a substantial impact on the size and cost of the superconducting magnets. After screening, the radiation damage on the metallic components of a superconducting coil (steel, copper, NiTi and Ni<sub>3</sub>Sn), is not critical and partly recovers (e.g., for copper) upon warming up at room temperature. The actual weak link for radiation damage is the organic fraction of the electrical insulation. In pressed windings, the glass-epoxy is broadly used, either as laminates or prepreg wraps or vacuum-impregnated fabrics, to bond together the winding turns and to provide the required dielectric strength. The neutron and gamma act on the long molecular chain of the resin, irreversibly breaking the atomic links. The mechanical strength of the composite, mostly the shear strength, is affected and macroscopic cracking may occur under operating loads, eventually leading to a short circuit between winding sections. To limit this risk, the magnets must be designed to have low stress in the insulation, that is, limit the risk of crack propagation. Another design approach is to separate the mechanical and electrical functions, for example, including a redundant electrical insulation layer, either interleaved or overlapped to the glass-epoxy, to stop the crack propagation in the resin. The free radicals originated from the broken organic polymers are chemically active and evolve into gaseous molecules. The most severe consequence of radiation-induced chemical reactions at 4 K is the accumulation of frozen gas bubbles (mostly hydrogen). Upon warming-up of large windings, the internal pressure of the evolved gas increases dramatically due to the little permeation and may lead eventually to swelling in the insulation (20). A possible cure against postirradiation hazards of organic insulation is to reduce the resin volume fraction and select the resin composition to maximize the gas evolution rate. On the other hand, there is a broad reluctance to start an expensive and time-consuming task for the industrial development of innovative insulation systems, which will be actually needed only when a fusion reactor will work at full power on a time scale of several years. The full replacement of organic insulation systems by ceramic

materials with adequate mechanical properties may be the ultimate, long-term goal to solve the issue of the electrical insulation on the heavily irradiated fusion magnets.

### Quench Protection

In case of quench, the huge amount of energy stored in a fusion magnet must be actively dumped in an outer resistor. If a quench fails to be detected, the ohmic power locally dissipated in the slowly expanding normal zone is sufficient within one minute or less to melt the conductor and start a chain of serious failures (vacuum break, electric arc, mechanical collapse). A number of quench detectors have been developed and are currently applied in superconducting magnets, from the easy ones (voltage balancer of different winding sections, monitoring of outler mass flow rate) to the most sophisticated, including the laser interference on optical fibers used as distributed thermometer, transmission, and reflection of super high-frequency waves in the coolant channel, acoustic emission, magnetization change at the normal zone (21). However, a redundant and intrusive instrumentation is not welcome in a fusion reactor, as it may increase the risk of leaks and insulation failure, due to the large number of feedthrough required. Whatever the quench detector is, the ultimate question always arises: What happens if the active quench protection fails? The design approach for an actual fusion magnet (i.e., not for an experimental device) will need to offer both a reliable and robust quench-protection system and a conductor/magnet layout that intrinsically limits the damage in case of failure of the protection system, for example, enhancing the quench propagation and the enthalpy at intermediate temperature.

### Cost Optimization

In several applications of the superconducting magnets (e.g., accelerators, detectors, high-field magnets, prototypes), the achievement of the technical goal is the main care of the designer, while the cost of the device does not play a major role. However, after completion of the demonstration phase for the fusion magnets, the cost optimization will be a key issue for the commercial success of fusion. On one side, the behavior of the superconductor needs to be mastered by the designer (e.g., ac loss, stability, mechanical properties), in order to set the design margins at a safe but realistic level and make effective

use of the expensive superconductors. On the other hand, the choice of the manufacturing methods and tooling may have a very strong impact on the cost of the coil and should be included as a driving factor in the design. Two examples are given to show how a design choice may affect the cost.

A high electrical conductivity material (stabilizer) needs to be added to the superconductor cross-section, to allow effective current-sharing and fast recovery for small thermal disturbances. The required stabilizer cross-section may be much larger than the superconductor. In cable-in-conduit conductors, the straight choice is to equally distribute the stabilizer cross-section in each superconducting strand, specifying a high Cu:non-Cu ratio. However, the cost of the Nb<sub>3</sub>Sn strand is independent of the copper ratio. If the designer masters the mechanism of the current-sharing among strands and knows the operating values of the interstrand resistance, he or she may select a much smaller Cu:non-Cu ratio in the Nb<sub>3</sub>Sn strand and add extra copper wires in the strand bundle. Keeping the same superconductor cross-section, that is, without affecting the operating margins, the amount of Nb<sub>3</sub>Sn strand can be significantly reduced with a large cost saving.

A Nb<sub>3</sub>Sn conductor needs a heat treatment at 650 °C to 700 °C to form the brittle intermetallic composite by solid-state diffusion. If the designer does not master heat resistant electrical insulation systems, he or she will conservatively choose to first heat-treat the conductor and then insulate it and wind in the final shape (e.g., react and wind or wind and react and transfer methods). As the Nb<sub>3</sub>Sn after heat treatment is degraded for permanent deformation as large as 0.2% to 0.3%, the handling for post-heat treatment insulation and final assembly requires sophisticated tooling and continuous adjustment (e.g., shimming of each turn) to achieve the required tolerance with minimum strain on the conductor. If a reliable insulation system is selected, compatible with the heat treatment procedure, the coil can be wound in the final form and to the final tolerance before the heat treatment (wind and react method), saving the cost of a large number of tools and manufacturing steps and avoiding the risk associated to the post-heat-treatment handling.

#### Risk and Quality Assurance

Large superconducting magnets are usually unique items for which a thorough quality assurance program cannot be conveniently established in advance, as is the case for series production, due to the lack of iterative improvements in the manufacturing procedures. The global acceptance tests of the magnets tend to replace the quality assessment of the individual procedures, achievable only on the basis of a broad statistical database. However, an individual acceptance test of a fusion magnet cannot reproduce all the actual operating conditions, for example, mechanical load and peak field from other coils, nuclear radiation, mechanical and thermal cycling. Moreover, a global acceptance test should not be pushed to the failure limit, that is, the operating margin cannot be assessed. The lack of confidence may push the designer in a circle of overconservative choices, for example, assuming minimum performance for material properties, welds, assembly tolerances, which adversely affect the cost and the effectiveness of the design. To avoid this trend, the designer should identify the critical area for quality assurance and select a low risk design and procedure. For example, the resistance of a joint between conductor sections cannot be checked during the manufacture. In this case, the

designer should aim for a joint layout where the resistance performance is only marginally affected by incorrect assembly procedure.

#### BIBLIOGRAPHY

1. E. Teller (ed.), *Fusion*, New York: Academic Press, 1961.
2. N. A. Chernogolov, Superconducting magnet systems for plasma physics research in the USSR, *Proc. Magnet Technology Conf. MT-6*, 3 Bratislava, Czechoslovakia, 1977.
3. T. A. Kozman et al., Magnets for the mirror fusion test facility: Testing of the first Yin-Yang and the design and development of the other magnets, *IEEE Trans Magn.*, **19**: 859, 1993.
4. B. P. Ivanov et al., Test results of "tokamak-7" superconducting magnet system (SMS) sections, *IEEE Trans Magn.*, **16**: 550, 1979.
5. E. N. Bondarchuk et al., Tokamak-15 electromagnetic system: Design and test results, *Plasma Device Oper.*, **2** (1), 1992.
6. Y. Nakamura et al., Reliable and stable operation of the high field superconducting tokamak TRAM-1M, *Proc. Magnet Technol. Conf. MT-11*, 767 Tuskuba, New York: Elsevier, 1990.
7. B. Turck and A. Tornatou, Operating experience of three supra superconducting magnets, *Proc. 15th IEEE/NPSS SOFE*, Hyannis, MA, 1994, p. 303.
8. The IEA Large Coil Team, *Fusion Eng. Des.*, **7** (1-2): 1-229, 1989.
9. K. P. Jaeger et al., Superconducting torus "TESPE" at design values, *Proc. Magnet Technol. Conf. MT-9*, 36 Zurich, 1985.
10. M. Yamaguchi et al., Development of a 12 T forced-cooling toroidal field coil, *Proc. ICFE 10*, Helsinki, 1984, p. 169.
11. E. Salgredo et al., Construction of a toroidal field model coil (TFMC) for ITER, *Proc. Magnet Technol. Conf. MT-15*, Beijing, 1987.
12. Y. Takahashi et al., Experimental results of the Nb<sub>3</sub>Sn demountable poloidal coil (DPC-EX), *Cryogenics*, **31**: 640, 1991.
13. M. M. Stevens et al., Test results from the Nb<sub>3</sub>Sn US-demonstration poloidal coil, *Adv. Cryog. Eng.*, **37A**: 545, 1992.
14. M. Qiu et al., Charging test results of the DPC-TJ, a high current-density large superconducting coil for fusion machines, *IEEE Trans Appl. Supercond.*, **3**: 490, 1993.
15. M. Darweschad et al., Development and test of the poloidal field prototype coil POLO at the Forschungszentrum Karlsruhe, *Fusion Eng. Des.*, **36**: 227, 1997.
16. S. Mitchell et al., ITER CS model coil project, *Proc. 16th ICFE/ICMC*, Kitakyushu, New York: Elsevier, 1997, p. 763.
17. S. Imazawa et al., Helical coils for LH2, *Proc. Symp. Cryogenic Syst. for Large Scale Superconducting Applications*, Tokai, NIFS-PROC-28, 1996, p. 112.
18. K. Takahata et al., Lopoloidal coils for the Large Helical Device (LHD), *Proc. Symp. Cryogenic Syst. for Large Scale Superconducting Applications*, Tokai, NIFS-PROC-28, 1996, p. 118.
19. J. Sapper, The superconducting magnet system for the Wendelstein 7-X stellarator, *Proc. Ann. Meeting Amer. Nuclear Soc. (ANS)*, Reno, NV, 1996.
20. D. Evans, R. P. Reed, and N. J. Simons, Possible hazards following irradiation of superconducting magnet insulation, *Proc. 16th ICFE/ICMC*, Kitakyushu, New York: Elsevier, 1997, p. 2017.
21. A. Anghel et al., The quench experiment on long length. Final report to ITER, Villigen PSI, 1997.

## SUPERCONDUCTING MAGNETS FOR PARTICLE ACCELERATORS AND STORAGE RINGS

### TYPES OF PARTICLE ACCELERATOR

#### Accelerator Systems

One of the main activities in nuclear and high-energy physics is the study of internal structures of charged particles. The research is carried out by smashing particles into pieces and then analyzing the nature and characteristics of the pieces. The particles are broken by accelerating them to high momenta and either blasting them against a fixed target or colliding them among themselves. To increase the event rate, the particles are bunched into a high-intensity beam. The more elementary the particles, the higher the energy needed to smash them. Experiments at the proton scale require beam energies of the order of 1 TeV or more.

The beams of charged particles are produced by accelerator systems made up of several stages, which progressively raise the energy. In the largest machines, the last stage of the accelerator chain, usually referred to as main ring, can have a circumference of several tens of kilometers and is installed in an underground tunnel. Such a ring is operated in three phases: (1) *injection*, during which the beam, which has been prepared in various preaccelerators, is injected at low energy; (2) *acceleration*, during which the beam is accelerated to nominal energy; and (3) *storage*, during which the beam is circulated at nominal energy for as long as possible (typically up to 24 hr) and is made available for physics experiments. As mentioned above, there are two types of experiments: (1) *fixed-target* experiments, for which the beam is extracted from the main ring to be blasted against a fixed target, and (2) *colliding-beam* experiments, for which two counterrotating beams are blasted at each other. The breakage products are analyzed in large detector arrays surrounding the targets or collision points.

A main ring of a large accelerator system is designed as a synchrotron-type accelerator, and the beam is circulated on an ideally circular orbit, which remains the same throughout injection, acceleration, and storage (1). The charged particles are accelerated by means of electrical fields and are guided and focused by means of magnetic fields. The electrical fields are provided by RF cavities. In large machines, the bending and focusing functions are separated: the former is provided by dipole magnets, whereas the latter is provided by pairs of focusing/defocusing quadrupole magnets (see the discussion that follows). The magnets are arranged around the ring in a regular lattice of cells, which are made up of a focusing quadrupole, a set of bending dipoles, a defocusing quadrupole, and another set of bending dipoles. During acceleration, the field and field gradient of the magnets are raised in proportion to particle momentum to maintain the beam on the design orbit and to preserve its size and intensity.

#### Bending and Focusing Magnets

**Coordinate System Definitions.** Let  $(O, u, v, w)$  designate a rectangular coordinate system, and let  $(C)$  be a circle of center  $O$ , located in the  $(u, v)$  plane and representing the design orbit of an accelerator ring. Furthermore, let  $P$  be a given point of  $(C)$ , and let  $(P, x, y, z)$  designate a rectangular coordinate system associated with  $P$ , such that  $x$  is a unit vector parallel to  $(OP)$ ,

$y$  and  $w$  are one and the same, and  $z$  is tangent to  $(C)$  at  $P$ . The  $x$  axis defines the horizontal direction, the  $y$  axis defines the vertical direction and the  $z$  axis corresponds to the main direction of particle motion.

**Normal Dipole Magnet.** A normal dipole magnet is a magnet, which, when positioned at  $P$ , produces within its aperture a magnetic flux density parallel to the  $(x, y)$  plane and such that

$$B_x = 0 \quad \text{and} \quad B_y = B_1 \quad (1)$$

where  $B_x$  and  $B_y$  are the  $x$  and  $y$  components of the flux density and  $B_1$  is a constant.

According to Lorentz' law, a charged particle traveling along the direction of the  $z$  axis through the aperture of such a magnet is deflected on a circular trajectory parallel to the horizontal  $(x, z)$  plane. The trajectory radius of curvature  $r$  can be estimated from

$$r = \frac{E}{0.3 q B_1} \quad (2)$$

Here,  $r$  is in meters,  $B_1$  is in teslas,  $q$  is the particle charge in units of electron charge, and  $E$  is the particle energy in gigaelectron-volts (GeV). The effect of a dipole magnet on a beam of charged particles is similar in some respects to that of a prism on a light ray.

Equation (2) shows that, to maintain a constant radius of curvature as the particle is accelerated, the dipole field must be ramped up in proportion to particle energy.

**Normal Quadrupole Magnet.** A normal quadrupole magnet is a magnet, which, when positioned at  $P$ , produces within its aperture a magnetic flux density parallel to the  $(x, y)$  plane and such that

$$B_x = gy \quad \text{and} \quad B_y = gx \quad (3)$$

where  $g$  is a constant referred to as the *quadrupole field gradient* (in teslas per meter).

According to Lorentz' law, a beam of positively charged particles traveling along the direction of the  $z$  axis through the aperture of such a magnet is horizontally focused and vertically defocused when  $g$  is positive, and vertically focused and horizontally defocused when  $g$  is negative. In reference to its action along the  $x$  axis on a beam of positively charged particles traveling in the  $z$  direction, a magnet with a positive gradient is called a *focusing* quadrupole, while a magnet with a negative gradient is called a *defocusing* quadrupole. To obtain a net focusing effect along both  $x$  and  $y$  axes, focusing and defocusing quadrupoles must be alternated in the magnet lattice. For either type of quadrupole magnets, the focal length  $f$  can be estimated from

$$f = \frac{E}{0.3 q g l_q} \quad (4)$$

Here,  $f$  is in meters,  $E$  is in GeV,  $q$  is in units of electron charge,  $g$  is in teslas per meter, and  $l_q$  is the quadrupole magnetic length in meters. The effect of focusing/defocusing quadrupoles on a beam of charged particles is similar to that of convex/concave lenses on a light ray.

Equation (1) shows that to maintain  $f$  constant as the particle beam is accelerated, the quadrupole field gradient must be ramped up in proportion to beam energy.

## PARTICLE ACCELERATORS AND SUPERCONDUCTIVITY

### Why Superconductivity?

Throughout the years, the quest for elementary particles has promoted the development of accelerator systems producing beams of increasingly higher energies. Equation (2) shows that, for a synchrotron, the particle energy is directly related to the product  $\lambda R_1$ . Hence, to reach higher energies, we must increase either the accelerator radius or the dipole field (or both). Increasing the accelerator radius means a bigger tunnel. Increasing the dipole field above 2 T implies the use of superconducting magnets. The trade-off between tunneling costs, magnet development costs, and accelerator operating costs is, since the late 1970s, in favor of using superconducting magnets generating the highest possible field and field gradient (2).

Superconductivity is a unique property exhibited by some materials at low temperatures when the resistivity drops to zero. As a result, materials in the superconducting state can transport current without power dissipation by the Joule effect. This offers at least two advantages for large magnet systems such as those needed in accelerator main rings: (1) significant reduction in electrical power consumption and (2) the possibility of relying on much higher overall current densities in the magnets coils. There are, however, at least three drawbacks in using superconducting magnets: (1) the superconductor generates magnetization effects that result in field distortions that have to be corrected (see section on field quality), (2) the magnets must be cooled down and maintained at low temperatures, which requires large cryogenic systems (see section on magnet cooling), and (3) it may happen that an energized magnet, initially in the superconducting state, abruptly and irreversibly switches back to the normal resistive state in a phenomenon referred to as a *quench* (see section on quench performance).

The occurrence of a quench causes an instantaneous beam loss and requires that all or part of the magnet ring be rapidly ramped down to limit conductor heating and possible damage to the quenching magnet (see section on quench protection). Once the quenching magnet is discharged, it can be cooled down again and restored into the superconducting state, and the machine operations can resume. A quench is seldom fatal but is always a serious disturbance. All must be done to prevent it from happening, and all cautions must be taken to ensure the safety of the installation when it does happen.

### Review of Superconducting Particle Accelerators

**Tevatron.** The first large-scale application of superconductivity was the Tevatron, a proton synchrotron with a circumference of 6.3 km built at Fermi National Accelerator Laboratory (FNAL) near Chicago, IL, and commissioned in 1983 (3). The Tevatron now operates as a proton/antiproton collider with a maximum energy of 900 GeV per beam. It relies on about 1000 superconducting dipole and quadrupole magnets, with a maximum operating dipole field of 4 T (4).

**HERA.** The next large particle accelerator to rely massively on superconducting magnet technology was HERA (Hadron

Electron Ring Anlage) built at DESY (Deutsches Elektronen-Synchrotron) near Hamburg, Germany, and commissioned in 1990 (5). HERA is an electron/proton collider with a circumference of 6.3 km. It includes two large rings: (1) an electron ring, relying on conventional magnets (maximum energy: 30 GeV), and (2) a proton ring, relying on superconducting magnets (maximum energy: 820 GeV). The maximum operating field of the superconducting dipole magnets is 4.7 T (6).

**UNK.** Since the early 1980s, the Institute for High Energy Physics (IHEP) located in Plovdiv, near Moscow, Russia, has been working on a proton accelerator project named UNK (Uskoritelno-Nakopitelny Komplex). The circumference of the UNK main ring is 21 km for a maximum energy of 3 TeV in a fixed target mode (7). The maximum operating dipole field is 5 T (8). A number of superconducting dipole and quadrupole magnet prototypes have been built and cold tested, and the tunnel is almost completed, but, given the present (1998) economic situation in Russia, the future of the machine is undecided.

**SSC.** In the mid 1980s, the United States started the Superconducting Super Collider (SSC) project, a giant proton-proton collider with a maximum energy of 20 TeV per beam (9). The last stage of the SSC complex would have been made up of two identical rings of superconducting magnets installed on top of each other in a tunnel with a circumference of 87 km. The maximum operating dipole field was 6.8 T. The project was eventually canceled in October 1993 by decision of the US Congress, after 12 miles of tunnel had been dug near Dallas, TX, and a successful superconducting magnet R&D program had been carried out (10).

**RHIC.** Brookhaven National Laboratory (BNL), located on Long Island, NY, will complete the construction in 1999 on its site of the Relativistic Heavy Ion Collider (RHIC). RHIC is designed to collide beams of nuclei as heavy as gold, accelerated in two identical rings to energies between 7 and 100 GeV per beam and per unit of atomic mass (11). Each ring has a circumference of 3.6 km; the maximum operating dipole field is 8.4 T (12).

**LHC.** In December 1994, the European Laboratory for Particle Physics (CERN) approved the construction of the Large Hadron Collider (LHC) in its existing 27-km circumference tunnel located at the Swiss/French border, near Geneva, Switzerland (13). LHC will be a proton/proton collider with a maximum energy of 7 TeV per beam. It will have a single ring of so-called *flat aperture* superconducting magnets, housing within the same mechanical structure, the pipes for two counter-propagating proton beams (14). The maximum operating dipole field is set at 8.36 T. Commissioning is planned for 2005.

### Prominent Features of Superconducting Accelerator Magnets

Selected parameters of the major projects of superconducting particle accelerators are summarized in Table 1, whereas Fig. 1(a-e) presents cross-sectional views of the Tevatron, HERA, SSC, RHIC, and LHC dipole magnets (15).

The magnets rely on similar design principles, which are detailed in the following sections. The field is produced by saddle-shaped coils that, in their long straight sections, approximate cosine distributions of conductors, with  $n = 1$  for dipole magnets



**Table 1. Selected Parameters of Major Superconducting Particle Accelerators**

| Laboratory                       | FNAL        | DESY | HEP*      | SSCL      | BNL                    | CEBN              |
|----------------------------------|-------------|------|-----------|-----------|------------------------|-------------------|
| Name                             | Tevatron    | HERA | UNK       | SSC       | RHIC                   | LHC               |
| Circumference (km)               | 6.3         | 6.3  | 21        | 57        | 3.8                    | 27                |
| Particle type                    | p $\bar{p}$ | ep   | pp        | pp        | heavy ions             | pp                |
| Energy/beam (TeV)                | 0.9         | 0.62 | 3         | 20        | up to 0.1 <sup>†</sup> | 7                 |
| Number of dipoles                | 774         | 416  | 2168      | 7914      | 264                    | 1232 <sup>‡</sup> |
| Aperture (mm)                    | 76.2        | 75   | 70        | 50        | 80                     | 56                |
| Magnetic length (m)              | 6.1         | 9.6  | 5.9       | 35        | 9.7                    | 14.2              |
| Field (T)                        | 4           | 4.66 | 5.9       | 6.79      | 3.4                    | 6.36              |
| Number of quadrupoles            | 216         | 256  | 322       | 1096      | 276                    | 366 <sup>†</sup>  |
| Aperture (mm)                    | 58.9        | 75   | 70        | 50        | 40                     | 56                |
| Magnetic length <sup>§</sup> (m) | 1.7         | 1.9  | 3.0       | 5.7       | 1.1                    | 3.1               |
| Gradient (T/m)                   | 76          | 91.2 | 97        | 194       | 71                     | 221               |
| Commissioning                    | 1983        | 1990 | undecided | cancelled | 1999                   | 2005              |

\* Per unit of atomic mass.

† Two in-site magnets.

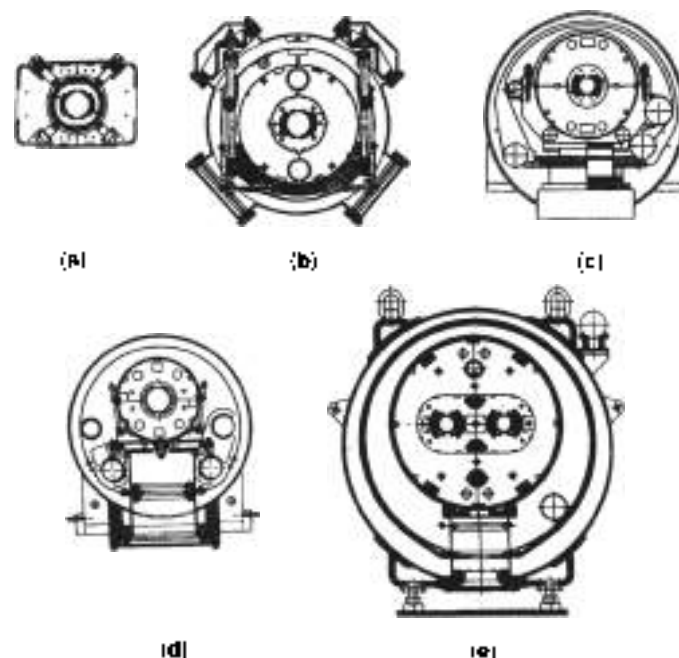
‡ Quadrupoles come in several lengths.

and  $n = 2$  for quadrupole magnets). The coils are wound from Rutherford-type cables made of NbTi multifilamentary strands and are mechanically restrained by means of laminated collars. The collared-coil assembly is placed within an iron yoke providing a return path for the magnetic flux. In the case of the Tevatron, the collared-coil assembly is cold, whereas the iron yoke is warm. Starting with HERA, the iron yoke is included in the magnet cryostat, and the cold mass is completed by an outer shell delimiting the region of helium circulation. The cold mass of the LHC magnets includes two collared-coil assemblies within a common yoke. Tevatron, HERA, UNK, SSC, and RHIC magnets are cooled by boiling helium at 1 atm (4.2 K) or supercritical helium at 8 atm to 5 atm (between 4.5 K and 6 K),

whereas LHC magnets are cooled by superfluid helium at 1.9 K (1 atm  $\approx$  0.1 MPa).

#### Superconducting Accelerator Magnet R&D

A number of laboratories are presently involved in R&D work on high field or high field gradient accelerator magnets. Among them is Twente University, located near Enschede in the Netherlands, which, in 1995, cold-tested at CERN a short model dipole magnet (made with Nb<sub>3</sub>Sn cable), which reached 11 T on its first quench at 4.4 K (16). Soon after, in early 1997, Lawrence Berkeley National Laboratory (LBNL), located in Berkeley, CA, cold-tested a short dipole magnet model (also made with Nb<sub>3</sub>Sn cable), which, after a number of training quenches, reached a record dipole field of 13.5 T at 1.8 K (17).



**Figure 1.** Cross-sectional views of superconducting dipole magnets for large particle accelerators: (a) Tevatron, (b) HERA, (c) SSC, (d) RHIC, and (e) LHC.

#### CONDUCTOR AND CONDUCTOR INSULATION

##### Superconducting Material

The most widely used superconducting material is a metallic alloy of niobium and titanium (NbTi), with a Ti content between 45% and 50% in weight (18). NbTi is easy to mass-produce and has good mechanical properties. It is a type-II superconductor, with a coherence length  $\xi$  of 5 nm, and a London penetration depth  $\lambda_L$  of 300 nm (chapter 2 of Ref. 2).

The upper critical magnetic flux density of NbTi,  $B_{c2}$ , can be estimated as a function of temperature  $T$  using

$$B_{c2}(T) = B_{c20} \left[ 1 - \left( \frac{T}{T_{c0}} \right)^2 \right] \quad (5)$$

where  $B_{c20}$  is the upper critical magnetic flux density at zero temperature (about 14.5 T) and  $T_{c0}$  is the critical temperature at zero field (about 9.2 K).

The critical current density of NbTi,  $J_{c1}$ , can be parametrized as a function of temperature, magnetic flux density,  $B$ , and

critical current density at 4.2 K and 5 T,  $J_{c0}$ , using (19):

$$\frac{J_{c0} B_c T_c}{J_{c0} T_c} = \frac{C_1}{B} \left[ \frac{B}{B_c \beta T_c} \right]^\nu \left[ 1 - \frac{B}{B_c \beta T_c} \right]^\beta \left[ 1 - \left( \frac{T}{T_c} \right)^\gamma \right]^\gamma \quad (16)$$

where  $C_1$ ,  $\nu$ ,  $\beta$ , and  $\gamma$  are fitting parameters. (Typical values for LHC strands are  $C_1 = 30$  T,  $\nu = 0.6$ ,  $\beta = 1.0$ , and  $\gamma = 2.0$ .) Since the time of the Tevatron, a factor of about 2 has been gained on the critical current density at 4.2 K and 5 T, and values in excess of 3000 A/mm<sup>2</sup> are now obtained in industrial production (20).

The highest dipole field reached on a NbTi magnet is 10.5 T at 1.77 K (21). Magnet designers consider that this is about the limit for NbTi and that to produce higher fields, it is necessary to change the material. The only other material that is readily available at (small) industrial scale is an intermetallic compound of niobium and tin (Nb<sub>3</sub>Sn) belonging to the A15 crystallographic family (18). Nb<sub>3</sub>Sn presents interesting superconducting properties (e.g., its upper critical field at zero temperature and zero strain is in excess of 25 T) (22). However, its formation requires a heat treatment at temperatures up to 700 C for times up to 300 h. Furthermore, once it is reacted, it becomes very brittle, and its superconducting properties are strain-sensitive. Hence, Nb<sub>3</sub>Sn calls for special fabrication techniques, which, so far, have limited its use. In recent years, significant R&D work has been carried out to improve the performance of Nb<sub>3</sub>Sn wires, thanks to the International Thermonuclear Experimental Reactor (ITER) program (23,24).

Although great progress has been made in the development of so-called *high temperature superconductor* (HTS), such as bismuth copper oxides, Bi<sub>2</sub>Sr<sub>2</sub>CaCu<sub>2</sub>O<sub>8</sub>, and (Bi,Pb)<sub>2</sub>Sr<sub>2</sub>Ca<sub>2</sub>Cu<sub>3</sub>O<sub>10</sub>, and yttrium copper oxides, YBa<sub>2</sub>Cu<sub>3</sub>O<sub>7</sub>, these materials are not ready yet for applications requiring low-cost, mass-production, and high-critical current density (25).

### Rutherford-Type Cable

Superconducting accelerator magnet coils are wound from so-called *Rutherford-type* cables. As illustrated in Fig. 2(a), a Rutherford type cable consists of a few tens of strands, twisted together, and shaped into a flat, two-layer, slightly keystoneed cable (26). The strands themselves consist of thousands of superconducting filaments, twisted together and embedded in a matrix of normal metal (18). Except for the cables used in a few R&D model magnets, the filaments are made of NbTi, and

the matrix is high-purity copper. The strand diameter ranges from 0.5 mm to 1.3 mm, and the filament diameter ranges from 5 μm to 15 μm. Figure 2(b) presents a cross-sectional view of a typical SSC strand.

The small radii of curvature of the coil ends preclude the use of a monolithic conductor because it would be too hard to bend. A multi-strand cable is preferred to a single wire for at least four reasons: (1) it limits the piece length requirement for wire manufacturing (a coil wound with a  $N$ -strand cable requires piece lengths which are  $1/N$  shorter than for a similar coil wound with a single wire), (2) it allows strand-to-strand current redistribution in the case of a localized defect or when a quench originates in one strand (27,28), (3) it limits the number of turns and facilitates coil winding, and (4) it limits coil inductance (the inductance of a coil wound with a  $N$ -strand cable is  $1/N^2$  smaller than that of a similar coil wound with a single wire). A smaller inductance reduces the voltage requirement on the power supply to ramp-up the magnets to their operating current in a given time and limits the maximum voltage to ground in case of a quench (see quench protection section). The main disadvantage of using a cable is the high operating current (over a few thousand amperes), which requires large current supplies and large current leads.

The main issues for strand design and manufacturing are (1) copper-to-superconductor ratio, which should not be too small to limit conductor heating in case of a quench while achieving a high overall critical current, (2) filament size, to limit field distortions resulting from superconductor magnetization at low field (see section on field quality); (3) superconductor critical current density, which can be improved by improving pinning and filament uniformity (15) and (4) piece length.

The main issues for cable design and fabrication are (1) construction, which should be large enough to ensure good mechanical stability and high overall current density while leaving enough void (typically a few percent in volume) for liquid helium cooling; (2) control of outer dimensions to achieve suitable coil geometry and mechanical properties; (3) limitation of critical current degradation due to strand and filament deformations at the cable edges (29,30), and (4) control of interstrand resistance, which should not be too small to limit field distortions induced by coupling currents while ramping (see section on field quality) and should not be too large to allow current redistribution among cable strands.

The interstrand resistance can be modified by oxidizing or by coating the strand surface (31,32). Also, a thin, insulating foil (such as stainless steel) can be inserted between the two-strand layers of the cable (33). The strands used in HERA and LHC cables are coated with a silver-tin solder, called *stabrite*. Half of the strands of the Tevatron cable are coated with stabrite, whereas the other half are insulated with a black copper oxide, called *chobid*. LHC, SSC, and RHIC cables rely on natural oxidation. Up to now, no foiled cable has been used in a magnet.

Note that at the end of cabling, the high purity copper of the strand matrix is heavily cold worked and that it may require an annealing procedure.

### Cable Insulation

The main requirements for cable insulation are: (1) good dielectric strength in a helium environment and under high transverse pressure (up to 100 MPa), (2) small thickness (to maximize overall current density in the magnet coil) and good

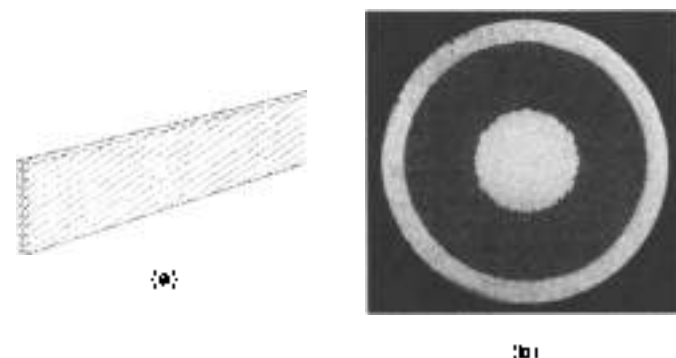


Figure 2. Rutherford-type cable for accelerator magnet: (a) cable sketch and (b) cross-sectional view of a cable strand.

physical uniformity (to ensure proper conductor positioning for field quality), (3) retention of mechanical properties over a wide temperature range (from helium temperature to coil curing temperature—see the discussion that follows), and (4) ability to withstand radiation in an accelerator environment. In addition, the insulation system is required to provide a means of bonding the coil turns together to give the coil a semirigid shape and facilitate its manipulation during the subsequent steps of magnet assembly. It is also desirable that the insulation be somewhat porous to helium for conductor cooling. Note that the dielectric strength of helium gas at 4.2 K is far worse than that of liquid helium and that it degrades significantly with increasing temperature (34).

The insulation of Tevatron, HERA, and UNK magnets, of most SSC magnets, and of the early LHC models is made up of one or two inner layers of polyimide film, wrapped helically with a 50% to 60% overlap, completed by an outer layer of resin-impregnated glass tape, wrapped helically with a small gap. The inner layer is wrapped with an overlap for at least two reasons: (1) the polyimide film may contain pinholes that must be covered (the probability of having two superimposed pinholes in the overlapping layer is very low) and (2) the Tevatron experience has shown that it was preferable to prevent the resin impregnating the glass wrap from entering in contact with the NbTi cable (the energy released by cracks in the resin is believed to be sufficient to initiate a quench) (Ref. 4, p. 784). The outer layer is wrapped with a gap to set up helium cooling channels between coil turns. The resin is of a thermosetting type and requires heat to increase cross-link density and cure into a rigid bonding agent. The curing is realized after winding completion in a mold of very accurate dimensions to control coil geometry and Young's modulus (35).

RHIC magnets and the most recent LHC models use a so-called *all-polyimide* insulation where the outer glass wrap is replaced by another layer of polyimide film with a polyimide adhesive on its surface (36). The all-polyimide insulation has a better resistance to puncture, but the softening temperature of the adhesive can be higher than the temperature needed to cure a conventional resin (225 C for RHIC-type all-polyimide insulation compared to 135 C for SSC-type polyimide-glass insulation).

## MAGNETIC DESIGN

### Field Produced by Simple Current-Line Distributions

**Single Current-Line in Free Space.** Let  $(O, x, y, z)$  designate a rectangular coordinate system, and let  $(-I, R, \theta)$  designate a current-line of intensity  $(-I)$ , parallel to the  $x$  axis, and located at a position  $\mathbf{s} = R \exp(i\theta)$  in the complex  $(O, x, y)$  plane, as represented in Fig. 3(a). (The current-line intensity is chosen to be negative to end up with a positive factor in the right member of Eq. (8).) The magnetic flux density  $\mathbf{B}$ , produced by this current-line in free space, can be computed using Biot and Savart's law. It is uniform in  $z$  and parallel to the  $(x, y)$  plane. It can be expanded into a power series of the form (37)

$$B_x + iB_y = \sum_{k=0}^{\infty} B_k + iA_k z^k \quad \text{for } z = x + iy, |z| < R \quad (7)$$

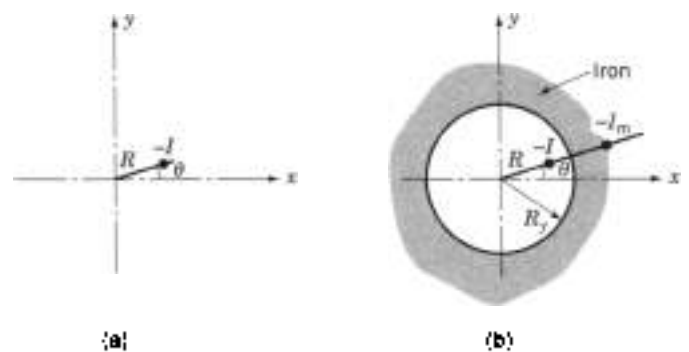


Figure 3. Representations of a single current-line: (a) in a vacuum and (b) inside a circular iron yoke.

where  $B_x$  and  $B_y$  are the  $x$  and  $y$  components of  $\mathbf{B}$  and  $A_k$  and  $B_k$  are constant coefficients, referred to as *skew* and *normal*  $2k$ -pole field coefficients, given by

$$A_k + iB_k = \frac{\mu_0 I}{2\pi R^k} [\cos k\theta e^{ik\theta} + i \sin k\theta e^{ik\theta}] \quad (8)$$

**Single Current-Line within a Circular Iron Yoke.** Let us now assume that the current-line of Fig. 3(a) is located inside a circular iron yoke of inner radius  $R_i$ , as represented in Fig. 3(b). The contribution of the iron yoke to the magnetic flux density can be shown to be the same as that of a mirror current-line, of intensity  $(-I_m)$ , and position  $s_m$  in the complex plane, where

$$I_m = \frac{\mu - 1}{\mu + 1} I \quad \text{and} \quad s_m = \frac{R_i^2}{s^*} \quad (9)$$

Here  $\mu$  designates the relative magnetic permeability of the iron yoke, and  $s^*$  designates the complex conjugate of  $s$ . Note that the mirror image method is applicable only if the iron yoke is not saturated and as long as its permeability is uniform.

**Quadruplet of Current-Lines with Dipole Symmetry.** Using these expressions, the magnetic flux density produced by the quadruplet of current-lines  $(I, R, \theta), (I, R, \pi - \theta), (I, R, \pi + \theta)$  and  $(I, R, -\theta)$ , represented in Fig. 3(a), can be estimated from the power series expansion

$$B_x + iB_y = \sum_{k=0}^{\infty} B_{2k+1} z^{2k} \quad \text{for } z = x + iy, |z| < R \quad (10)$$

where

$$B_{2k+1} = \frac{2\mu_0 I}{\pi R^{2k+1}} [\cos 12k\theta + 1] \quad (11)$$

The first term ( $k = 0$ ) of the series corresponds to a pure normal dipole field parallel to the  $y$  axis. The  $B_{2k+1}$  coefficients are called the *allowed* multipole field coefficients of this current distribution.

**Octuplet of Current-Lines with Quadrupole Symmetry.** Similarly, the magnetic flux density produced by the octuplet of

current-lines represented in Fig. 4(b) is given by

$$B_z = B_0 - \sum_{k=2}^{\infty} B_{4k-2} x^{4k-2} \quad \text{for } z = x + iy, \quad x < R \quad (12)$$

where

$$B_{4k-2} = \frac{2j_0 d}{\pi R^{2k-1}} \cos[(4k-2)\theta] \quad (13)$$

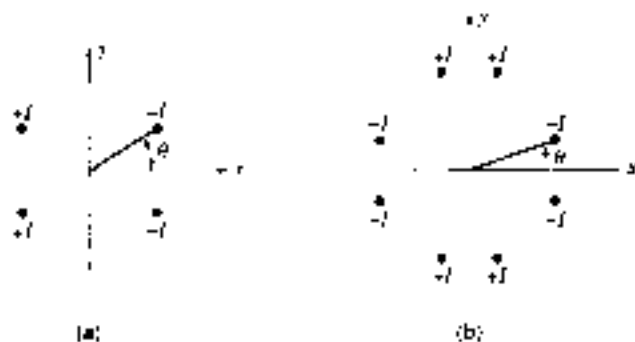
The first term ( $k = 0$ ) of the series corresponds to a pure normal quadrupole field whose axes are parallel to the first and second bisectors. For this current distribution, the allowed multipole field coefficients are the normal  $(4k + 2)$ -pole field coefficients.

**Two-Dimensional Geometry**

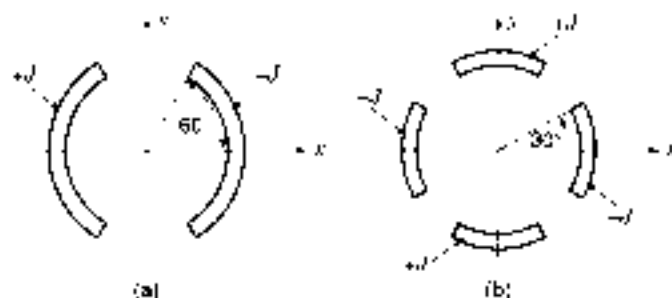
**Symmetry Considerations.** The field computations presented in the previous section showed that current distributions with the symmetries of Fig. 4(a) (i.e., even with respect to the  $x$  axis and odd with respect to the  $y$  axis) were suitable for generating dipole fields, whereas current distributions with the symmetries of Fig. 4(b) (i.e., even with respect to the  $x$  and  $y$  axes and odd with respect to the first and second bisectors) were suitable for generating quadrupole fields. Starting from these premises, the coil geometry can be optimized to obtain the required dipole or quadrupole field strength within the magnet aperture. In addition, in most accelerator designs, it is required that high-order multipole field coefficients be as small as possible. Hence, the coil geometry optimization is also carried out to minimize the contributions from nondipole or nonquadrupole terms.

**cos  $\theta$  Coil Designs.** The coil geometry most commonly used for a dipole magnet is composed of the cylindrical current shells shown in Fig. 5(a). The magnetic flux density produced by such shells can be computed by dividing them up into quadruplets of current-lines having the symmetry of Fig. 4(a) and by summing their contributions over a shell quadrant. It follows that the magnetic flux density is again given by Eq. (10), but the expressions of the multipole field coefficients become

$$B_k = \frac{2j_0 d}{\pi} (R_0 - R_1) \sin \theta_0 \quad (14a)$$



**Figure 4.** Examples of current-line distributions with selected symmetries: (a) quadruplet of current-lines with an even symmetry about the  $x$  axis and an odd symmetry about the  $y$  axis and (b) quadruplet of current-lines with even symmetries with respect to the  $x$  and  $y$  axes and odd symmetries with respect to the first and second bisectors.



**Figure 5.** Current shell approximations for the generation of multipole fields: (a) dipole field and (b) quadrupole field.

and

$$B_{4k-2} = \frac{2j_0 d}{\pi(2k+1)(2k-1)} \left( \frac{1}{R_0^{2k+1}} - \frac{1}{R_1^{2k+1}} \right) \sin[(2k+1)\theta_0] \quad \text{for } k, k \geq 1 \quad (14b)$$

Here,  $R_0$  and  $R_1$  are the inner and outer radii of the shells,  $\theta_0$  is the pole angle, and  $j_0$  is the overall current density, which is assumed to be uniform. Note that  $B_0$  (first allowed multipole field coefficient after  $B_1$  in a current distribution with a dipole symmetry) is nil for  $\theta_0 = \pi/3$ .

Similarly, the coil geometry most commonly used for a quadrupole magnet is made up of the cylindrical current shells shown in Fig. 5(b). The magnetic flux density is here given by Eq. (12), where

$$B_2 = \frac{2j_0 d}{\pi} \ln \left( \frac{R_0}{R_1} \right) \sin 2\theta_0 \quad (15a)$$

and

$$B_{4k-2} = \frac{j_0 d}{\pi k(4k+2)} \left( \frac{1}{R_1^{2k}} - \frac{1}{R_0^{2k}} \right) \sin[(4k-2)\theta_0] \quad \text{for } k, k \geq 1 \quad (15b)$$

Note that  $B_2$  corresponds to the quadrupole field gradient  $g$  and that  $B_0$  (first allowed multipole field coefficient after  $B_1$  in a current distribution with a quadrupole symmetry) is nil for  $\theta_0 = \pi/6$ .

By reference to the conductor distribution around the circular inner bore, such coil geometries are referred to as cos  $\theta$  and cos  $2\theta$  designs. They are very compact and make the most effective use of conductors by bringing them close to the useful aperture.

**Current Shell Approximations.** In practice, the current shells of Figs. 5(a) and 5(b) are approximated by stacking into an arch the slightly keystoneed cables described in the conductor section. High field dipole or high-field-gradient quadrupole magnets usually rely on two coil layers whose contributions add up. Also, wedges are introduced between some of the coil turns to separate the conductors into blocks. The blocks angles are then optimized to eliminate high-order multipole field coefficients (37).

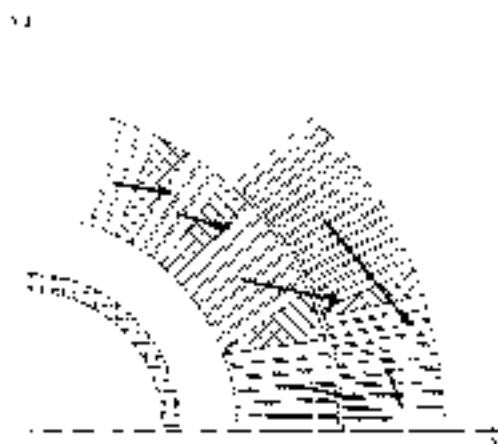


Figure 6. Conductor and Lorentz force distributions in a quadrant of a 50-mm-aperture SSC dipole magnet coil (38).

In the case of Tevatron, HERA, and UNK magnets, the cable keystone angle is large enough to allow the formation of an arch with the desired aperture. Furthermore, each coil turn lies along a radius vector pointing towards the aperture center. In the case of SSC and LHC magnets, the coil aperture is reduced to minimize the volume of superconductor. This results in a keystone angle requirement deemed unacceptable from the point of view of cabling degradation. Hence, in these magnets, the cables are not sufficiently keystoneed to assume an arch shape, and the wedges between conductor blocks must be made asymmetrical to compensate for this lack (36). Also, the coil turns end up being nonradial, as illustrated in Fig. 6, which shows the conductor distribution in a quadrant of a 50-mm-aperture SSC dipole magnet coil (the vectors represent the components of the Lorentz force discussed in the section on mechanical design).

Note that the magnetic flux density produced by the coil of Fig. 16) can be accurately computed by dividing each turn into two rows of elementary current-lines parallel to the  $x$  axis and approximately equal in number to the number of cable strands (Ref. 39, p. 226).

**Iron Yoke Contribution.** The coils of particle accelerator magnets are usually surrounded by an iron yoke, which provides a return path for the magnetic flux while enhancing the central field or field gradient.

As an illustration, let us place the cylindrical current shells of Fig. 5(a) within a circular iron yoke of inner radius  $R_i$ . The contribution of the iron yoke to the normal  $(2k + 1)$ -pole field coefficient  $B_{2k+1}^y$  can be estimated as (Ref. 2, p. 53)

$$B_{2k+1}^y = \frac{\mu - 1}{\mu + 1} \left( \frac{R_i R_o}{R_i^2} \right)^{2k+1} B_{2k+1} \quad (16)$$

where  $\mu$  is the relative magnetic permeability of the iron yoke,  $R_i$  and  $R_o$  are the current shell inner and outer radii, and  $B_{2k+1}$  is the  $(2k + 1)$ -pole field coefficient produced by the current shell alone.

Equation (16) shows that the smaller  $R_i$ , the larger the field enhancement. However, there are two limitations on how close the iron can be brought to the coils: (1) room must be left for

the support structure, and (2) iron saturates for fields above 2 T, resulting in undesirable distortions (see section on field quality).

As already mentioned, the Tevatron magnets use a warm iron yoke (i.e., placed outside the helium containment and vacuum vessel), but starting with HERA magnets, the iron yoke is included within the magnet cold mass. For SSC dipole magnets, the field enhancement due to the iron yoke is of the order of 20%. In LHC magnets, two coil assemblies (powered with opposite polarity) are placed within a common iron yoke. This two-aperture design results in left-right asymmetries in the yoke surrounding each coil assembly taken individually. These asymmetries must be taken into account when calculating the field quality.

**Operating Margin.** Equations 14(a) and 15(a) show that, to achieve high fields and high field gradients, it is desirable to maximize the overall current density in the magnet coil. This can be done by three means: (1) maximizing the superconductor performance, (2) minimizing the copper-to-superconductor ratio in the cable strands, and (3) minimizing the turn-to-turn insulation thickness. As explained in other sections, there are lower bounds on the values of copper-to-superconductor ratio and insulation thickness in order to limit conductor heating in case of quenching and to ensure proper electrical insulation. As for the superconductor, the upper limit is the critical current density at the given temperature and magnetic flux density.

The magnetic flux density to which the conductor is exposed is nonuniform over the magnet coil, but the maximum current-carrying capability of the conductor is determined by the section where the magnetic flux density is the highest. In most cases, this corresponds to the pole turn of the innermost coil layer. Let  $B_p = f(I)$  designate the peak magnetic flux density on the coil as a function of supplied current  $I$ , and let  $I_c = f(B)$  designate the supposedly known cable critical current as a function of applied magnetic flux density  $B$ . The intersection between these two curves determines the maximum quench current of the magnet  $I_{qm}$ .

In practice, magnets must be operated below  $I_{qm}$  so as to ensure that the superconductor is in the superconducting state and to limit the risks of quenching. Let  $I_o$  designate the operating current. Then, the current margin of the magnet  $m_1$  is defined as

$$m_1 = 1 - \frac{I_o}{I_{qm}} \quad (17)$$

The excellent quench performance of the HERA magnets (6) suggests that the current margin can be set to as little as 10%, but it is safer to aim for 20%.

In comparison to other superconducting magnets, such as solenoids for magnetic resonance imaging, a current margin of 10% to 20% is quite small. This implies that accelerator magnets are operated very close to the superconductor critical surface and that they are very sensitive to any kind of disturbances that may cause the magnet to cross the critical surface and lead to a quench.

A peculiarity of a two-layer, cos  $\theta$  dipole magnet coil design is that the peak field in the outermost layer is quite a bit lower than in the innermost layer. Hence, when using the same cable and current for both layers, the inner layer is operated with a

much higher current margin than the inner layer, which can be considered as a waste of costly superconductor. SSC and LHC dipole magnet coils use a smaller conductor for the outer layer than for the inner layer. This results in a higher overall current density in the outer layer and reduces the difference in current margins. Such action is referred to as *conductor grading*. The main disadvantage of grading is that it requires splices between inner and outer layer cables (which, of course, are connected electrically in series and only require one power supply).

**Limits of cos  $\theta$  Design.** The cos  $\theta$  coil design has been very successful until now, with a record dipole field of 13.5 T reached by the LBNL short dipole magnet model using Nb<sub>3</sub>Sn cables at 1.8 K). However, it has two main drawbacks: (1) the coil ends are difficult to make (see the section on coil ends), and (2) due to the Lorentz force distribution, a stress accumulation in the azimuthal direction results in high transverse pressures on the midplane conductors (see Fig. 6). For very high field magnets, requiring the use of A15 (or even possibly HTS) superconductors, which are strain-sensitive, these high transverse pressures can result in significant critical current degradation (40).

Alternative coil designs, which may allow better management of the Lorentz stresses within the magnet coil, are being investigated. As an illustration, Fig. 7 presents a conceptual block or *window/frame* design developed at BNL for a twin-aperture dipole magnet relying only on simple, *racetrack* coils (41). Note, however, that such designs make less effective use of superconductors.

### Coil End Design

One of the main difficulties of the cos  $\theta$  design is the realization of coil ends. In the coil straight section, the conductors run parallel to the magnet axis, but, in the coil ends, the conductors must be bent sharply with small radii of curvature to make U-turns over the beam tube that is inserted within the magnet aperture. This confers to the coil a *saddle shape* as illustrated in Fig. 8.

Sophisticated algorithms have been developed to determine the conductor trajectories that minimize strain energy (42). These algorithms, which often require winding tests to

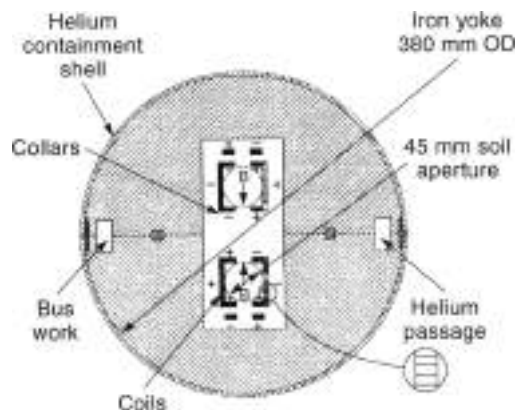


Figure 7. Conceptual block design developed at BNL for a high field, twin-aperture dipole magnet (41).

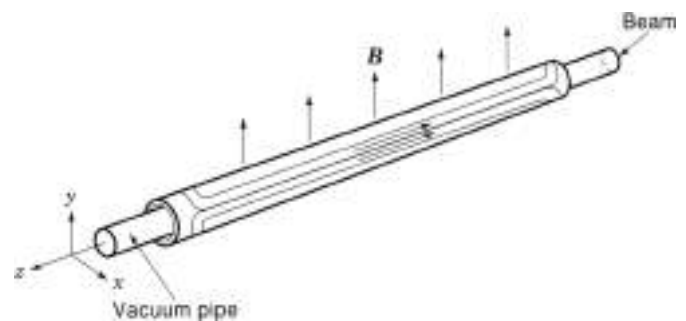


Figure 8. Perspective view of a saddle-shaped coil assembly for a dipole magnet.

determine correction factors, are coupled with electromagnetic computations to minimize field distortions. SSC and LHC magnets use precisely machined end spacers, designed by the optimization programs, which are positioned between conductor blocks (43). In addition, the iron yoke does not extend over the coil ends, to reduce the field on the conductors and ensure that the peak field is located in the coil straight section where the conductors can be better supported.

### Sagitta

To limit the number of coil ends and of magnet interconnects around the accelerator ring, the arc dipole and quadrupole magnets are made as long as possible. The circulation of a charged beam in a dipole magnet, of magnetic length  $l_m$ , results in an angular deflection  $\phi$  of the particle trajectory, which can be estimated as

$$\phi = \frac{q B_1 l_m}{E} = \frac{l_m}{r} \quad (18)$$

Here,  $\phi$  is in radians and  $l_m$  is in meters.  $B_1$  is the dipole magnetic flux density in teslas,  $q$  is the particle charge in units of electron charge, and  $E$  is the particle energy in GeV.

As a result, long dipole magnets must be slightly bent to accompany the particle trajectory. This bending, which is implemented in the  $(x, z)$  plane, is referred to as *sagitta*.

## FIELD QUALITY

### Multipole Expansion

Except near the short coil ends, the magnetic flux density produced in the horn of a particle accelerator magnet can be considered as two-dimensional. The power series expansion of Eq. (7) is usually rewritten in the more convenient form

$$B_x + iB_y = B_0 10^{-4} \sum_{n=1}^{\infty} b_n + i a_n \left( \frac{z}{R_0} \right)^{n-1} \quad (19)$$

for  $z = x + iy$ ,  $|z| < R_0$

where  $B_x$  and  $B_y$  are the  $x$  and  $y$  components of the magnetic flux density.  $R_0$  is a reference radius representative of the maximum beam size ( $R_0$  was 10 mm for the SSC and is now 17 mm for the LHC),  $B_0$  is the absolute value of the dipole or quadrupole component at  $R_0$ ,  $a_n$  and  $b_n$  are the dimensionless skew and

normal 2n-pole coefficients, and  $R_c$  is the coil inner radius. Note the presence of the  $10^{-4}$  scale factor.

Given the symmetries of current distributions in magnet coil assemblies, and as explained in the previous section, only selected normal multipole coefficients are expected to be nonzero. These allowed multipole coefficients can be tuned up by iterating on the electromagnetic design. In practice, however, nonuniformities in material properties and manufacturing errors result in symmetry violations that produce unallowed multipole coefficients. For instance, a top/bottom asymmetry in a dipole magnet produces a nonzero skew quadrupole coefficient ( $a_2$ ), whereas a left/right asymmetry produces a nonzero normal quadrupole coefficient ( $b_2$ ). These unwanted coefficients can be eliminated only by improving material selection, tooling, and assembly procedures.

### Field Quality Requirements

From the accelerator point of view, the beam optics is primarily governed by integrated field effects over the magnet ring. The main field quality requirements are (1) suitable dipole field integral and small dipole field angle variations (the former to ensure that the integrated bending angle over the magnet ring is  $(2\pi)$  and the latter to ensure that the particle trajectory is planar); (2) accurate quadrupole alignment and suitable quadrupole field integral (the former to avoid coupling of particle motions along the x and y axes and the latter to ensure proper focusing); and (3) small high-order multipole coefficients (to ensure large beam dynamic aperture). In the case of high-order multipole coefficients, it is customary to specify tables of mean values and standard deviations over the entire magnet population (44). The tables of mean values are referred to as systematic multipole specifications, whereas those of standard deviations are referred to as random multipole specifications. The specified values are all expressed at the reference radius  $R_c$ .

In large machines such as SSC or LHC, the dipole and quadrupole field integrals must be controlled with a relative precision of the order of  $10^{-8}$ . The variations in dipole field angles must be kept within a few milliradians, and the tolerance on quadrupole alignment is of the order of 0.1 mm. Systematic and random multipole specifications are given up to the 15th or 20th pole and get tighter with increasing pole order. For SSC magnets at 10 mm, the specifications went from a few tenths of a unit for low-order coefficients to a few thousandths of a unit for higher-order coefficients.

### Geometric Errors

**Types of Geometric Errors.** The specifications on multipole coefficients require that the individual conductors and the yoke surrounding the coil assembly be positioned with a very good accuracy (typically a few hundredths of a millimeter in the two-dimensional cross section). Improper positioning results in geometric errors that distort the central field and produce unwanted multipole coefficients.

The geometric errors can be classified in at least five categories: (1) errors in coil inner and outer radii and in yoke inner radius; (2) errors in coil pole angle, wedge angle, and conductor angular distribution; (3) symmetry violations in coil assembly; (4) centering errors with respect to the iron yoke; and (5) residual twist of magnet assembly.

**Effects of Azimuthal Coil Size Mismatch.** A common cause of geometric error is a mismatch between the azimuthal sizes of the various coils constituting a coil assembly. Such mismatch results in displacements of the coil assembly symmetry planes that produce nonzero, low-order unallowed multipole coefficients (45). For instance, a mismatch between the azimuthal sizes of the top and bottom coils used in a dipole magnet coil assembly causes an upward or downward displacement of the coil parting planes, which produces a nonzero skew quadrupole coefficient  $a_2$ . Similarly, a systematic mismatch between the left and right sides of the coils used in a dipole magnet coil assembly causes a rotation of the coil parting planes, which produces a nonzero skew sextupole coefficient  $a_3$ . A systematic  $a_3$  can be limited by randomly mixing coil production, whereas the occurrence of a systematic  $a_1$  can be avoided only by correcting tooling.

### Iron Saturation

When the field in the iron yoke is less than 2 T, the relative magnetic permeability of the yoke can be considered as very large and uniform, and the iron contribution to the central field increases linearly as a function of transport current in the magnet coil. For fields above 2 T, parts of the iron start to saturate, and their relative magnetic permeability drops. As a result, the iron contribution becomes a less-than-linear function of transport current. This relative decrease in iron contribution appears as a sag in the magnet transfer function (38). (The transfer function is defined as the ratio of  $B_z$  to the transport current.) The transfer function sag can exceed a few percent in dipole magnets but is usually negligible in quadrupole magnets.

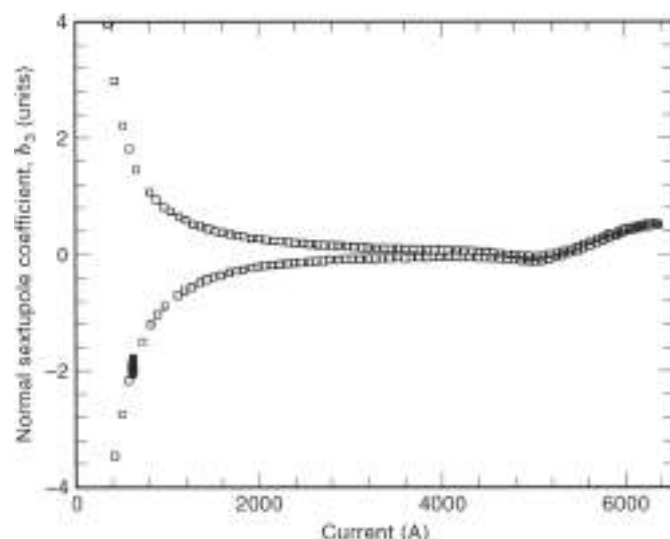
In the case of a single-aperture magnet with a symmetrical iron yoke, the saturation first occurs in the pole areas producing a positive shift in normal sextupole coefficient  $b_3$ . At higher currents, the saturation reaches the midplane areas, producing a negative shift in  $b_3$ , which partially compensates for the effects of pole saturation. The midplane saturation can be forced to occur sooner by punching notches (i.e., removing material at appropriate locations) in the yoke. As an illustration, Fig. 9 presents measurements of  $b_3$  as a function of current in the central part of a SSC dipole magnet prototype. The measurements above 8 kA clearly show the effect of pole saturation at high currents; the origin of the hysteresis is explained in the next section.

In the case of a twin-aperture dipole, the central part of the yoke saturates before the outer parts, resulting in left/right asymmetries in the yoke contributions to each aperture, which affect the normal quadrupole coefficient  $b_2$ . The saturation effects in  $b_2$  are of opposite sign in the two apertures.

In any case, the iron contribution depends on the packing factor of the yoke laminations, which must be tightly controlled over the magnet length. Also, the iron yoke must be carefully aligned to limit magnet assembly twist.

### Superconductor Magnetization

**Critical State Model.** According to the so-called critical state model, bipolar magnetization currents are induced at the periphery of the superconducting filaments in the cable strands each time the field to which the filaments are exposed is varied (46). The magnetization currents distribute themselves with a density equal to the superconductor critical current density at the given temperature and field  $J_c$ , in order to screen the



**Figure 9.** Measurements of normal sextupole coefficient  $b_3$  as a function of current at the central part of a SSC dipole magnet showing the hysteresis resulting from superconductor magnetization and the distortions at high currents resulting from iron saturation.

filament cores from the applied field change. Unlike regular eddy currents, the magnetization currents do not depend on the rate of field variations. Also, because they can flow with zero resistance, they do not decay as soon as the field ramp is stopped. They are called *persistent magnetization currents*.

**Effects of Superconductor Magnetization.** When an accelerator magnet is cycled in current, the bipolar shells of magnetization currents induced in the filaments behave as small magnetic moments which contribute to—and distort—the central field. The magnetic moments depend on  $J_c$  and are proportional to filament diameter. Their distribution follows the symmetries of the transport-current field (i.e., the field produced by the transport current in the magnet coil), and, if the superconductor properties are uniform, only the allowed multipole coefficients are affected. Computer models that can accurately predict the field distortions resulting from superconductor magnetization have been developed [47].

The field distortions are the most significant at low transport current, where the transport-current field is low and  $J_c$  is large. They are progressively overcome as the transport-current field increases and  $J_c$  diminishes and become negligible at high transport current. They change sign and regain influence as the transport current is ramped down. As a result, the allowed multipole coefficients exhibit sizable hystereses as a function of transport current, which depend on magnet excitation history. This is illustrated in Fig. 9, which shows measurements of  $b_3$  as a function of current in the central part of a SSC dipole magnet. In Fig. 9, the magnetization effects can be seen at currents below 3 kA (as explained in the previous section, the distortions at high field result from iron yoke saturation).

The field distortions resulting from superconductor magnetization are one of the major drawbacks of using superconducting magnets in a particle accelerator. They can be reduced by reducing filament size (typically, to 5  $\mu\text{m}$  for SSC and LHC strands), but they cannot be eliminated. The powering cycle of the magnets must be adapted to avoid brutal jumps between the

two branches of the multipole coefficient hystereses while the beam circulates. Also, elaborate beam optics correction schemes must be developed. This can include superconducting, high-order multipole magnets (chapter 9 of Ref. 2).

**Time Decay.** In addition, the effects of superconductor magnetization are not indefinitely persistent, but exhibit a slow time decay, which, at low transport current, can result in significant drifts of the allowed multipole coefficients (48,49). These drifts are particularly disturbing during the injection phase of machine operation, where the magnet current is maintained at a constant and low level for some period of time (50). Also, they complicate the early stages of acceleration. For, as the current is increased at the end of injection, the drifting multipoles rapidly snap back to values on the hysteresis curves (51). Part of the observed time decay can be attributed to flux creep in the superconductor (52), but flux creep cannot account for the large drifts observed after a high current cycle (49). The nature of the other mechanisms that may be involved is not well understood.

#### Coupling Currents

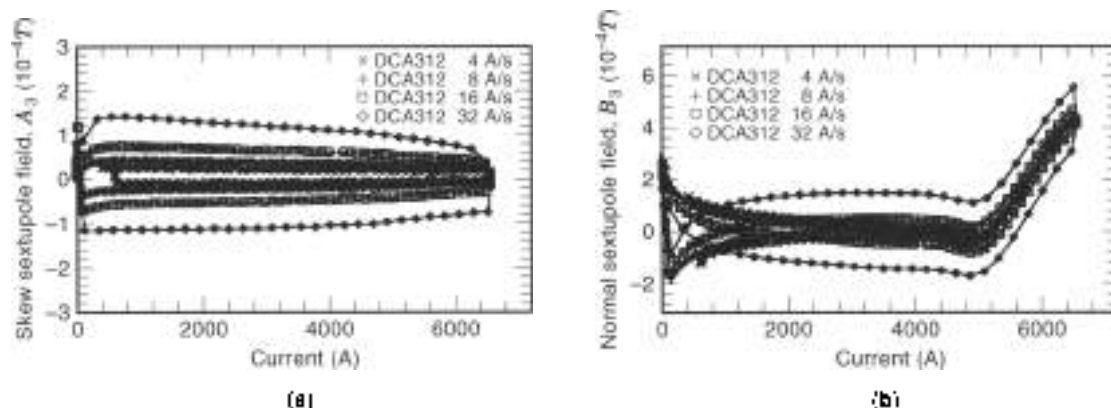
As described in the conductor section, accelerator magnet coils are wound from Rutherford-type cables, which consist of a few tens of strands twisted together and shaped into a flat, two-layer, slightly keystoned cable. The cable mid-thickness is smaller than twice the strand diameter, which results in strand deformation and large contact surfaces at the crossovers between the strands of the two layers. Furthermore, and as explained in the mechanical design section, the coils are precompressed azimuthally during magnet assembly. Large pressures that keep the strands firmly in contact are thus applied perpendicularly to the cable. The large contact surfaces and the high pressures can result in low contact resistances at the strand crossovers.

In the steady state, the transport current flows in the superconducting filaments, which offer no resistance. When the cable is subjected to a transverse varying field, the network of low interstrand resistances allows the formation of current loops, which are superimposed on the transport current. The loop currents, referred to as *interstrand coupling currents*, circulate along the superconducting filaments and cross over from strand to strand through the interstrand resistances. Unlike persistent magnetization currents, the interstrand coupling currents are directly proportional to the rate of field variations, and they start to decay as soon as the field ramp is stopped.

Interstrand coupling currents have three main effects on magnet performance: (1) quench current degradation (for they are superimposed on the transport current), (2) heat dissipation (when crossing the interstrand resistances), and (3) field distortions. This last issue is the most critical for accelerator magnet applications.

The coupling current contribution to the central field does not depend on transport current and increases linearly as a function of current ramp rate. If the interstrand resistance is uniform throughout the coil assembly, the coupling current distribution follows the symmetries of the transport-current field, and only the allowed multipole coefficients are affected. In practice, however, there can be large end-to-end differences as well as large nonuniformities within the coils themselves, which result in sizable effects in the unallowed multipole coefficients.





**Figure 10.** Effects of interstrand coupling currents on multipole field coefficients as measured as a function of ramp rate in the central part of a SSC dipole magnet (39): (a) skew sextupole field coefficient  $A_3$  and (b) normal sextupole field coefficient  $B_3$ . The transport-current contribution is subtracted from the data.

This is illustrated in Fig. 10 (a, b), which presents plots of the skew and normal sextupole field coefficients ( $A_3$  and  $B_3$ ) as functions of current, measured at various ramp rates in the central part of a SSC dipole magnet prototype. (Note that the transport-current contribution has been subtracted from the data.) No particular treatment (such as stabrite) was applied to the strands of the cable used in this prototype.

The effects of interstrand coupling currents can be limited by ensuring that the interstrand resistances are not too low. However, and as mentioned in the conductor section, the interstrand resistances should not be too large either to allow some possibility of current redistribution among cable strands.

### Longitudinal Periodicity

When measuring the field with fine spatial resolution along the axis of an accelerator magnet, all multipole coefficients appear to exhibit periodic oscillations (53,54). The amplitude of the oscillations varies as a function of space, transport current, excitation history, and time, but the wavelength is always approximately equal to the twist pitch length of the cable used in the innermost coil layer.

The longitudinal periodic oscillations are believed to result from imbalances in the current distribution among cable strands. The current imbalances may have at least three origins: (1) nonuniformities in the properties of cable strands, (2) nonuniformities in the solder joints connecting the coils in series to the current leads, and (3) large and long-lasting interstrand coupling current loops superimposed on the transport current (55). Such current loops could be induced by spatial variations in the time derivative of the field to which the cable is exposed as it turns around the coil ends or exits toward the current leads (56–58).

The oscillation wavelength is too short to affect beam optics but may be an issue for magnetic measurements. It is recommended that the measurements be averaged over an integer number of cable pitch lengths. Also, the slow decay of the large interstrand coupling current loops associated with these periodic oscillations may contribute to the drifts of the allowed multipole coefficients observed at low and constant transport current (see the section on superconductor magnetization (59).

## MECHANICAL DESIGN

### Support Against the Lorentz Force

**Components of the Lorentz Force.** The high currents and fields in an accelerator magnet coil produce a large Lorentz force on the conductors. In a dipole coil, the Lorentz force has three main components, which are represented in Fig. 6 (38,60): (1) an azimuthal component, which tends to squeeze the coil toward the coil assembly midplane (which, in the coordinate system defined previously, corresponds, for a dipole magnet, to the horizontal ( $x, z$ ) plane), (2) a radial component, which tends to bend the coil outwardly, with a maximum displacement at the coil assembly midplane (along the horizontal  $x$  axis for a dipole magnet), and (3) an axial component, arising from the solenoidal field produced by the conductor turnaround at the coil ends and which tends to stretch the coil outwardly along the  $z$  axis.

**Stability against Mechanical Disturbances.** Because accelerator magnets are operated close to the critical current limit of their cables, their minimum quench energy (MQE), defined as the minimum energy deposition needed to trigger a quench, is very small. As a matter of fact, the MQE of accelerator magnets is of the same order of magnitude as the electromagnetic work produced by minute wire motions in the coil (61). If the motions are purely elastic, no heat is dissipated, and the coil remains superconducting, but if the motions are frictional, the associated heat dissipation may be sufficient to initiate a quench. This leaves two possibilities: either to prevent wire or coil motion by providing a rigid support against the various components of the Lorentz force or to reduce to a minimum the friction coefficients between potentially moving parts of magnet assembly.

**Conceptual Design.** The mechanical design concepts used in present accelerator magnets are more or less the same and were developed at the time of the Tevatron (4,62). In the radial direction, the coils are confined within a rigid cavity defined by laminated collars, which are locked around the coils by means of keys or tie rods. In the azimuthal direction, the collars are assembled so as to precompress the coils. In the axial direction,

the coils either are free to expand or are restrained by means of stiff end plates.

The use of laminated collars, pioneered at the Tevatron, was a real breakthrough in achieving a rigid mechanical support while keeping tight tolerances over magnet assemblies, which are a few meters in length and which must be mass produced. The laminations are usually stamped by a fine blanking process allowing a dimensional accuracy on the order of one hundredth of a millimeter to be achieved.

### Azimuthal Precompression

**Preventing Collar Pole Unloading.** As described previously, the azimuthal component of the Lorentz force tends to squeeze the coil toward the midplane. At high fields, it may happen that the coil pole turns move away from the collar poles, resulting in variations of the coil pole angle that distort the central field and creating a risk of mechanical disturbances. To prevent conductor displacements, the collars are assembled and locked around the coils so as to apply an azimuthal precompression. The precompression is applied at room temperature and must be sufficient to ensure that, after cool-down and energization, there is still contact between coil pole turns and collar poles.

**Precompression Requirement.** To determine the proper level of room temperature azimuthal precompression, at least three effects must be taken into account: (1) stress relaxation and insulation creep following the collaring operation, (2) thermal shrinkage differentials between coil and collars during cool-down (if any), and (3) stress redistribution resulting from the azimuthal component of the Lorentz force. In addition, the collaring procedure must be optimized to ensure that the peak pressure seen by the coils during the operation (which may be significantly higher than the residual precompression) does not overstress the insulation (Ref. 60, p. 1326).

The precompression loss during cool-down,  $\Delta\sigma$ , can be estimated from

$$\Delta\sigma = E_{\theta}(\alpha_c - \alpha_s) \quad (20)$$

where  $E_{\theta}$  is the coil Young's modulus in the azimuthal direction, and  $\alpha_c$  and  $\alpha_s$  are the thermal expansion coefficients of the coil (in the azimuthal direction) and of the collars, integrated between room and operating temperatures. Note that Eq. (20) is derived with the assumptions that  $E_{\theta}$  does not depend on temperature and that the collars are infinitely rigid.

**Choice of Collar Material.** To limit cool-down loss, it is preferable to use for the collars a material whose integrated thermal expansion coefficient matches more or less that of the coil. For NbTi coils with polyimide-glass or all-polyimide insulation, this suggests aluminum alloy (see Table 2). However, and as will be described in the next section, it is also desirable that the collars be as rigid as possible or have an integrated thermal expansion coefficient approaching that of the low carbon steel used for the yoke. This favors austenitic stainless steel, which has a lower integrated thermal expansion coefficient and whose Young's modulus is 195 GPa at room temperature and 203 GPa at 4.2 K, compared to 72 GPa at room temperature and 80 GPa at 4.2 K for aluminum alloy.

When assessing the respective merits of austenitic stainless steel and aluminum alloy, note that austenitic stainless steel

presents a better resistance to stress cycling at low temperature (53), but that it has a higher density (7800 kg/m<sup>3</sup> compared to 2800 kg/m<sup>3</sup> for aluminum alloy) and is more expensive.

There is no ideal choice between stainless steel and aluminum alloy, and magnets with both types of collar materials have been built: HERA dipole magnets and most LHC dipole magnet prototypes use aluminum alloy collars, whereas Tevatron dipole magnets and most SSC dipole magnet prototypes rely on stainless steel collars. In any case, and whichever collar material is chosen, a thorough mechanical analysis of the structure under the various loading conditions is required.

### Radial Support

**Limiting Radial Deflections.** As described previously, the radial component of the Lorentz force tends to bend the coil outwardly, with a maximum displacement at the coil assembly midplane. At high fields, this bending results in shear stresses between coil turns and in an ovalization of the coil assembly (along the horizontal  $x$  axis for a dipole magnet), which generates field distortions. To prevent displacements or deformations, the radial deflections of the coil assembly must be limited to, typically, less than 0.05 mm.

**Seeking Yoke Support.** The main support against the radial component of the Lorentz force is provided by the collars, whose stiffness and radial width must be optimized to limit collared-coil assembly deflections. However, in the magnetic design of high field magnets, the field enhancement provided by the iron yoke is maximized by bringing it as close as possible to the coil. This reduces the space left for the collars, whose rigidity then becomes insufficient to hold the Lorentz force, and the yoke and helium containment shell must also be used as part of the coil support system.

The mechanical design of magnets where the yoke is needed to support the collared-coil assembly is complicated by the fact that the collar material (stainless steel or aluminum) shrinks more during cool-down than the low carbon steel used for the yoke (see Table 2). This thermal shrinkage differential must be compensated for to ensure that, when the magnet is cold and energized, there is a proper contact between the collared-coil assembly and the yoke along the axis of maximum potential displacements. Such contact limits the deformations of the collared-coil assembly and allows a partial transfer (up to 50% in some LHC dipole magnet prototypes) of the radial component of the Lorentz force to the yoke and the shell.

The aforementioned thermal shrinkage differential  $\Delta r$  can be estimated as

$$\Delta r = R_{\theta}(\alpha_c - \alpha_y) \quad (21)$$

Table 2. Integrated Thermal Expansion Coefficients between 4.2 K and Room Temperature (10<sup>-7</sup> m/m)

|                                   |      |
|-----------------------------------|------|
| Low carbon steel                  | 2.0  |
| Stainless steel (304/316)         | 2.9  |
| Copper (OFHC)                     | 3.1  |
| Aluminum                          | 4.2  |
| Insulated cable (polyimide/glass) | 5.1* |
| Insulated cable (all polyimide)   | 5.6  |

\* Transverse direction; SSC inner cable

where  $R_{co}$  is the collar outer radius and  $\alpha_{yA}$  is the thermal expansion coefficient of the yoke, integrated between room and operating temperatures.

To limit contact loss due to thermal shrinkage differentials, it is preferable to use for the collars a material whose integrated thermal expansion coefficient approaches that of low carbon steel. This suggests the use of austenitic stainless steel (see Table 2). However, and as was described in the section on choice of collar material, it is also desirable to limit the cool-down loss of coil precompression, which favors the use of aluminum alloy

**Mechanical Design with Fully Mated Yoke Assembly.** To facilitate assembly, the yoke of dipole magnets is usually split into two halves, which are mounted around the collared-coil assembly. The shell, which is also made up of two halves, is then placed around the yoke and welded. If the thermal shrinkage differential between collar and yoke is not too large (as in the case of stainless steel collars), it can be compensated for by designing and assembling the structure so that the two yoke halves apply a compressive load over selected areas of the collared-coil assembly. This compressive load is obtained by introducing a shrinkage allowance into the geometry of either the collars or the yoke and by welding the shell so as to press radially onto the two yoke halves and as to force them to mate at room temperature. During cool-down, the collared-coil assembly shrinks away from the two yoke halves, which remain fully mated. This results in a progressive decrease of the compressive load exerted by the yoke, but a suitable contact can be maintained over the designated areas of the collared-coil assembly.

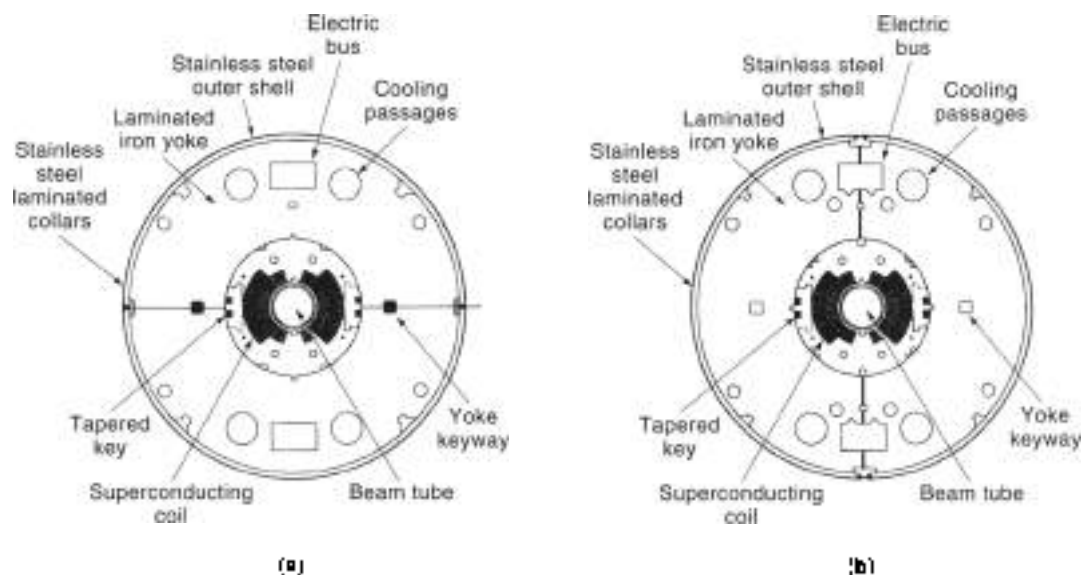
In practice, the compressive load provided by the yoke is directed along a given axis. The choice of the axis drives the choice of yoke split orientation. The SSC dipole magnet prototypes built at BNL use a horizontally split yoke with a yoke-collar compressive load directed along the vertical  $y$  axis as shown in Fig. 11(a), while the SSC dipole magnet prototypes built at FNAL use a vertically split yoke with a yoke-collar compressive load directed along the horizontal  $x$  axis as shown in Fig. 11(b) (64). Both types of magnets performed very well.

**Mechanical Design with Yoke Midplane Gap at Room Temperature.** For large thermal shrinkage differentials (as in the case of aluminum collars), the yoke-collar compressive load required at room temperature for a full compensation would overstress the collared-coil assembly, and a more sophisticated mechanical design must be used. The twin-aperture LHC dipole magnet prototypes with aluminum collars rely on a two-piece, vertically split yoke with an open gap at room temperature and a welded outer shell made of a material (stainless steel or aluminum) that shrinks more during cool-down than the low-carbon steel yoke (65).

In these magnets, the yoke is designed so that, when placed around the collared-coil assembly at room temperature with no pressure applied to it, there remains an opening between the two yoke halves of the order of the expected thermal shrinkage differential. The yoke midplane gap is then closed in two stages: (1) during shell welding, as a result of the compressive load arising from weld shrinkage, and (2) during cool-down, as a result of the compressive load arising from thermal shrinkage differential between yoke and shell. The initial gap closure during shell welding is limited to avoid overstressing the collared-coil assembly. The closure is completed during cool-down thanks to the radial pressure exerted by the shell, which forces the two yoke halves to follow the shrinkage of the collared-coil assembly and to maintain contact along the horizontal  $x$  axis. The yoke midplane gap must be fully closed at the end of cool-down to ensure that the structure is very rigid and to avoid any risk of oscillation during energization.

A crucial issue in such a design is the ability to perform the shell welding operation in a reproducible way during mass production so as to achieve the desired yoke midplane gap value at room temperature and to keep a tight tolerance on this value (of the order of 0.1 mm). As we have seen, a gap too close may result in coil overstressing at room temperature, whereas a gap too open may result in contact loss during cool-down.

In some LHC prototypes, the closure of the yoke midplane gap is controlled by means of aluminum spacers located between the two yoke halves (66). The spacers are dimensioned



**Figure 11.** SSC dipole magnet cross-sections (64): (a) BNL-style with horizontally split yoke and (b) FNAL-style with vertically split yoke.

to have a spring rate similar to that of the collared-coil assembly, and they prevent the gap from closing at room temperature. During cool-down, however, they shrink more than the yoke and cease to be effective.

**RHIC Magnets.** In RHIC magnets, collar and yoke designs are altogether simplified by replacing the collars by reinforced plastic spacers and by using directly the yoke to precompress the one-layer coils (67). It remains to be seen if this structure could be scaled up to higher-field magnets.

#### End Support

As described previously, the axial component of the Lorentz force tends to stretch the coil outwardly along the  $z$  axis. In magnets where the yoke is not needed to support the collared-coil assembly, a clearance can be left between the two. If the axial stresses resulting from the Lorentz force do not exceed the yield stress of the coil, it is possible to let the collared-coil assembly expand freely within the iron yoke. This is the case of the quadrupole magnets designed at Commissariat à l'Énergie Atomique/Saclay for HERA, SSC and LHC (68). However, in magnets where there is contact between collar and yoke, it is essential to prevent stick/slip motions of the laminated collars against the laminated yoke and to provide a stiff support against the axial component of the Lorentz force (60,69). The ends of SSC and LHC dipole magnet coils are contained by thick stainless steel end plates welded to the shell.

## MAGNET COOLING

### Superconductor Critical Temperature

The superconducting state exists only at temperatures below the so-called *critical temperature*  $T_c$ . For NbTi,  $T_c$  can be estimated as a function of applied magnetic flux density  $B$  using

$$T_c(B) = T_{c0} \left( 1 - \frac{B}{B_{c2}} \right)^2 \quad (22)$$

where  $T_{c0}$  is the critical temperature at zero field (about 9.2 K) and  $B_{c2}$  is the upper critical magnetic flux density at zero temperature (about 14.5 T).

**Boiling and Supercritical Helium Cooling.** To achieve low temperatures and ensure stable operations against thermal disturbances, the accelerator magnet coils are immersed in liquid helium (70). Helium is a cryogenic fluid whose pressure-temperature phase diagram is presented in Fig. 12. Its boiling temperature is 4.22 K at 1 atm (1 atm = 0.1 MPa).

Small superconducting magnet systems usually rely on boiling helium at 1 atm (71). Boiling helium offers the advantage that, as long as the two phases are present, the temperature is well determined. However, in large-scale applications, such as superconducting particle accelerators, the fluid is forced to flow through numerous magnet cryostats and long cryogenic lines, where heat leaks are unavoidable. The heat leaks result in increases in vapor contents and create a risk of gas pocket formation that may block circulation.

The aforementioned difficulty can be circumvented by taking advantage of the fact that helium exhibits a critical point at a temperature of 5.2 K and a pressure of 0.226 MPa (see Fig. 12).

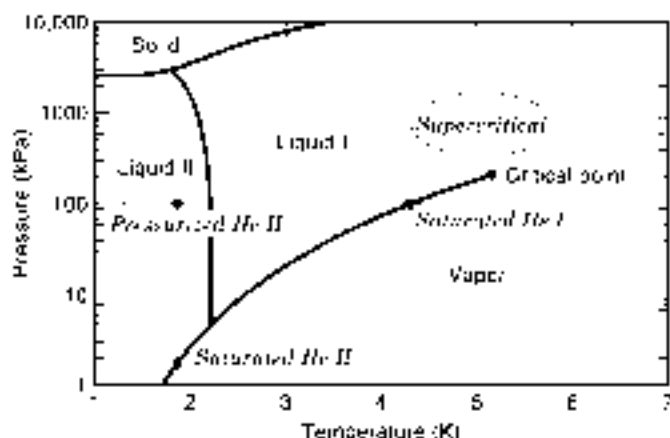


Figure 12. Pressure-temperature phase diagram of helium (71).

For temperatures and pressures beyond the critical point, the liquid and vapor phases become indistinguishable. The single-phase fluid, which is called *supercritical*, can be handled in a large system without risk of forming gas pockets. However, its temperature, unlike that of boiling helium, is not constant and may fluctuate as the fluid circulates and is subjected to heat losses.

The cryogenic systems of the Tevatron, HERA, and RHIC, and that designed for the SSC, combine single-phase and two-phase helium (71). In the case of the Tevatron and HERA, the insides of the magnet cold masses are cooled by a forced flow of supercritical helium, while two-phase helium is circulated in a pipe running at the cold mass periphery (around the collared-coil assembly for Tevatron magnets, in a bypass hole in the iron yoke for HERA magnets). In the case of the SSC, it was planned to only circulate supercritical helium through the magnet cold masses, while *removers*, consisting of heat exchangers using two-phase helium as primary fluid, would have been implemented at regular intervals along the cryogenic lines. The cryogenic system used for the RHIC is inspired by that of the SSC. In all these schemes, the boiling liquid is used to limit temperature rises in the single-phase fluid.

### Superfluid Helium Cooling

A peculiarity of helium is the occurrence of *superfluidity* (70). When boiling helium is cooled down at 1 atm, it stays liquid until a temperature of the order of 2.17 K, where a phase transition appears. For temperatures below 2.17 K at 1 atm helium loses its viscosity and becomes a superconductor of heat. This property, unique to helium, is called *superfluidity*. Superfluidity is very similar to superconductivity, except that, instead of electrical conductivity, it is the thermal conductivity that becomes infinite. The transition temperature between the liquid and superfluid phases depends on pressure. It is called the *lambda temperature*  $T_\lambda$ .

The LHC magnets are cooled by superfluid helium, and their operating temperature is set at 1.9 K (72). Decreasing the temperature improves the current-carrying capability of NbTi dramatically and allows higher fields to be reached. (For NbTi, the curve of critical current density as a function of field is shifted by a about +3 T when lowering the temperature from 4.2 K to 1.9 K.) The feasibility of a large-scale cryogenic installation relying on superfluid helium has been demonstrated by Tore Supra, a superconducting tokamak built at Commissariat

à l'Énergie Atomique, Cadarache near Aix en Provence in the South of France and operating reliably since 1988 (73).

### Magnet Crystal

To maintain the magnet cold masses at low temperature, it is necessary to limit heat losses. There are three mechanisms of heat transfer (74): (1) convection, (2) radiation, and (3) conduction. The convection losses are eliminated by mounting the cold masses into cryostats, which are evacuated (71,75). The radiation losses, which scale in proportion with the effective emissivities of the surfaces facing each other and with the fourth power of their temperatures, are reduced by surrounding the cold masses with blankets of multilayer insulation and thermal shields at intermediate temperatures. The main sources of conduction losses are the support posts, the power leads, and the cryogenic feedthroughs, which are designed to offer large thermal resistances.

## QUENCH PERFORMANCE

As explained in the operating margin section, the maximum quench current  $I_{mq}$  of a magnet at a given operating temperature can be estimated from the critical current of the cable and the peak field on the coil. It corresponds to the ultimate current-carrying capability of the cable and can be raised only by lowering the operating temperature.

When energizing a superconducting magnet, the first quenches usually occur at currents below  $I_{mq}$  (chapter 5 of Ref. 76). In most cases, however, it appears that, upon successive energizations, the quench currents gradually increase. This gradual improvement is called the magnet's *training*. The training often leads to a stable *plateau* corresponding to the maximum quench current.

Quenches below the expected maximum quench current have at least four origins: (1) energy deposition in the magnet coil resulting from frictional motions under the Lorentz force, (2) energy deposition from synchrotron radiation and beam losses, (3) heat dissipation from coupling currents in the cable, and (4) current imbalances among cable strands. Quenches of the first origin reveal flaws in the mechanical design or in the assembly procedures that must be analyzed and corrected. The effects of synchrotron radiation can be reduced by implementing an intercepting screen within the beam tube. Coupling losses and current imbalances are only of concern for fast current cycles.

When operating an accelerator made up of several hundred or even several thousand superconducting magnets, it cannot be tolerated that magnets quench at random. Hence, the magnets must be designed with a safe margin above the maximum operating current of the machine. In addition, systematic tests must be carried out before installing the magnets in the tunnel to ensure that their quench performance is adequate and does not degrade upon extended current and thermal cycling (77).

## QUENCH PROTECTION

### Conductor Heating

Although most R&D programs have been successful in developing magnet designs that can be mass-produced and meet ac-

celerator requirements, quenches do occur in accelerator operations. These quenches must be handled in order to avoid any damage of the quenching magnet, to ensure the safety of the installation, and to minimize down time.

The most damaging effect of a quench is that, once a volume of conductor has switched to the normal resistive state, it dissipates power by the Joule effect (Chapter 9 of Ref. 76). Most of this power is consumed locally in heating up the conductor. In a very short time (typically a few tenths of a second), the conductor temperature can reach room temperature, and, if the magnet is not discharged, keep on increasing.

**Maximum-Temperature Requirement.** The temperature rise subsequent to a quench must be limited for at least three reasons: (1) to restrict the thermal stresses induced in the quenching coil, (2) to prevent degradation of superconductor properties, and (3) to avoid insulation damage.

For most materials, thermal expansion starts to be significant for temperatures above 100 K. The critical current density of NbTi is affected by exposure to temperatures above 250 K. The extent of degradation depends on the temperature level and on the duration of the exposure: at 250 K, it takes of the order of 1 h for significant degradation to occur, while it may take less than a minute at 400 K to 450 K (78). Finally, the polyimide materials used to insulate NbTi cables lose most of their mechanical properties for temperatures above 500 K.

It follows that an upper limit for conductor heating subsequent to a quench is 400 K. Most magnets are designed not to exceed 300 K to 400 K, and whenever possible, the limit should be set at 100 K.

### Protecting a Quenching Magnet

The source of conductor heating in a quenching magnet is power dissipation by the Joule effect. Power keeps being dissipated as long as there is current in the magnet coil. To eliminate the heat source and limit the temperature rise, it is thus necessary to ramp the current down.

To discharge a quenching magnet, all its stored magnetic energy must be converted into resistive power. If the zone where the conductor has switched to the normal state remains confined to a small volume, there is a risk that a large fraction of the stored energy will be dissipated in this small volume. In the case of a string of magnets connected electrically in series, it may even happen that the energy of the whole string will be dissipated in the quenching magnet. Hence, to prevent burnout, it is necessary to ensure that the normal resistive zone spreads rapidly throughout the quenching coil. This can be done by means of heaters, implemented near the magnet coils and fired as soon as a quench is detected. These heaters are referred to as *quench protection heaters*.

In comparison with other superconducting magnets, most accelerator magnets require an active quench protection system because of the rapidity of the temperature rise resulting from the high current density and the low fraction of stabilizing copper in the cable strands. One notable exception is the RHIC dipole magnets, whose one-layer coil assemblies are wound from a cable with a high copper-to-superconductor ratio (2.25 to 1), and which do not rely on quench protection heaters.

### Hot Spot Temperature

**Estimating Hot Spot Temperature.** The volume of conductor that heats up the most significantly during a quench is the spot where the quench first originated. It is called the *hot spot*. An upper limit of the hot spot temperature,  $T_{max}$ , can be determined by assuming that, near the hot spot, all the power dissipated by the Joule effect is used to heat up the conductor. Integrating the heat balance equation yields

$$S \int_{T_0}^{T_{max}} dT \frac{C(T)}{\rho(T)} = \int_{t_0}^{t_1} dt I(t)^2 \quad (23)$$

where  $C$  is the overall specific heat per unit volume of conductor,  $\rho$  is the overall conductor resistivity in the normal state,  $S$  is the conductor cross-sectional area,  $I$  is the current,  $t_0$  is the time of quench start, and  $T_0$  is the coil temperature at  $t_0$ .

The left member of Eq. (23) depends only on conductor properties whereas the right member depends only on the characteristics of current decay. The right-hand side integral, divided by  $10^6$ , is called the *MIT integral* (Mega  $I$  times  $I$  versus Time integral) and its value is referred to as the *number of MITs*. The maximum temperatures computed from the numbers of MITs have been shown to be in fairly good agreement with actual measurements of hot spot temperatures on quenching magnets (79).

**Limiting Hot Spot Temperature.** The hot spot temperature can be limited by acting on either member of Eq. (23). Regarding the left member, the only conceivable action is to reduce the overall conductor resistivity by increasing the copper-to-superconductor ratio. However, and as explained in the conductor section, the copper-to-superconductor ratio must also be optimized to ensure a high overall critical current. Regarding the right member, the MIT integral can be minimized by (1) detecting the quench as soon as possible, (2) turning off the power supply (case of a single magnet) or forcing the current to bypass the quenching magnet (case of a magnet string), (3) firing the quench protection heaters, and (4) discharging the quenching magnet or the magnet string.

### Quench Detection

The magnets are connected to quench detection systems that monitor the occurrence of a resistive voltage in the coil windings or the coils leads. The resistive voltage must be discriminated from inductive voltages arising from magnet ramping. The inductive components are cancelled out by considering voltage differences across two identical coil assemblies or two identical parts of a given coil assembly (e.g., the upper and lower half coils in a dipole magnet). When the resistive voltage exceeds a preset threshold over a time exceeding a preset duration, the detection system generates a trigger that signals the occurrence of a quench.

### Protection of a Single Magnet

**Current Decay.** Let us first consider the case of a single magnet, and let us assume that, once a quench is detected, the power supply is turned off and the magnet is switched to an external dump resistor,  $R_{ext}$ . The current decay is determined

by

$$L_m \frac{dI}{dt} + [R_q(t) + R_{ext}] I = 0 \quad (24)$$

where  $L_m$  is the magnet inductance and  $R_q(t)$  is the developing resistance in the quenching coils. Furthermore, the total voltage across the magnet  $V_m$  is given by

$$V_m = R_q I(t) \quad (25)$$

where  $I$  is the current intensity.

To limit the number of MITs, it is desirable to have a fast current decay. Equation (24) shows that fast decay rates are obtained either by means of a large  $R_{ext}$  or by ensuring that  $R_q(t)$  increases rapidly. For some magnets, an external resistor can be used to extract a significant fraction of the stored magnetic energy. However, it is also required to keep  $V_m$  to a reasonable level (typically less than 1 kV) to avoid insulation breakdown. Given the order of magnitude of  $I$  (up to 15 kA), this imposes a small  $R_{ext}$  (typically a few hundredths of an ohm), which, during a quench, is soon overcome by  $R_q(t)$ . Hence, for accelerator magnets, the current decay is largely dominated by the resistance development in the quenching coils, and the decay rate can be increased only by speeding up  $R_q(t)$ .

**Maximum Voltage to Ground.** The developing resistance in the quenching coil separates the coil impedance into several parts (Ref. 2, p. 137): unquenched parts across which the voltage is mainly inductive and quenched parts across which the voltage is mainly resistive. The resistive and inductive voltages compensate each other partially so that their sum equals  $V_m$ . The voltage distribution with respect to ground depends on the respective sizes and locations of these various parts. The more uniform the quench development, the lower the maximum voltage to ground. As an illustration, Fig. 13 shows the voltage distribution in a quenching magnet. Here,  $V_m$  is assumed to be nil, and  $R_q$  is assumed to be concentrated near two-thirds of the magnet length.

**Quench Protection Heaters.** As described earlier, to speed up and uniformize quench development, most accelerator magnets rely on quench protection heaters, which are fired as soon as a quench is detected. The heaters are usually made of stainless steel strips, which are copper clad at regular intervals along their lengths and which are placed on the outer surface of the coil assemblies. Note, however, that the heater firing unit relies on a capacitor bank and that it takes some time for the energy to be released. Note also that the heaters must be electrically insulated from the coil and that this electrical insulation introduces a thermal barrier. As a result, there is a nonnegligible delay between the firing of the heaters and their effect on the coils, during which we must rely on natural quench propagation (80). The heaters and their implementations in the magnet assembly are optimized to reduce this delay.

### Protection of a Magnet String

In an accelerator, the magnet ring is divided into several sectors made up of series-connected magnets. The sectors are powered independently and are electrically independent. Once a quench is detected in a magnet, the power supply of the sector to which

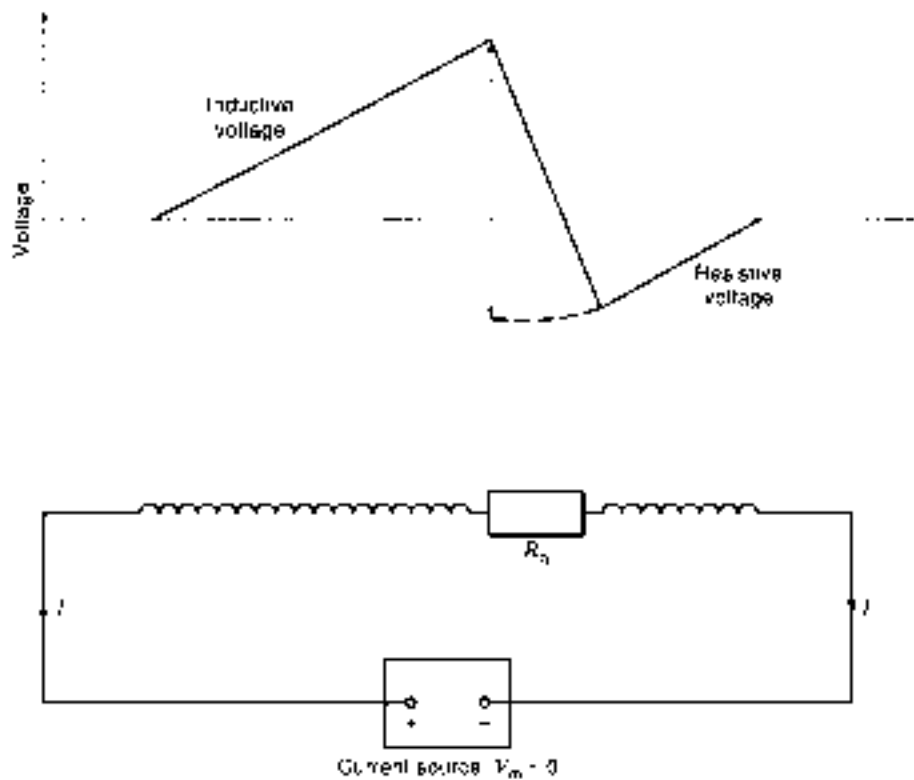


Figure 13. Voltage distribution in a quenching magnet. The total voltage across the magnet is assumed to be nil and the developing resistance is assumed to be concentrated near two-thirds of the magnet length (2).

the magnet belongs is turned off and the sector is discharged over a dump resistor.

Unlike in the case of a single magnet, the current decay rate in the sector must be limited for at least two reasons: (1) to prevent the induction of large coupling currents in the magnet coils (which may quench the remaining magnets in the sector, resulting in general warming and significant helium venting); and (2) to avoid the occurrence of unacceptable voltages to ground (because of the large overall inductance of the sector). A too slow decay rate, however, creates the risk that a significant fraction of the total energy stored in the sector be dissipated in the quenching magnet, resulting in destructive overheating.

These contradictory considerations can be reconciled by forcing the current to bypass the quenching magnet and by ramping the current down at the desired rate in the remaining unquenched magnets. The bypass elements consist of diodes (or thyristors) connected in parallel to individual or small groups

of magnets, as shown in Fig. 14. As long as the magnets are superconducting, the current flows through the magnets. Once a magnet has quenched and starts to develop a resistive voltage, the main current is bypassed through the diode connected in parallel, and the quenching magnet is discharged over the diode circuit. The current decay is determined by an equation similar to Eq. (24), except that  $R_0$  must be replaced by the resistance associated with the bypass element  $R_b$ .

HERA, RHIC, and LHC rely on silicon diodes that are mounted inside the helium cryostats and operate at cryogenic temperatures. The main requirements for these cold diodes are (81): (1) small forward voltage and low dynamic resistance (to limit power dissipation in the diodes); (2) good radiation hardness, and (3) large backward voltage. In the case of the Tevatron, which has a short current ramp time resulting in large inductive voltages across the bypass elements, the diodes are replaced by thyristors operating as fast switches (82). The

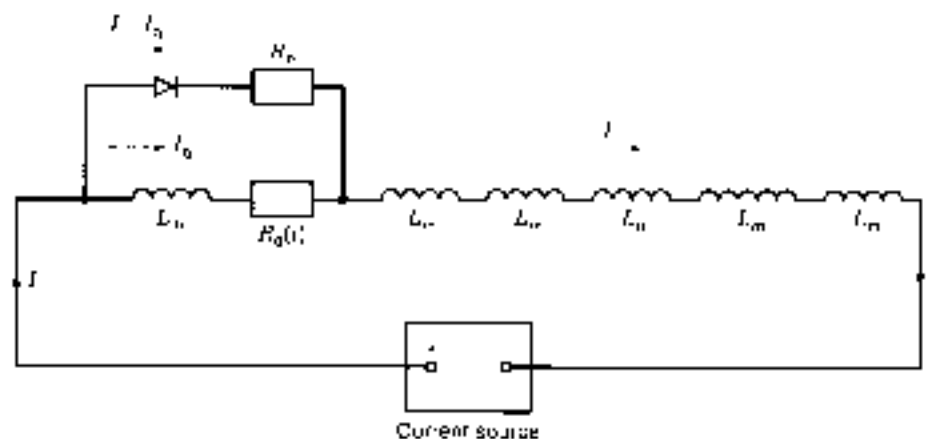


Figure 14. Electrical circuit of a quenching magnet in a magnet string (2).

thyristors are located outside the magnet cryostats and require additional power leads and cryogenic feedthroughs.

The protection system of the magnet ring must be carefully designed and thoroughly tested before starting up the machine. The system tests are usually carried out on a cell or a half-cell representative of the magnet lattice, and all failure modes are investigated (83–85).

## SUMMARY

As of today, two large superconducting accelerator rings, Tevatron and HERA, have been built and are reliably operating, and work is under way on two other superconducting colliders—RHIC and LHC. The construction of RHIC is near completion, and the industrial contracts for the mass production of LHC magnets will be awarded in 1999.

Since the time of the Tevatron (late 1970s), a factor of about two has been gained on the critical current density of NbTi at 4.2 K and 5 T, and a dipole field of 10.5 T has been reached on a short magnet model relying on NbTi cables at 1.8 K. In recent years, encouraging results have been obtained on a couple of short dipole magnet models relying on Nb<sub>3</sub>Sn cables, which may open the range 10 T to 15 T.

## BIBLIOGRAPHY

1. D. A. Edwards and M. J. Squires, *An Introduction to the Physics of High Energy Particle Accelerators*. New York: Wiley, 1993.
2. K. H. Mess, F. Schmuser, and S. Wolf, *Superconducting Accelerator Magnets*. Singapore: World Scientific, 1993.
3. H. T. Edwards, The Tevatron energy doubler: A superconducting accelerator. *Annu. Rev. Nucl. Part. Sci.*, **35**: 605–660, 1985.
4. A. V. Tallstra, Superconducting magnets. *AIP Conf. Proc.*, **87**: 1982, pp. 889–904.
5. B. H. Wik, The status of HERA. *Conf. Rec. 1991 IEEE Part. Accel. Conf.*, IEEE Catalogue 91CH3333-7, 1991, pp. 2805–2809.
6. R. Meinke, Superconducting magnet system for HERA. *IEEE Trans. Magn.*, **27**: 1728–1734, 1991.
7. V. I. Balbekov and G. G. Ginz, HEP accelerating and storage complex: Status and possibility of D-factory. *Nucl. Instrum. Meth. Phys. Res., A*, **333**: 198–195, 1993.
8. A. V. Zaban, CNK superconducting magnets development. *Nucl. Instrum. Meth. Phys. Res., A*, **333**: 196–203, 1993.
9. R. F. Schwitters, Future hadron collider: The SSC. *AIP Conf. Proc.*, **272**: 1893, pp. 306–323.
10. R. I. Scherner, Status of superconducting magnets for the superconducting super collider. *IEEE Trans. Magn.*, **30**: 1567–1594, 1994.
11. H. Fetscher et al., The relativistic heavy ion collider, RHIC, in H. Schopper (ed.), *Advances of Accelerator Physics and Technology*, vol. 12, Adv. Ser. Directions High Energy Phys., Singapore: World Scientific, 1993, pp. 104–131.
12. A. Greene et al., The magnet system of the relativistic heavy ion collider (RHIC). *IEEE Trans. Magn.*, **32**: 2043–2046, 1996.
13. L. B. Evans, The large hadron collider project, in T. Haruyama, T. Mitsui, and K. Yamafuji (eds.), *Proc. 16th Int. Cryogenic Eng. Conf. Int. Cryogenic Mater. Conf.*, London: Elsevier, 1997, pp. 45–52.
14. R. Perin, Status of LHC programme and magnet development. *IEEE Trans. Appl. Supercond.*, **6**: 189–195, 1996.
15. M. N. Wilson, Superconducting magnets for accelerators: A review. *IEEE Trans. Appl. Supercond.*, **7**: 727–732, 1997.
16. A. den Ouden et al., Application of Nb<sub>3</sub>Sn superconductors to high-field accelerator magnets. *IEEE Trans. Appl. Supercond.*, **7**: 733–738, 1997.
17. R. M. Scanlan et al., Preliminary test results of a 13-tesla niobium tri-dipole. *Inst. Phys. Conf. Ser.*, **168**: 1503–1506, 1997.
18. E. Gregory, Conventional wire and cable technology. *AIP Conf. Proc.*, **249**: part 2, 1992, pp. 1196–1229.
19. L. Bottura, private communication, 1997.
20. D. C. Larbaletier and P. J. Lee, New development in niobium-titanium superconductors. *Proc. 1995 IEEE Part. Accel. Conf.*, vol. 2, 1995, pp. 1279–1281.
21. D. Leroy et al., Design features and performance of a 10 T twin-aperture model dipole for LHC. *Proc. 16th Int. Conf. Magn. Technol. MT-15*, Beijing, China, 1997, submitted.
22. L. T. Summers et al., A model for the prediction of Nb<sub>3</sub>Sn critical current as a function of field, temperature, strain and radiation damage. *IEEE Trans. Magn.*, **27**: 3041–3044, 1991.
23. R. Ayari, Overview of the ITER project, in T. Haruyama, T. Mitsui, and K. Yamafuji (eds.), *Proc. 16th Int. Cryogenic Eng. Conf. Int. Cryogenic Mater. Conf.*, London: Elsevier, 1997, pp. 53–59.
24. P. Hruzovic et al., Testing of industrial Nb<sub>3</sub>Sn strands for high field fusion magnets. *IEEE Trans. Magn.*, **30**: 1986–1989, 1994.
25. D. C. Larbaletier, The road to conductors of high temperature superconductors—10 years do make a difference!. *IEEE Trans. Appl. Supercond.*, **7**: 93–97, 1997.
26. J. M. Rivet and R. M. Scanlan, Manufacture of key-stored flat superconducting cables for use in SSC dipoles. *IEEE Trans. Magn.*, **23**: 440–483, 1987.
27. M. N. Wilson and R. Wolf, Calculation of minimum quench energies in Rutherford-type cables. *IEEE Trans. Appl. Supercond.*, **7**: 950–953, 1997.
28. A. K. Ghosh, W. B. Sampson, and M. N. Wilson, Minimum quench energies of Rutherford cables and single wires. *IEEE Trans. Appl. Supercond.*, **7**: 954–957, 1997.
29. T. Skintans et al., Development of large keystone angle cable for dipole magnet with ideal arch structure. *Adv. Cryog. Eng. (Mater.)*, **36**: Pt. 3: 323–329, 1990.
30. R. M. Scanlan and M. Rayer, Recent improvements in superconducting cable for accelerator-dipole magnets. *Conf. Rec. 1991 IEEE Part. Accel. Conf.*, 1991, pp. 2155–2157.
31. D. Richter et al., DC measurements of electrical contacts between strands in superconducting cables for the LHC magnets. *IEEE Trans. Appl. Supercond.*, **7**: 796–792, 1997.
32. J. M. Depond et al., Examination of contacts between strands by electrical measurement and topographical analysis. *IEEE Trans. Appl. Supercond.*, **7**: 793–796, 1997.
33. J. D. Adam et al., Rutherford cables with anisotropic transverse resistance. *IEEE Trans. Appl. Supercond.*, **7**: 954–956, 1997.
34. P. A. Houghton and K. D. Williamson, Jr., Electrical breakdown and tracking characteristics of pulsed high voltages in cryogenic helium and nitrogen. *Adv. Cryog. Eng.*, **21**: 107–108, 1975.
35. P. Vedrine et al., Mechanical tests on the prototype LHC lattice quadrupole. *IEEE Trans. Magn.*, **30**: 2475–2479, 1994.
36. M. Anerella et al., Improved cable insulation for superconducting magnets. *Proc. 1993 Part. Accel. Conf.*, 1993, pp. 2790–2792.
37. G. H. Morgan, Shaping of magnetic fields in beam transport magnets. *AIP Conf. Proc.*, **249**: part 2, 1992, pp. 1242–1261.
38. R. C. Gupta, S. A. Kahn, and G. H. Morgan, SSC 50 mm dipole cross section. in *J. Nucl. Sci., Superconduct.*, New York: Plenum, 1991, pp. 537–539.
39. A. Devred and T. Ogitsu, Ramp-rate sensitivity of SSC dipole



magnet prototypes, in S. I. Kurokawa, M. Month, and S. Turner (eds.), *Frontiers of Accelerator Technology*, Singapore: World Scientific, 1996, pp. 194-304.

43. H. H. J. ten Kate et al., The reduction of the critical current in Nb<sub>3</sub>Sn cables under transverse forces, *IEEE Trans Magn.*, **28**: 714-718, 1992.

44. K. C. Gupta, A common coil design for high field 2-in-1 accelerator magnets, *Conf. Rec. 1997 IEEE Part. Accel. Conf.*, in press.

45. J. M. Cook, Strain energy minimization in SSC magnet winding, *IEEE Trans Magn.*, **27**: 1976-1980, 1991.

46. J. S. Brandt et al., Coil end design for the SSC collider dipole magnet, *Conf. Rec. 1991 IEEE Part. Accel. Conf.*, 1991, pp. 2182-2184.

47. T. Garavaglia et al., Application of the SSC/BTK numerical simulation program to the evaluation of the SSC magnet aperture, in M. Mankashin (ed.), *Supercollider 2*, New York: Plenum, 1990, pp. 59-76.

48. T. Ogitsu and A. Deved, Influence of azimuthal coil size variations on magnetic field harmonics of superconducting particle accelerator magnets, *Rev. Sci. Instrum.*, **65**(10): 1998-2005, 1994.

49. U. P. Bean, Magnetization of high-field superconductors, *Rev. Mod. Phys.*, **36**(1): 31-39, 1964.

50. H. Brück et al., Field distortions from persistent magnetization currents in the superconducting HERA magnets, *Z. Phys. C - Part. Fields*, **44**: 395-399, 1989.

51. R. W. Easift et al., Studies of time dependent field distortions from magnetization currents in Tevatron superconducting dipole magnets, *IEEE Trans Magn.*, **25**: 1647-1651, 1989.

52. A. Dwyer et al., Time decay measurements of the sextupole component of the magnetic field in a 4-cm aperture, 17-m long SSC dipole magnet prototype, *Conf. Rec. 1991 IEEE Part. Accel. Conf.*, 1991, pp. 2450-2462.

53. P. Witteke and F. Zimmermann, The impact of persistent current field errors on the stability of the proton beam in the HERA proton ring, *Conf. Rec. 1991 IEEE Part. Accel. Conf.*, 1991, pp. 2483-2487.

54. L. Bottura, L. Walckiers, and R. Wolf, Field errors decay and "snap-back" in LHC model dipoles, *IEEE Trans Appl. Supercond.*, **7**: 602-605, 1997.

55. P. W. Anderson, Flux creep in hard superconductors, *Phys. Rev. Lett.*, **9**(7): 309-311, 1962.

56. H. Brück et al., Observation of a periodic pattern in the persistent-current fields of the superconducting HERA magnets, *Conf. Rec. 1991 IEEE Part. Accel. Conf.*, 1991, pp. 2149-2151.

57. A. K. Chosh et al., Axial variations in the magnetic field of superconducting dipoles and quadrupoles, *Proc. 1993 Part. Accel. Conf.*, 1993, pp. 2742-2743.

58. A. Akhmetov et al., Current loop decay in Rutherford-type cables, in P. H. Deder, (ed.), *Supercollider 5*, New York: Plenum, 1994, pp. 443-446.

59. L. Kremparski and C. Schmidt, Influence of a longitudinal variation of  $\alpha$  on the magnetic field distribution of superconducting accelerator magnets, *Appl. Phys. Lett.*, **66**(12): 1545-1547, 1995.

60. A. P. Verweij, Boundary-induced coupling currents in a 1.7 m Rutherford type cable due to a locally applied field change, *IEEE Trans Appl. Supercond.*, **7**: 270-273, 1997.

61. L. Hecura, L. Walckiers, and Z. Ang, Experimental evidence of boundary induced coupling currents in LHC prototypes, *IEEE Trans Appl. Supercond.*, **7**: 801-804, 1997.

62. K. Stiening, private communication 1997.

63. A. Dwyer et al., About the mechanics of SSC dipole magnet prototypes, *Adv. Conf. Proc.*, **240**(2): 1992, pp. 1769-1774.

64. T. Ogitsu, K. Tsuchiya, and A. Deved, Investigation of wire motion in superconducting magnets, *IEEE Trans Magn.*, **27**: 2132-2135, 1991.

65. K. Koepke et al., Ferrilab doubler magnet design and fabrication techniques, *IEEE Trans Magn.*, **MAG-15**: 658-661, 1979.

66. D. Perrin et al., Measurements of the resistance to stress cycling at 4.2 K of LHC dipole collars, *IEEE Trans Magn.*, **30**: 1754-1757, 1994.

67. A. Deved et al., Review of SSC dipole magnet mechanics and quench performance, in J. Nonte (ed.), *Supercollider 4*, New York: Plenum, 1992, pp. 113-136.

68. D. Leroy et al., Structural analysis of the LHC 10 T twin-aperture dipole, in T. Sekiguchi and S. Shimamoto (eds.), *Proc. 11th Int. Conf. Magn. Technol.*, London: Elsevier Applied Science, 1990, pp. 159-164.

69. E. Acerro et al., Development and fabrication of the first 10 m long superconducting dipole prototype for the LHC, *IEEE Trans Magn.*, **30**: 1791-1796, 1994.

70. P. Wanders et al., Construction and testing of arc dipoles and quadrupoles for the relativistic heavy ion collider (RHIC) at BNL, *Proc. 1995 Part. Accel. Conf.*, 1995, pp. 1291-1297.

71. J. M. Baze et al., Design and fabrication of the prototype superconducting quadrupole for the CERN LHC project, *IEEE Trans Magn.*, **28**: 335-337, 1992.

72. J. Strait et al., Tests of full scale SSC R&D dipole magnets, *IEEE Trans Magn.*, **25**: 1456-1458, 1989.

73. S. W. Van Siver, *Helium Cryogenics*, New York: Plenum, 1986.

74. P. Lebrun, Cryogenic systems for accelerators, in S. I. Kurokawa, M. Month, and S. Turner (eds.), *Frontiers of Accelerator Technology*, Singapore: World Scientific, 1996, pp. 691-700.

75. P. Lebrun, Superfluid helium cryogenics for the large hadron collider project at CERN, *Cryogenics, ICCC Suppl.*, **34**: 1-9, 1994.

76. G. Claudet and P. Aymar, Two-stage and He II cooling of superconducting magnets, *Adv. Cryog. Eng.*, **35**(A): 55-67, 1990.

77. W. Frost (ed.), *Heat Transfer at Low Temperatures*, New York: Plenum, 1975.

78. T. H. Neal, Cryostat design for the superconducting super collider, *ATP Conf. Proc.*, **249**: part 2, 1992, pp. 1230-1241.

79. M. N. Wilson, *Superconducting Magnets*, Oxford: Clarendon, 1983.

80. P. Wanders et al., Partial lifetime test of an SSC collider dipole, *IEEE Trans Magn.*, **30**: 1738-1741, 1994.

81. L. Schneider and P. Turawski, Critical current degradation of a NbTi multifilament conductor due to heat treatment, *IEEE Trans Magn.*, **30**: 2391-2394, 1994.

82. A. Deved et al., Investigation of heater-induced quenches in a full-length SSC R&D dipole, in T. Sekiguchi and S. Shimamoto (eds.), *Proc. 11th Int. Conf. Magn. Technol.*, London: Elsevier, 1990, pp. 91-95.

83. A. Deved et al., Quench characteristics of full-length SSC R&D dipole magnets, *Adv. Cryog. Eng.*, **35**(A): 599-608, 1990.

84. L. Cault et al., LHC magnet-quench protection system, *IEEE Trans Magn.*, **30**: 1742-1745, 1994.

85. B. Stiening et al., A superconducting synchrotron power supply and quench protection scheme, *IEEE Trans Magn.*, **MAG-15**: 670-672, 1979.

86. K. Koepke et al., Two-bell doubler system test, *IEEE Trans Magn.*, **MAG-17**: 713-715, 1991.

87. W. Burgett et al., Full power test of a string of magnets comprising a half-cell of the superconducting super collider, *Part. Accel.*, **43**: 41-75, 1993.

88. E. Rodriguez Mateos et al., Electrical performance of a string of magnets representing a half-cell of the LHC machine, *IEEE Trans Magn.*, **32**: 2105-2108, 1996.

## SUPERCONDUCTING TRANSFORMERS

In an electrical power system, power is generated far away from the consuming areas. The electric power is transmitted from generating locations to consuming locations through transmission lines. A high voltage is desirable for transmitting large amounts of power in order to minimize the current and the associated  $I^2R$  losses, and reduce the amount of conductor used in transmission lines. A much lower voltage, on the other hand, is required for distribution, for various reasons connected with safety and convenience. The transformer makes the reduction in voltage easy and economically possible. Generally, electricity is transformed three or four times between the location of generation and the location of consumption making transformers one of the basic elements of an electric power system.

The physical basis of the transformer is mutual inductance between two circuits linked by a common magnetic field, as shown in Fig. 1. The power transformer transfers electrical energy from one circuit to another, via the medium of the pulsating mutual magnetic field. Magnetic iron enhances the flux linkage between the circuits. The transformer units are therefore made to embrace an iron core, which serves as a conduit for the mutual magnetic flux, ensuring that the flux links each coil fairly completely. The use of an iron core permits greater freedom in shape and relative position of the primary and secondary coils (Fig. 2), since the majority of the mutual flux is conveyed by the core regardless of the relative positions of the two sets of coils—primary and secondary.

Since transformers are employed extensively in a power system, their efficiency and losses are considered a serious issue. The transformer design selection is normally made on the basis of its lifetime cost which consists of the initial cost plus the cost of operating it over its lifetime. The lifetime cost of even a small loss could be significant. In addition, most larger transformers employ oil for cooling the windings and the iron core. Although the oil is an excellent cooling medium and is a good high voltage insulator, it has attracted the ire of environmentalists and fire departments due to the possibility of oil spills and fire hazards. Moreover, larger units are too heavy and bulky for normal transportation channels. Space in an urban environment is quite valuable. A more compact, light-weight transformer could more easily be sited, possibly even above the ground floor or basement levels. This could be a significant advantage. Some utilities have indicated that they might be willing to pay a premium for such an advantage. These difficulties have inspired

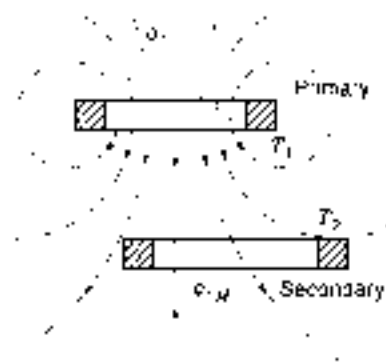


Figure 1. Flux linkage between primary and secondary in air.

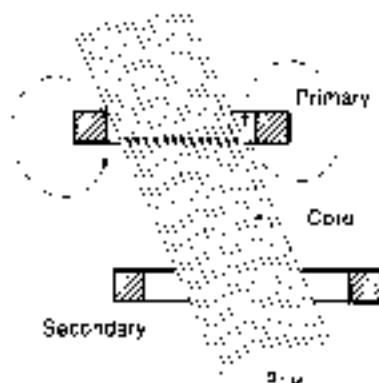


Figure 2. Flux linkage between primary and secondary with iron core.

designers and users to look for alternative transformer solutions. Superconducting transformers appear to offer a solution to most of these problems.

A superconductor only operates within a space bounded by three parameters: current density in the superconductor, magnetic field experienced by the superconductor, and its operating temperature. The maximum operating current of a given superconducting wire is a function of these parameters. If any of these parameters is violated, the superconducting wire loses its superconducting property and becomes resistive. Once in the resistive state, it generates joule heating. This heating must be limited to a safe value in order to prevent permanent damage to the windings. The circuit breaker feeding the transformer could be used to disconnect it if the winding temperature exceeds a given upper limit. Since superconductors can sustain large current densities with potentially low losses, a superconducting transformer is expected to be smaller, lighter, and more efficient. In addition, since the superconducting transformer uses cryogenic liquids as dielectric and coolant, it is also free of environmentally unacceptable oil. During the eighties, several groups designed, built, and tested small transformers employing low temperature superconductors (LTS) Niobium-Titanium (NbTi) superconducting windings (1, 6), cooled with liquid helium to around 4 K. Several problems were observed:

- First, the superconductor must operate in the presence of fluctuating ac currents in a moderately high magnetic field resulting in decreased stability and increased ac losses.
- Second, under system fault conditions, if the critical current of the conductor is exceeded, recovery to the superconducting state is too slow to allow automatic circuit reclosure.
- Third, the need to cool with liquid helium reduces economic benefits of such a device.

The discovery of high temperature superconductors (HTS) has revived interest in superconducting transformers. Presently, attempts are being made to design, build, and test transformers with HTS windings cooled with liquid nitrogen at 77 K. These HTS transformers are less likely to have the problems associated with NbTi transformers. For example, the use of HTS material greatly improves stability, the transformer could be designed to prevent quenching (i.e., transitioning to

the normal state, and the consequences of ac losses could be potentially overcome by developing low-ac-loss HTS conductors and operating it at 77 K. References 7–9 have summarized the evolution of the superconducting transformers. Recently two HTS prototype transformers have been tested—a 630 kVA, three-phase transformer by ABB (10) and a 500 kVA, one-phase transformer by Kyushu University (11). HTS transformers are also being considered for railroad rolling stock in Japan (11A).

The HTS transformer could also be designed to provide fault current limiting functionality and to improve system stability (11B). This article discusses potential system benefits associated with superconducting transformers, and reviews and summarizes the design requirements for such a device. It also compares and contrasts the requirements for transformers utilizing LTS and HTS conductors.

## SUPERCONDUCTING TRANSFORMER CONFIGURATION

A superconducting transformer operates using the same principles and constituent parts as found in a conventional transformer. Both employ an iron core to contain magnetic flux, and primary and secondary windings to exchange power. However, the construction of the two types of transformers is quite different. For example, in conventional transformers primary and secondary coils are directly wound on the iron core and both iron core and coil assemblies are immersed in a tank filled with oil that cools both the iron core and the coils. On the other hand, in superconducting transformers, the iron core is usually maintained at room-temperature while the superconducting coils operate at cryogenic temperatures. The decision to maintain the iron core at room temperature is dictated by the fact that the iron core losses, due to hysteresis and eddy-currents, are substantial (almost 1 watt per pound) and they go up when the iron core is operated at cryogenic temperatures (12). The iron core permeability also goes down at low temperature which means more iron core is required to carry the same flux at low temperature than at room-temperature. The core losses would represent a major load on the refrigerator if the iron core were operated at cryogenic temperature. On the other hand, the superconducting windings must be cooled to cryogenic temperatures (between 4.5 K and 77 K) which necessitates that these windings be enclosed in containers which could hold vacuum or cryogen or both. These containers surround the iron core limbs and take the shape of hollow toroids. Since they surround the iron core, they must be constructed from nonmetallic material lest they form a closed circuit around the iron core and thus form a shorted secondary for the transformer, making the transformer unworkable. It is possible to employ metallic containers but a dielectric break must be included in the circumferential direction to prevent flow of current in the container walls. The dielectric break makes these metallic cryogen containers more expensive and less reliable. The superconducting windings must be cooled with suitable cryogen (liquid helium for NbTi windings, liquid nitrogen for HTS BiPbSrCaCuO-2223 (BSCCO-2223) windings, or an intermediate temperature for Nb<sub>3</sub>Sn or BSCCO-2212 windings). Reference 13 describes the status of the HTS conductor technology and (14) discusses coils made from BSCCO-2212 material. Since these containers hold windings at low temperatures, their walls must be thermally insulating. For low temperature operation at around

4 K, a double wall construction is employed with multi-layer-insulation (MLI) insulation in the vacuum space between the walls. Additionally, an intermediate temperature (77 K) shield is also inserted between warm and cold walls of the cryostat. This makes the container design complex and expensive. On the other hand, if the windings operate at 77 K then single wall construction could be used for these containers. This makes design and construction simpler and the cost is substantially lower as compared to those of the low temperature coils. The cost of refrigerator (both capital and lifetime) is also much lower for devices operating at 20 K to 77 K than those operating at 4 K.

## TRANSFORMER DESIGN AND ANALYSIS

The transformer design is obtained with an optimization process which involves varying several significant parameters which are interrelated in complex ways. A transformer is sized on the basis of its power rating, voltage, number of phases, frequency, and short-circuit reactance. It is also necessary to pay close attention to transformer type, service conditions, cost of losses, and the relative costs of conductor, iron, insulation, labor, machinery, and configuration. All these factors are considered when designing a superconducting as well as normal transformer.

### Design Issues

It is possible in principle to construct a superconducting transformer without an iron core. Such transformers are characterized by a larger reduction in losses, size, and weight than those employing iron core but they require much larger excitation current (15–16). On the other hand, an iron core offers the following benefits:

- The core contains the mutual flux between the windings, and thus, reduces stray field which has adverse impacts on people and other equipment in the vicinity of the transformer.
- The magnetic field experienced by the superconducting windings is reduced, thus reducing the amount of superconductor required.
- The ac losses in the windings are reduced, which reduces the size of the refrigeration system.

Since losses in the iron core are large, it is normally preferable to keep the iron core at room-temperature. This requirement to operate the iron core at room-temperature forces superconducting transformer configurations which differ from those of conventional transformers.

A generic three-phase superconducting transformer is shown in Fig. 1. It has a set of concentric primary and secondary windings surrounding each leg of a three-leg transformer core. Since the windings operate at cryogenic temperatures, primary-secondary winding pairs are enclosed in individual cold containers (identified as cold wall of cryostat in Fig. 3). This cold wall must be made of nonconductive materials in order to prevent a shorted turn. A common warm wall encloses all three cold walls—this wall could be metallic.

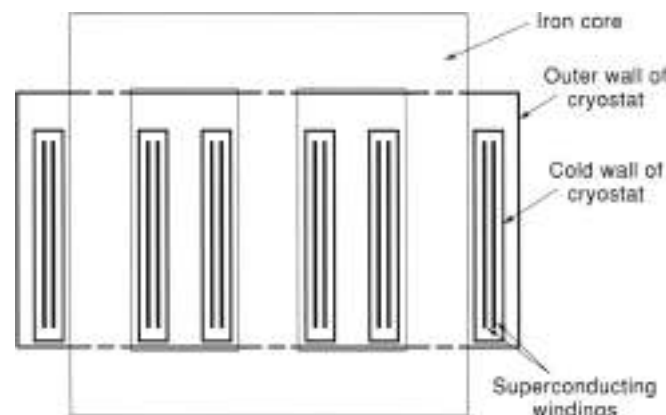


Figure 3. HTS transformer configuration

### Conductor Concepts

Superconductors being considered for transformers could be divided into two broad categories on the basis of their operating temperatures. LTS operate at liquid helium temperature ( $\sim 4.2$  K) and HTS operate at temperatures ranging from 20 K to 77 K (the temperature of liquid nitrogen). Two LTS conductor options are available: Niobium-Titanium (NbTi) and Niobium-Tin (Nb<sub>3</sub>Sn). Likewise, two HTS conductor choices are available in long lengths, BSCCO-2212 for operation at  $\sim 20$  K to 30 K, and BSCCO-2223 for operation  $\sim 50$  K to 77 K. This section discusses characteristics of these conductors and their pros and cons for application in a superconducting transformer.

**Niobium-Titanium Conductors.** Superconducting NbTi wires for 50 Hz to 60 Hz applications require very small diameters in order to minimize ac losses and improve intrinsic stability (17–19). These NbTi wires are characterized by very fine filaments (0.1  $\mu\text{m}$  or less), high resistance CuNi matrix, and small wire diameter ( $\sim 0.2$  mm). A practical size conductor, capable of carrying hundreds of amperes, consists of many such wires.

Primary and secondary coils made with this conductor are generally housed in a common container filled with liquid helium. Since it is expensive to remove heat generated at liquid helium temperature (4 K) with a refrigerator, designs are always optimized to minimize heat generated at low temperatures. Since the specific heat of metals is very low at the liquid helium temperatures, a small heat input forces the conductor into its normal conducting state. Protection against the consequences of an unexpected quench is one of the most significant challenges for superconducting coils—the intensive and localized Joule heating can produce catastrophic damage (20). Moreover, the protection of ac windings is more critical than of dc windings, because of high matrix resistivity—it is not permissible for wire to carry its nominal current longer than a few milliseconds, otherwise permanent damage could occur. Thus, transformers employing NbTi windings must be designed carefully to avoid these problems. Similar design approaches must be used for Nb<sub>3</sub>Sn conductors. These conductors have higher critical temperature than NbTi, and it is therefore possible to operate them at higher temperatures ( $\sim 10$  K) as compared to 4 K for NbTi. Nevertheless, most problems associated with low temperature persist. The lack of economic feasibility and high cost of refrigeration stopped LTS transformer activities.

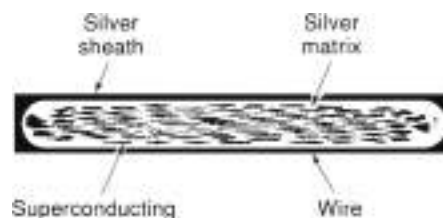


Figure 4. Highly textured HTS conductor

**BSCCO Conductors.** Although BSCCO conductors have a potential of operating at  $\sim 77$  K and being cooled with environmentally friendly liquid nitrogen, no suitable design of an ac conductor exists at this time. The major problem is posed by high aspect ratios (1:10) of such conductors as shown in Fig. 4. Since hysteresis losses are directly proportional to filament diameter (17), losses induced by magnetic field perpendicular to the wide surface of the conductor are high and they pose coil cooling and overall transformer efficiency challenges. Furthermore, since a single strand of BSCCO wire could only carry limited current, it is necessary to wind several strands in parallel. If these strands are not transposed they can cause significant coupling losses. Iwakuma (21) has proposed a transposition scheme to minimize these coupling losses. Losses due to field parallel to the wide face of the conductor are likely to become acceptable.

The perpendicular field losses only occur in the end regions (10% of the coil axial length) of coils and with clever winding schemes, it might be possible to minimize these losses. Research is continuing at AMSC and other places around the world but no one has yet published a credible conductor design. Until a credible HTS conductor is developed for ac application, the HTS transformer will remain elusive.

Since the specific heat of metals is high at higher temperatures, it is possible to absorb larger amount of heat at the higher operating temperatures of HTS for a modest temperature rise of the conductor. Nevertheless, ac losses in presently available conductors are still unacceptable due to cost/efficiency considerations. ABB has made a 630 kVA transformer (10,22) which employed BSCCO 2223 wire from AMSC but its ac losses are unacceptably high for commercial deployment. On the other hand, this device has demonstrated that if a suitable HTS wire was available then it would be possible to construct a practical transformer. Figure 5 (taken from Ref. 22) shows a Model coil made by ABB, from wire provided by American Superconductor Corporation (AMSC), for the 630-kVA transformer project in Geneva, Switzerland.

Coated conductors employing YBCO films on a substrate are emerging as an alternative to BSCCO conductors. High critical current in HTS films deposited on crystallographically oriented substrates has been demonstrated recently by Los Alamos National Laboratory (LANL) (23) and independently by Oak Ridge National Laboratory (ORNL). Joint industry-laboratory programs are underway to scale up this coated conductor technology. A number of technological problems must be solved before a practical conductor could emerge. Production conductors are expected by a 2004 time frame.

The coated conductor promises significantly higher performance than the BSCCO, with projected cost/performance below the much discussed \$10/kA-m commercialization



Figure 5. Model coil made by ABB, from wire provided by American Superconductor Corporation (AMSC), for the 630 kVA transformer project in Geneva, Switzerland.

benchmark. Overall strand current densities of up to 50,000 A/cm<sup>2</sup> are expected to be achievable. In addition, the coated conductor films may ultimately be engineered to optimize filament dimensions and to eliminate filament coupling through the careful selection of layers within the conductor architecture.

#### AC Losses

Losses in a superconducting coil are quite small under dc operation. However, these losses become significant if the coil current is ramped rapidly or if it carries ac as in a transformer. The ac losses are generally quite small, but the refrigeration penalty amplifies their effect. When the applied field is low, superconductor tends to screen penetration of field into interior of the conductor. However, higher fields fully penetrate a conductor. Under such a condition, hysteresis loss ( $W_H$ ) is given by the following formula (17,24) for decoupled filaments:

$$Q_H = 3H_c d_c J_c$$

where

- $\Delta H$  = field variation (peak-to-peak) (T)
- $J_c$  = critical current density of superconducting filament (A/cm<sup>2</sup>)
- $r$  = radius of filament (m)
- $f$  = frequency (Hz)

The largest loss component is usually hysteresis loss in the superconductor filaments. One way to reduce these losses is to make the superconducting filament diameter ( $2r$ ) as small as possible.

Another component of ac losses is due to coupling between filaments. This coupling takes place when the electric field between adjacent filaments is sufficiently large to cause a current flow between filaments through the matrix. Filaments are tightly twisted in a helical fashion and are surrounded by high resistivity matrix in order to reduce these coupling currents and the associated losses. The coupling losses are given by the

following equations taken from (25):

$$P_c = \frac{1}{144} \left( \frac{c}{d} \right)^2 \left( \frac{t_c^2}{\rho_c} \cdot B^2 \right) \quad (1)$$

$$P_c = \frac{1}{16} \cdot \left( \frac{r}{d} \right)^2 \cdot \left( \frac{t_c^2}{\rho_c} \cdot B^2 \right) \quad (2)$$

where

- $P_c$  = loss per unit volume of the conductor (W/cm<sup>3</sup>)
- $c$  and  $d$  = conductor cross-sectional dimensions (width and thickness) (m)
- $t_c$  = twist pitch length (m)
- $B$  = rate of change of magnetic field (T/s)
- $\rho_c$  = matrix resistivity ( $\mu\Omega\cdot\text{m}$ )

Equation (1) is used when the field is parallel to the wider face Eq. (2) of the conductor ( $d \ll r$ ). Eq. (2) is for the case when  $r \gg d$ .

It has been shown (26,27) that NbTi coils operating at liquid helium temperature must have a filament diameter in the order of 0.1  $\mu\text{m}$  or less to make these coupling losses comparable to the losses in copper windings of a conventional transformer.

Both of these ac loss phenomena also apply to HTS conductors. HTS conductors are currently made in highly aspected tape shapes (1:10) as shown in Fig. 4. The hysteresis losses due to magnetic field parallel to the broad face of the conductor are acceptable since the dimension transverse to the field is small, but losses due to magnetic field perpendicular to the broad face of the conductor are excessively large. Although several industrial groups are attacking this problem, no method has yet been published for making a conductor capable of carrying large currents (comparable to those of LTS) while keeping hysteresis and coupling losses low. HTS conductors, however, have a couple of significant advantages over the LTS conductors: the temperature rise due to a transient heat input (i.e., by conductor movement) is lower at the higher HTS operating temperatures than when operating at liquid helium temperature, and they have a slow transition from superconducting to normal state which makes them inherently more stable. This advantage may translate into relaxed requirements on the acceptable filament size and other conductor configuration parameters. However, the ac losses in these wires must be low in order to make HTS transformers economically acceptable.

Another significant loss component is the heat conduction through current leads. One end of a current lead is at room temperature and the other end is at low temperature, and the heat is conducted along the length of the lead from warm to cold regions. In case of LTS transformers, conduction heat can be intercepted at the intermediate thermal shield, which is usually kept at  $\sim 70$  K. The heat conduction between the 70 K and 4 K winding regions could be controlled by employing HTS current leads. However, in case of HTS transformers operating at the liquid nitrogen temperature (77 K), all of the lead conduction must be cooled by the liquid nitrogen coolant. A typical 100 A pair of leads conducts 8 W of heat load into the cold region.

### Cryostat

To maintain the low temperature environment essential for operation of LTS and HTS magnets, they must be placed in special vessels or cryostats. These are vacuum insulated containers. Designs for LTS and HTS could be widely different: HTS cryostats are likely to be easier to design and fabricate than LTS cryostats. Since the cost of removing losses from low temperatures (4.2 K) is very high, usually a double or triple wall construction is employed. The innermost space contains the liquid helium and outermost wall operates at room temperature. An intermediate wall is normally employed which operates at an intermediate temperature such as liquid nitrogen temperature (77 K). In some applications, even another wall is introduced at 20 K to 30 K in order to minimize total refrigeration load.

On the other hand, a HTS cryostat operating at 77 K could employ a simple double wall construction. The outer wall is at room temperature and the inner wall is at liquid nitrogen temperature. The space between the two walls is filled with multi-layer-insulation (MLI) or some other suitable thermal insulation such as various types of foam. Normally this simple construction reduces heat leak to an acceptable level.

### Cooling System

Superconducting magnets can be cooled with a pool of liquid cryogen or cooled by conduction with a cryocooler. The majority of LTS magnets are cooled with liquid helium. Helium has the lowest boiling point (4.2 K) at atmospheric pressure of any known cryogen and has been the cryogenic fluid used in the LTS magnets. On the other hand, cost of helium is very high (~ \$5/liter). Because of its cost, most facilities install a recovery and reliquefying system. A variety of devices are available worldwide. The choice of device is application specific as the designer must evaluate trade-offs: primary among these are first-cost versus operating cost; first-cost versus reliability and ease-of-use, and dollar per watt of cooling required. Devices are basically two types: open cycle or closed-cycle. The former are the simplest, lowest cost form of refrigeration available today and are quite simply open top "bucket" dewars. Closed cycle systems do not require use of liquefied cryogen but rely on the refrigeration capacities of the gas and the cycle design to achieve the cryogenic temperatures desired. Helium is the predominant gas used in these devices. For larger applications such as a transformer, a closed cycle helium system is substantially more economical in long run than purchased liquid helium. With a closed cycle system, the warm helium returning from the cryostat is reliquefied in a refrigerator and returned to the cryostat. A number of conduction cooled magnets which use no liquid cryogen have been built for operation at 4 K and higher temperatures using both HTS and LTS wires.

### Economic Considerations

From a utility perspective, a transformer must have low initial and operating costs, and be light weight, compact, and environmentally benign, with a lifetime of typically 30 years. To a great extent, a HTS transformer does have a potential to offer these advantages.

Conventional transformers are highly reliable and flexible in terms of their use in an electrical system. However, the

dominant component of losses is the  $I^2R$  loss in the windings. The capitalized cost of these losses over the life of a transformer could easily exceed its initial cost. The superconducting transformers are attractive because of the potentially lower winding losses. However, there is an energy penalty associated with the input power consumed by the refrigeration system. This energy penalty can be substantial. To remove 1 W of losses at 4 K, 600 W of refrigeration power is required. However, only 20 W is required to remove 1 W from 77 K. This reduction in the refrigeration power has generated a lot of interest in transformers employing HTS conductors. An optimally designed HTS transformer is likely to have lower losses and lower life cycle cost than a conventional unit.

The higher current density capacity of superconductors compared to copper leads to a more compact and lighter design of transformers. Even for the identical core diameters, the core window width (space between iron legs) could be reduced in proportion to the space saving due to the utilization of superconducting windings. This reduces the iron core weight. Lighter core size also leads to lower core losses. A compact and light weight transformer might see new applications which were not feasible with the conventional transformers. Lower weight and compact size would make them acceptable for more urban applications. Smaller core windows also lead to lower leakage inductance which helps to improve dynamic stability of a power system. The low leakage inductance also improves the voltage regulation to the load, and therefore, it might eliminate complex and expensive tap changers.

The environmentally friendly aspect of a superconducting transformer gives additional impetus to application of these devices. They do not have environmental, health, and safety concerns associated with conventional transformers. In urban settings, most transformers are installed in the basement of high rise buildings. Environmental concerns are forcing utilities to employ oil free transformers. In a superconducting transformer, oil is replaced with liquid helium or liquid nitrogen. They are much more benign—nontoxic, nonflammable, and noncarcinogenic. However, in the event of a quench, a large quantity of helium or nitrogen gas could be released which could displace oxygen in the surrounding air and present a personnel risk. This risk can be mitigated by controlled release and installation of exhaust fans.

### STATE-OF-THE-ART OF SUPERCONDUCTING TRANSFORMERS

With the advent of HTS conductors, the low temperature transformer design efforts have essentially been abandoned. Several HTS transformers programs are currently active in Japan, Europe, and North America for operation at temperatures ranging from 20 K to 77 K. Both air-core and iron-core options are being pursued. The major HTS programs being pursued are listed below.

- ABB has built (26–30) and tested a 630 kVA, 3-phase transformer in March 1997. This was the world's first demonstration of an HTS transformer, one that was designed and built by ABB with HTS wire that was developed and manufactured by AMSC. Following the success of this transformer, the Electricité de France (EDF), ABB, and ASC consortium is now developing a 10 MVA transformer (29).

The 10 MVA transformer is a crucial next step on the path to a commercial-scale 30 MVA transformer. The 10 MVA unit will be built by ABB and will be tested by EDF in its grid by the end of 2000.

- Waukesha Electric and Intermagnetic General Corporation (IGC) is building a 1 MVA transformer (8). Waukesha will build the unit using HTS wire supplied by IGC. This unit was tested recently (8A).
- Kyushu University consortium designed, built, and tested a 500 MVA-class single phase HTS transformer (11) in 1996. This was the first transformer cooled by liquid nitrogen at 77 K and was operated at a steady state with a 500 kVA secondary inductive load.

Although HTS conductors are much more tolerant of transient heat input (primarily due to increased heat capacity of materials at higher temperatures than at 4.2 K), the ac losses are still significant and an attractive ac conductor configuration is still elusive. The highly aspected configuration of HTS wire (width = 10  $\times$  thickness) generates excessive hysteresis losses caused by the magnetic field component perpendicular to the surface of the conductor. These losses are sufficiently high that the total refrigeration power needed to remove them from 77 K environment is comparable to the  $I^2R$  loss in the windings of a conventional transformer. Thus, the efficiency (or life cycle cost) advantage is lost. How the ac loss problem will be solved and how the market will value the benefits of a superconducting transformer over conventional transformers is not yet clear. By the year 2002, a clearer picture might emerge about the feasibility of HTS transformers.

## BIBLIOGRAPHY

1. H. Kemersina et al., Application of superconducting technology to power transformers, *IEEE Trans. Power Appar. Syst.*, **PAS-100**: 3398-3410, 1981.
2. Westinghouse Electric Corporation, Application of low temperature technology to power transformers, US Department of Energy DOE, ET-20024-1, February 1992.
3. A. Feyrer et al., Preliminary tests on a superconducting power transformer, *IEEE Trans. Magn.*, **MAG-24**: 1477-1480, 1988.
4. S. Hornfeldt et al., Power transformer with superconducting windings, *IEEE Trans. Magn.*, **MAG-20**: 3550-3558, 1984.
5. M. Iwakuma et al., Quench protection of superconducting transformers, *IEEE Trans. Magn.*, **MAG-27**: 2090-2093, 1991.
6. E. S. Yoneda et al., Tests on a 30 kVA class superconducting transformer, *Cryogenics*, **31**: 655-658, July 1991.
7. B. F. Giese, The status of progress towards high-temperature superconducting transformers. Work sponsored by the International Energy Agency Argonne National Laboratory, November 1996.
8. S. P. Mehta et al., Superconducting transforming transformers, *IEEE Spectrum*, **43**: 49, July 1997.
- 8A. S. P. Mehta et al., Progress in the design of a 1-10 MVA class HTS transformer. Presented at the *Appl. Supercond. Conf.*, Virginia Beach, VA, Sept. 2000.
9. R. C. Johnson, Status of superconducting power transformer development, *Proc. Amp. Power Conf.* (Chicago), April 1996, 58-1-89, subject 05-6, Illinois Institute of Technology.
10. D. Mikulic, HTS transformer for utility powering, *Modern Power Syst.*, June 1997.
11. K. Funnki et al., Preliminary tests of a 500 kVA-class oxide superconducting transformer cooled by subcooled nitrogen, *IEEE Trans. Appl. Supercond.*, **7**: 824-827, 1997.
- 11A. H. Hata et al., Study of superconductivity transformer for Shokusanen (doling) pack, Presented at the *Appl. Supercond. Conf.*, Virginia Beach, VA, Sept. 2000.
- 11B. N. Hayakawa et al., Study of superconductivity fault current limiting transformer, Presented at the *Appl. Supercond. Conf.*, Virginia Beach, VA, Sept. 2000.
12. P. K. Mukhopadhyay et al., Unexpected behaviour of core materials of electrical coils at low temperatures, *Cryogenics*, **34**(3): 241, 1994.
13. A. P. Malozemoff et al., Progress in Bi2220-2225 tape technology, M2S Conf., Beijing, China, February 26-March 3, 1997.
14. M. S. Walker et al., Performance of coils wound from long lengths of surface coated, coated Bi2220-2225 conductor, *IEEE Trans. Appl. Supercond.*, **7**: 889-892, 1997.
15. M. Yamamoto, A Study on a Coreless Superconducting Transformer, *IEEE Trans. Appl. Supercond.*, **3**: 468, 1993.
16. S. Okano et al., Fabrication and test of superconducting air-core autotransformers, *IEEE Trans. Magn.*, **MAG-28**: 430-433, 1992.
17. M. N. Wilson, *Superconducting Magnets*, Clarendon Press, Oxford, 1983.
18. T. Verhaege et al., A new class of ac superconducting conductors, *IEEE Trans. Appl. Supercond.*, **3**: 164, 1993.
19. K. Kamisada et al., Research and development of superconductors for commercial frequency, *IEEE Trans. Appl. Supercond.*, **28**: 291, 1992.
20. T. Vernege et al., Protection of superconducting air windings, *IEEE Trans. Appl. Supercond.*, **28**: 751, 1992.
21. M. Iwakuma et al., Electromagnetic properties in parallel conductors composed of Bi2220 multifilamentary wires for power transformer windings, *Appl. Supercond. Conf.*, August 25, Pittsburgh, PA, 1996.
22. A. M. Wolsky, IEA implementing agreement for a cooperative programme for assessing the impacts of high-temperature superconductivity on the electric power sector, Argonne National Laboratory, January 1997.
23. S. Fujita et al., 75 Spring MRS Meeting in San Francisco, California.
24. H. Roesch, *Superconducting Magnet Systems*, Berlin: Springer-Verlag, 1973.
25. K. Kwasnitza, AC losses of superconducting high-Tc multifilamentary Bi-2223/Ag sheathed tapes in perpendicular magnetic fields, *Physica C*, **233**: 425-435, 1994.
26. J. H. Murphy, Advanced applications of superconductors, *Advances in Cryogenic Engineering*, **20**: 173, 1984.
27. Ch. Gossens et al., Effect of the longitudinal magnetic field on a 2nd level superconducting cable, *Trans. IEEA*, **46** (1): 52-57, January 1997.
28. ABB and American Superconductor Corporation team up to demonstrate the world's first installation of a high temperature superconductor transformer, an ABB press release, Westborough, MA, March 13, 1997.
29. Electricité de France, ABB, and American Superconductor sign \$15 million parts accelerate HTS wire development for next generation power transformers, an ABB press release, Westborough, MA, January 12, 1999.
30. P. C. Thessand et al., High temperature 650 kVA superconducting transformer, Report 12-302, CIGRE Meeting, Aug. 30-Sept. 5, 1998, Paris.

SWAIN S. KAUST  
American Superconductor  
Corporation

## TEMPERATURE SENSORS

## TEMPERATURE AS A PHYSICAL QUANTITY

Every object or phenomenon existing in the real world may be described by a set of properties. Some of these properties are physical quantities, while others are descriptive ones. Physical quantities must be measurable. In order to make a property measurable, one has to establish a method with which to detect whether a state of a property is higher than another state, to detect if two states are the same, to propose a measure of the state, and finally to find a scale which transforms a given level of the property into an abstract symbol in the form of a number and a unit of measure. Considering temperature as a property of an object, all the problems mentioned above are rather complicated from both theoretical and practical points of view. The human sense of touch enables us to distinguish between higher and lower thermal levels, over a limited range of temperatures and with very limited repeatability, but nothing more, and there was a long way to go from our "feeling of heat" to the definition of temperature and temperature scales. The popular and often-quoted definition of temperature as an "intensity of heat" does not lead directly to solving the problem because of its lack of clarity.

Only the discovery of the fundamental laws of thermodynamics in the middle of nineteenth century allowed us to answer the question: What is temperature? The first law of thermodynamics says that thermal energy transfer is possible only from a system with higher temperature to a system with lower temperature. By observing the direction of thermal energy transfer, we are able both to tell which system is the one of a higher state of temperature and also to confirm the existence of equilibrium of temperature states when the heat transfer between two systems declines to zero. Furthermore, the works by Carnot, Lord Kelvin, and Clausius resulted in the formulation of the laws concerning the reversible thermodynamic cycle, called the Carnot cycle. The Carnot cycle consists of two isothermal heat conversion processes and two adiabatic heat transfer processes as illustrated in Fig. 1. By transferring a heat energy from the system with a higher state of temperature to a system with a lower temperature state, it is possible to transform a part of that energy (although a relatively small one) into mechanical energy. This constitutes a theoretical principle for all heat engines. The theoretical efficiency of a Carnot cycle engine is

$$\eta = \frac{Q_1 - Q_2}{Q_1} \quad (1)$$

where  $Q_1$  is thermal energy transferred from the system with higher temperature in the isothermal expansion process, and  $Q_2$  is thermal energy transferred to the system with lower temperature in the isothermal compression process. The theory of the Carnot cycle does not depend on the medium used, and Eq. (1) is a universal one. The engine efficiency depends only on the ratio  $Q_2/Q_1$  and this ratio was proposed by Lord Kelvin as the basis of a new "absolute" thermodynamic measure of temperature in the form

$$\frac{T_2}{T_1} = \frac{Q_2}{Q_1} \quad (2)$$

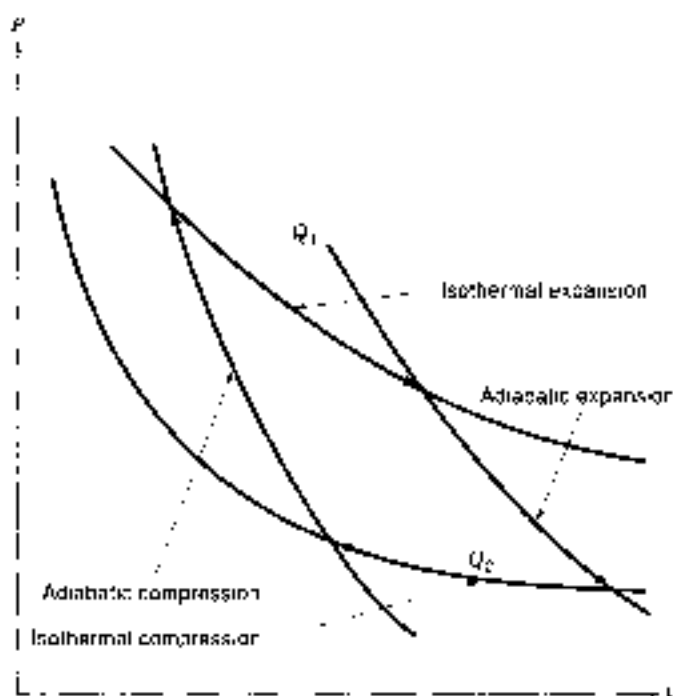


Figure 1. The Carnot cycle for an ideal heat engine. The arrows indicate the subsequent thermal processes.

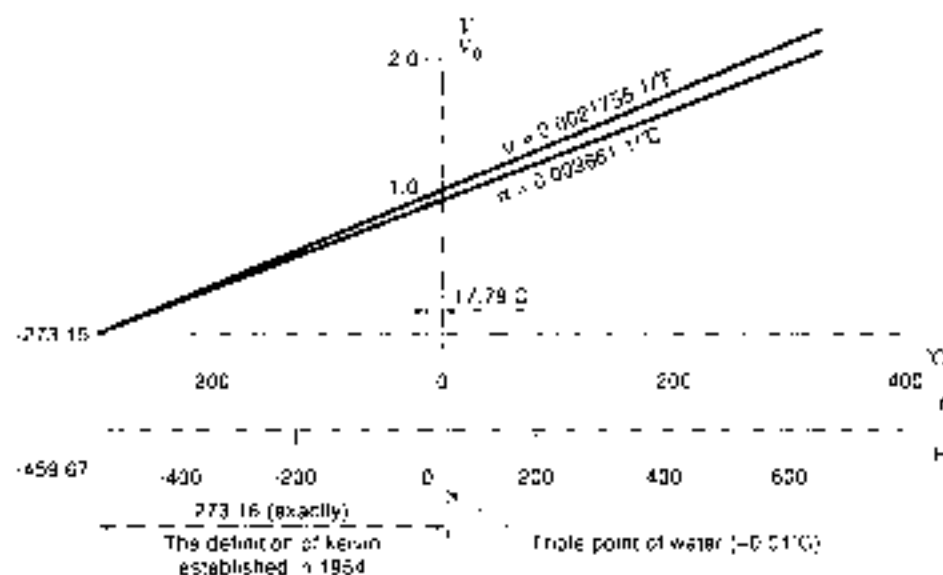
Such a measure is independent of the thermal properties of any material and requires only one fixed temperature point to create the temperature scale. Equation (2) indicates very clearly that temperature has an physical zero point. In creating any temperature scale the zero point is to be assumed arbitrarily. That is, temperature, in the same manner as time, can be measured only by the interval scale but not by the metric scale. The point defined as 0 K is only a mathematical point on the scale but not a physical null temperature. Even in outer space we are not able to achieve physical null temperature, because the insertion of any material body changes the state of the temperature previously existing.

In Lord Kelvin's lifetime the Celsius scale of temperature had been in use for 100 years, and the temperature differences (e.g., between the ice point and the boiling point of water) had already been expressed in Celsius degrees. Gay-Lussac's gas law had also been known for more than 100 years and expressed as

$$V = V_0[1 + \alpha(t - t_0)]_{p, \text{const}} \quad (3)$$

where  $V$  and  $V_0$  are the volumes of an ideal gas at temperatures  $t$  and  $t_0$ , respectively. Equation (3) proved to give results fitting quite well to the experimental results, especially for rarefied gases at low pressure, and the numerical value of the coefficient  $\alpha$  was known with a good accuracy. Taking a Celsius degree as a unit of temperature differences, assuming the ice point of water as a point of  $t_0 = 0$ , and accepting particularly for these conditions an experimentally obtained value of the coefficient  $\alpha = 0.003661 \text{ } 1/^\circ\text{C}$ , we are able to create a new temperature scale, with the zero value at the point when an ideal gas volume decreases to zero (Fig. 2). The scale is now known as the absolute or Kelvin scale and is shifted by a value of  $1/\alpha = 273.15$  with respect to the Celsius scale. All currently





**Figure 2.** The meaning of absolute temperature scale and its relation to the ideal gas law. The difference between ice point and triple point of water on the temperature scale is excessively enlarged in order to enhance clarity.

used temperature scales are in linear relations with each other.

$$T(\text{K}) = t(^{\circ}\text{C}) + 273.15 = (t(^{\circ}\text{F}) + 459.67) \frac{5}{9}$$

$$t(^{\circ}\text{C}) = T(\text{K}) - 273.15 = (t(^{\circ}\text{F}) - 32) \frac{5}{9}$$

$$t(^{\circ}\text{F}) = 1.8t(^{\circ}\text{C}) + 32 = 1.8T(\text{K}) - 255.37$$

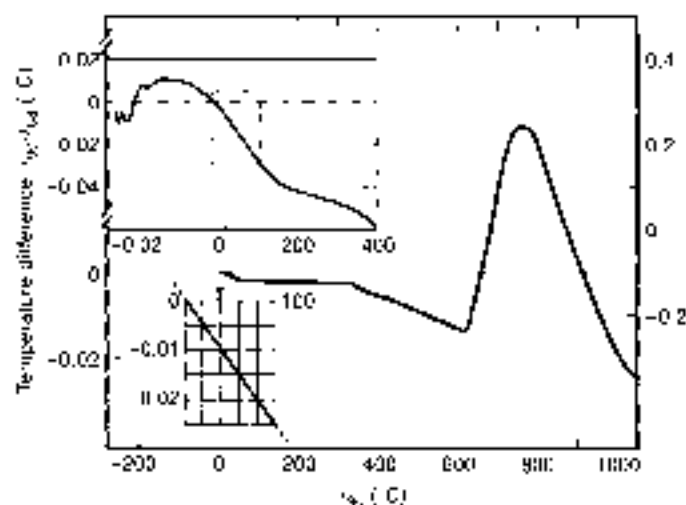
It is evident that Gay-Lussac's law is an excellent basis for developing not only the scale but also a thermometer: a gas thermometer. There is only one problem, albeit a significant one: An ideal gas does not exist. This inconvenience may be overcome either by the use of gases with properties close to the ideal gas ( $\text{He}$ ,  $\text{H}_2$ ) or by the use of rarefied gases at low pressures and by applying a very special measurement procedure. It has to be pointed out that in Eq. (3) there is an implied condition that  $p = \text{const}$ . In practically used gas thermometers, however, it is easier to fulfill the requirement of  $v = \text{const}$  and to observe the changes in  $p$ , instead of  $t$ . Therefore the majority of gas thermometers work on the constant volume principle.

### INTERNATIONAL TEMPERATURE SCALE OF 90

In spite of the great progress in measurement techniques achieved since the days of Lord Kelvin, the use of gas thermometers is not a way to establish a contemporary temperature scale because of the great difficulties regarding their performance. Absolute Kelvin temperature according to Eq. (2) remains the theoretical definition of temperature, but the temperature scale is reproduced with the highest accuracy by means of the International Temperature Scale (ITS) established first in 1927 and recently modified in 1990. Most high accuracy industrial temperature measurement requires the production of a reproducible temperature state rather than its value in terms of its absolute temperature value. ITS-90 allows this by setting up a scale that is highly reproducible but only approximates to an absolute Kelvin scale to the limits of technology available in the late 1980s. Temperatures corresponding to that

scale are sometimes marked using the subscript 90 in order to distinguish the differences with respect to the former scales (1927, 1948, 1988) and to the theoretical values of the thermodynamic scale ( $T$  [K] and  $t$  [C]) with no subscripts).

The ITS-90 describes a set of fixed temperature points (17 points) and the methods of measuring the temperature between these points. Fixed points are freezing or triple points of selected materials (except Gallium's melting point). The points and the methods have been chosen to ensure, according to the actual knowledge and technology, the best conformance to the absolute temperature scale. An additional document, "Supplementary Information for the ITS-90," gives a very in depth and exhaustive description of the instruments and the procedures which ensure the highest accuracy and traceability of temperature standard measurements. Figure 3 presents the difference in temperature values expressed by the previous IPTS 68 scale and by the present ITS-90. In some ranges, especially above 500 C, the differences are really great and are mainly caused



**Figure 3.** The differences in temperatures expressed by IPTS-68 and ITS-90. (After Ref. 1.)

by the incorrect reference function accepted by the IPTS-68 for the type S standard thermocouple. The correction of that reference function allows us to reduce the differences. Boiling points are rejected by the ITS-90 because of their poor stability and great sensitivity to pressure. The boiling point of water is no longer a fixed point. Some differences in standard instruments and methods have been introduced, too. The most important is that the standard platinum resistance thermometer now covers a much wider range of temperatures than previously, extending from about  $-260$  C up to the freezing point of silver  $+962$  C. Above that temperature the Planck radiation law is used as a principle for standard measurements. Thus, the PtRh-Pt thermocouple is no longer a standard thermometer (1).

Extremely complex reference functions have been defined in order to express sufficiently precisely the ratio of the resistance at a given temperature  $T_{20}$  to the resistance at the triple point of water ( $T_{20} = 273.16$  K) and vice versa, to express the temperature as a function of that ratio. The convenience of computerized calculations justify the complex forms of these functions.

They are not used, however, for numerical corrections of temperature values in the industrial, microprocessor-based instruments (e.g., for linearity correction). For that purpose much simpler, yet not so accurate, equations have been developed. Equation (4) is an adequate example.

ITS 90 serves as the best approximation of a realized temperature scale to the absolute thermodynamic scale and determines a highest level of the temperature standards. In every country, several levels of temperature standards are available, which are used for comparisons and to calibrate the thermometers in practical use. The uncertainty is higher as far as the calibration of the technical thermometers is concerned, and it ranges from a few millikelvin or less at the highest accuracy level and in the medium temperature range to the tenths of Kelvin for the case of industrial thermometers and thermometers used in everyday life. The uncertainty of both standard and practical thermometers in the higher temperature range (above 650 C) is always greater. According to the general idea of the ITS-90, it is evident that some modifications are inevitable in the future.

## THE GENERAL PRINCIPLES OF TEMPERATURE MEASUREMENTS

The measurement of temperature differs from the measurement of other fundamental quantities such as mass, force, longitude, or voltage not only because of the lack of physical zero

point of temperature, but primarily because of the inconvenience in direct comparison of the thermal state of the system of unknown temperature with the thermal state of the standard. The temperature is an intrinsic property of a material and hence does not permit scaling in the way of an extrinsic property such as length or mass. To measure temperature it is necessary to find an extrinsic property that varies in a predictable way with temperature and use this to construct a thermometer. That is why the practical measurements of temperature are always performed indirectly. The temperature sensor interfaces with the system whose temperature is to be measured. The interface may be realized by insertion, by point contact, or by visual contact with the system (Fig. 4). The sensor converts the thermal state of a system to a determined state of another quantity, which is defined as an output signal from the sensor. The output signal is then processed in the transducer 'T' and finally presented in numerical form as a result of the temperature measurement. However, it is not the only function that contemporary transducers perform. They are more and more frequently equipped with microprocessors and constitute a system which controls the measurement process, controls range changes, performs the numerical result correction, presents the results in appropriate units, and controls the standard interfaces such as RS 232, IEC 625, or others. Many control systems work according to the two-wire 4-20 mA standard. Therefore it happens very often that a temperature-measuring transducer provides the 4-20 mA output too, or even has an incorporated PID or on/off controller. Sometimes sensors are integrated with transducers either mechanically only or both mechanically and electrically on the same chip. At a large enough scale of integration, the term *integrated sensor* (IS) is justified. (The reading of the acronym IS as "intelligent sensor" is a commercial and marketing abuse.)

It is not the aim of this article to describe transducers but sensors. Therefore only input stages of the transducers—those stages which are directly interfacing with the sensors—will be presented. They are usually analog parts of measuring circuits, because all temperature sensors are analog devices. The principle of sensor operation depends on the physical phenomena used for conversion of temperature to the output signal. There are a lot of physical phenomena where temperature plays a significant role, but only a few of them are used in temperature sensors because they have to meet some additional requirements. These are as follows:

1. The monotonic calibration curve—that is, the relation between the temperature and an output signal over a sufficiently wide temperature range.

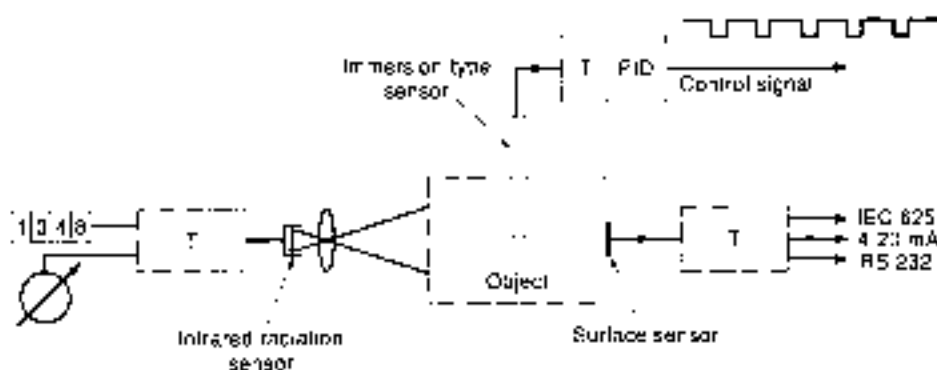


Figure 4. Temperature sensors and transducers. The outputs are arbitrarily assigned to the transducers.

2. Sensitivity to temperature that is much higher than the sensitivity to all other influencing variables
3. The output signal easily measurable with sufficiently low uncertainty.
4. Interchangeability of the sensors at least within the same type or the same manufacturing technology.
5. Repeatability of the calibration curve over the whole range of operating conditions.

Repeatability is mainly disturbed by hysteresis, relaxation, and aging. Hysteresis is observed when the calibration curve taken for increasing temperatures differs from that taken for decreasing temperatures. Relaxation means a delay in approaching the stable value of measured temperature after a rapid change in the sensor temperature. Aging is a long time process which causes the shift of the calibration curve due to many, sometimes unknown, reasons. Recalibration reduces, at least for a limited period of time, the errors caused by aging.

The linearity of the calibration curve and a large output signal are no longer essential requirements because of the progress in signal conditioning technology.

In order to produce a good sensor, the above requirements have to be fulfilled, but they are not sufficient to ensure a proper temperature measurement. The quality of temperature measurement depends to a great degree on the design of the sensor adequate to the conditions where the temperature has to be measured and on a proper measurement procedure. These two aspects are general ones, and are valid for all measurement techniques, but for the temperature measurements their importance is particularly great. It is due to the fact that every temperature sensor measures its own temperature and, more precisely, the temperature of its own sensitive part (the thermometric body). The designer and the user of a sensor have to ensure the sameness of that temperature with the temperature which is defined as a measurand (that one which has to be measured in the given particular circumstances). For that purpose the sensor should be brought into as close thermal equilibrium with the measurand as possible without disturbing the measurand's thermal state. It requires a good thermal connection to the thermometric body and a poor thermal connection to the environment. The difference in thermal conductivity between thermal insulating and conducting materials is not very high. This involves some difficulties in design of a good thermometer and leads to measurement errors. Errors and uncertainties in temperature measurements will be discussed in more detail in the last section of this article. The problem of the thermal burdening by a sensor does not exist in radiation thermometry, but other sources of errors occur in this case.

## RESISTANCE SENSORS

In all resistance sensors the change of their resistance follows the temperature changes, but the way it happens is different in metal sensors and in semiconductor sensors. It is therefore reasonable to separate the considerations about those two groups of resistance sensors. Furthermore, within the group of metal resistance sensors there is a great difference in the design of precise and industrial thermometers. This difference justifies further subdivision of the discussion.

### Precise Resistance Sensors

The resistivity of almost all metals depends on temperature, but only a few of them are used in resistance thermometers, those which meet the requirements listed in the previous section. Pure platinum is considered the best material for temperature sensors. It has a relatively high resistivity (15 times greater than copper); thus wires needed to form resistors do not need to be particularly thin. Platinum can be obtained in a pure form with few impurities ensuring repeatability and interchangeability of the sensors. However, the most important reason why platinum is so widely used for temperature sensors is its ability to withstand even severe environmental conditions at high temperatures. For this reason, only pure platinum is used in the standard temperature sensors.

The progress in the technology and in the design of platinum temperature sensors achieved in the past decades made it possible to eliminate the PtRh-Pt thermocouple from the list of standard thermometers, which define temperatures according to ITS-90. Now, the temperatures  $T_{\infty}$  in the range between 13 K and 1233 K (960 °C) silver point are defined by the standard platinum resistance thermometer (SPRT). SPRTs are situated at the top of the hierarchical system of propagation of standards and are used as a first link, which links ITS with all other temperature standards. They are used only occasionally for measurement of unknown temperature but more frequently for calibration purposes only. It is evident that to satisfy such high demands the quality of SPRTs must be the highest one. The purity of platinum used is secured by meeting two requirements:  $R_{102}/R_{100} \leq 0.544235$  and  $R_{100}/R_{100} > 1.11807$ , where  $R_{102}$ ,  $R_{100}$ , and  $R_{100}$  are resistances at Hg point, 100 point, and triple point of water, respectively. These requirements are much greater than those required for industrial thermometers. In order to achieve such high values, the purity of platinum has to be greater than 99.999%. The influence of impurities is much stronger at lower temperatures, limiting the temperature range of SPRTs. The next limitation is a very low resistance at 13 K with respect to the resistance at the triple point of water,  $R_{13}/R_{100}$  (approximately one-thousandth), which makes the calibration process more complicated and results in decreasing sensitivity.

Standard resistors are always wire-wound, but the cores differ according to the temperature range. For the lowest temperatures, resistors are encapsulated in a hermetically sealed platinum sheath filled with helium under the pressure of 30 kPa. Such a design makes the sensor short and vacuum-protected as required for calibration in cryostats used for realization of ITS-90 fixed points. For higher temperatures, SPRT sensors are fixed at the end of long Inconel or quartz glass tubes, because of the necessity of providing a deeper penetration in the calibration devices. For temperatures above 650 °C, some special materials such as silica, alumina, or sapphire must be used. The wire diameter of high-temperature SPRTs is greater, exceeding 0.5 mm, resulting in a lower resistance of 2.5  $\Omega$  or even 0.25  $\Omega$  as compared with 25  $\Omega$  for SPRTs used at lower temperatures and wound from 0.05 mm wire.

### Industrial Metal Resistance Thermometers

The industrial resistance thermometers used for measurement and control purposes in manufacturing plants, in the automotive industry, in housekeeping equipment, for environmental

measurements, and for many other everyday purposes are much less accurate than standard resistive thermometers. Industrial thermometers differ from SPRTs not only by design, technology, and material used, but also by the idea of its implementation. For the purposes of the ITS and of the calibration performed with SPRTs, the ratio of two resistances at two different temperatures is taken as a measure of the temperature. In "normal," not standard, temperature sensors the value of the resistance of the thermometer becomes a measure of temperature. In other words, the output of a standard thermometer is the resistance ratio, and the output of an industrial thermometer is its resistance. The abbreviation for industrial thermometers is PRT (without "S") or more frequently RTD (resistance temperature detector), indicating that sensors in use are not only platinum. Most of the RTDs all over the world are adjusted to the nominal value equal to 100.00  $\Omega$  at 0  $^{\circ}\text{C}$ , and hence termed as Pt-100 or Ni-100. The relationship between resistance and temperature for platinum RTDs is much simpler than for SPRTs and may be expressed in the form of the following equation:

$$R(t) = R_0[1 + At + Bt^2 + C(t - 100)^3] \quad (4)$$

where  $A = 3.90802 \times 10^{-3} \text{ } ^{\circ}\text{C}^{-1}$ ;  $B = -5.802 \times 10^{-7} \text{ } ^{\circ}\text{C}^{-2}$ ;  $C = 4.27 \times 10^{-12} \text{ } ^{\circ}\text{C}^{-3}$  for  $t < 0$ , and  $C = 0$  for  $t > 0$ . However, it is not the equation but the values of resistances corresponding to appropriate temperatures that are the subject to national and international (IEC) standards, in the form of reference tables. Sometimes a distinction is introduced between "European" sensors with  $R_{100}/R_0 = 1.365$  and "American" sensors with  $R_{100}/R_0 = 1.392$ . Furthermore, the uncertainties allowable for these sensors are also set in national standards, which are normally very close to the international IEC standards (Fig. 5). Standardization secures the reproducibility and hence interchangeability of RTDs, which is one of the most significant advantages of these sensors over all other temperature sensors.

It is a common practice that the repeatability of each individual sensor—especially over a limited temperature span—is in fact much better than the standard uncertainty limits. Therefore the individual recalibration of RTD sensors is recommended, because it allows further improvement of the accuracy. Due to the cost and available technical equipment, such calibration is commonly performed at one or two fixed points

only. One-point calibration enables us to take into account an additive component of the difference between the nominal and the actual value of the resistor (i.e., additive error). Two-point calibration enables us to account for the sensitivity difference (i.e., the multiplicative error too). Nonlinearity error remains unknown. If the range of measured temperatures is limited (as is usually the case in practical situations), the nonlinearity of the RTDs has little influence on the total uncertainty and may be neglected. For wider temperature ranges the reference tables or appropriate *R(t)* equations like Eq. (4) are useful for identifying the nonlinearity component and for applying it for the correction of the measurement result, together with the correction of the additive and multiplicative errors determined during the two-point calibration.

Besides platinum, some other metals and alloys are also used for temperature sensors. Nickel and copper are utilized for temperature measurements over a narrow range. The sensitivity of nickel sensors is higher than that of the platinum sensors, but their nonlinearity is greater. Copper sensors are known to be extremely linear, but due to their low resistivity, it takes very long and thin wires to produce a 100  $\Omega$  resistor. Therefore, lower nominal values of copper sensors are also allowed by standards.

With the wire-wound sensors two designs are usually used. In the first one the wire is bilinearly wound on a glass, quartz, or ceramic rod or pipe and coated with fired glass layer (Fig. 6a-b). Glass other than quartz glasses is unsuitable for sensors above about 250  $^{\circ}\text{C}$  as electrical insulation properties begin to deteriorate rapidly. For sensors working at temperatures above 600  $^{\circ}\text{C}$ , glass is not a proper material and is replaced by alumina ( $\text{Al}_2\text{O}_3$ ). The difference in the thermal expansion coefficient of core and platinum causes stress, which influences the long-term stability of the sensors. The second design is stress-free because the helical winding is placed in two holes drilled along the ceramic core and only sealed at the both ends of the core (Fig. 6c). Often, two independent sensors are placed in four separate holes in the body. One of them may be replaced by the second in case of damage, or, more frequently, one serves for measurement and recording purposes while the second serves as a control. It is also very important to ensure the high shunting resistance from the internal mounting structures of the RTDs.

Sensors are protected from mechanical and chemical influences by metal tubes of different length (up to 2 m) made

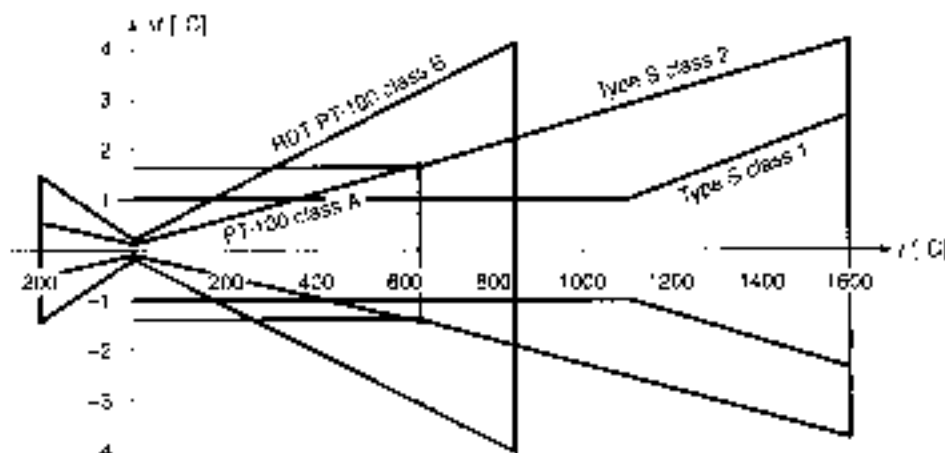
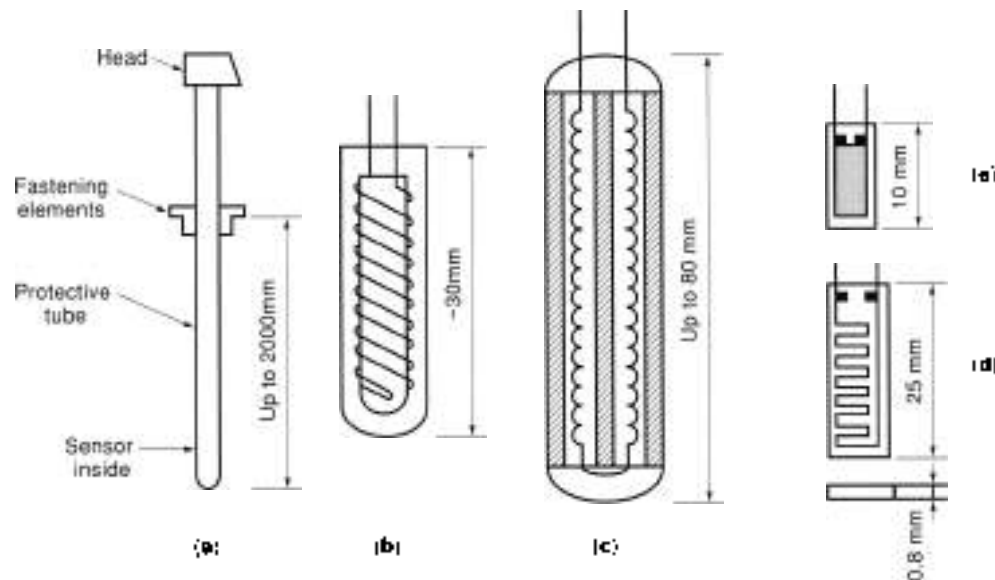


Figure 5. Comparison of allowable uncertainties of RTD Pt-100 after IEC Publication 751 and thermocouple type S after IEC Publication 584. Higher accuracy of resistance thermometers at lower temperatures is evident.



**Figure 6.** Immersion type thermometer (a) and four types of sensors: ceramic sealed type (b), free wire type (c), thick film (d), and thin film (e). Types b and c are commonly used in industrial thermometers.

of stainless steel, nickel alloys, and sintered alumina and equipped with fixing elements for fastening them to the objects where the temperature should be measured [Fig. 6(a)]. The choice depends on the kind of object and the range of measured temperatures, pressures, and other environmental conditions. Because the sensor is placed at the bottom of the tube, it is convenient to lead out all the wires from one end of the sensor. In fact, this is the reason why the bipolar winding is used. It protects against the induced noise voltage too. The best thermal contact between the sensor and the protecting tube is highly recommended.

Besides wire-wound RTDs, a group of sensors exists where a metallic layer is deposited on a flat or cylindrical core. The core material most frequently used is alumina, and a metal layer is deposited either as a thick film (5  $\mu\text{m}$  to 10  $\mu\text{m}$ ) in a screen printing process [Fig. 6(d)] or as a thin film (1  $\mu\text{m}$ ) by sputtering [Fig. 6(e)]. The laser cutting provides the adjustment to the required resistance value. Sensors are coated with a thin protective layer of overglaze. Short response times of such sensors result from the small dimensions and small mass of the sensors. Long-term stability is a bit worse, and the temperature range is restricted to 500 C, but it probably will change with the advances in technology. Deposited RTDs are used in laboratory and service hand-held thermometers and in instruments requiring relatively accurate but small sensors for thermal control or correction. Psychrometric humidity sensing instruments also use deposited sensors. The electric signal which is obtained from these sensors is smaller than the one from traditional sensors because of the lower magnitudes of applied supply current.

## SEMICONDUCTOR SENSORS

Semiconductor material may be used in resistance thermometers such as negative temperature coefficient thermistors, linear KTY sensors, and germanium resistance sensors used in cryogenic temperatures, but also in semiconductor devices such as diodes, transistors, and integrated circuits, the operation of

which is related to the properties of  $p$ - $n$  junctions being the essential part of each semiconductor device. Classification and terminology are not established, but for the purposes of this article the distinction will be made between semiconductor resistance sensors and semiconductor active sensors.

### Thermistors

Negative temperature coefficient (NTC) thermistors are prepared from a mixture of powdered metal oxides such as  $\text{MgTiO}_3$ ,  $\text{MgO}$ ,  $\text{CuO}$ ,  $\text{NiO}$ , and  $\text{CoO}$ , along with others sintered at the temperature of 1300 C. During that process, some  $p$ - and  $n$ -type semiconductor centers are created, thus enabling resistance-temperature relations to be described as semiconductorlike. In semiconductors, both electrons and holes are responsible for the conductivity.

$$\sigma = \frac{1}{\rho} = \frac{q^2 n \tau_n}{m_n^*} + \frac{q^2 p \tau_p}{m_p^*} \quad (5)$$

where:  $\sigma$  is the conductivity, reciprocal to resistivity  $\rho$ ;  $n$  and  $p$  are the numbers of electrons and holes in the valence band, respectively;  $\tau_n$  and  $\tau_p$  are their relaxation times; and  $m_n^*$  and  $m_p^*$  are their effective masses. In the semiconductors,  $\tau_n$  and  $\tau_p$  remain constant but  $n$  and  $p$  values change with temperature according to the relationship

$$n = p = 2 \left( \frac{kT}{2\pi h} \right)^{3/2} (m_n m_p)^{1/4} e^{-E_g/2kT} \quad (6)$$

where  $E_g$  is the energy of the band gap, and  $k$  and  $h$  are Boltzmann's and Planck's constants, respectively. From Eqs. (5) and (6) we obtain

$$\rho = CT^{-3} e^{E_g/2kT} \quad (7)$$

In the range of temperature in question (250 K to 400 K), the last term of Eq. (7) dominates. This fact leads to the well-known

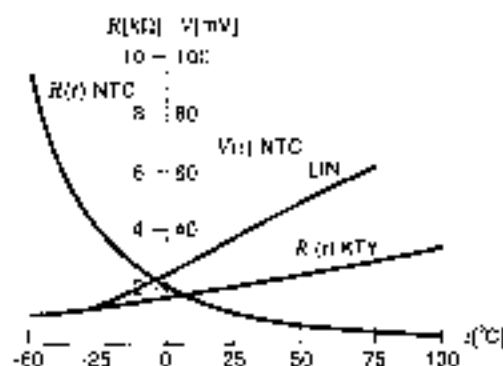


Figure 7. Calibration curves of an NTC thermistor, a KTY temperature sensor, and the output voltage of linearized circuit according to Eq. 115 for temperature span  $-25$  to  $+75$  C. The sensitivity of linearized NTC sensor is twice the sensitivity of KTY which does not need any linearization procedure.

relationship between temperature and resistance of NTC thermistors in the form of

$$R = R_0 e^{B/T} \quad (8)$$

$R_0$  has a very small value and no physical interpretation. More practical therefore is the equation

$$R = R_{25} e^{B(1/T - 1/298)} \quad (9)$$

where  $R_{25}$  is the thermistor resistance value at 25 C (500  $\Omega$  to 20 k $\Omega$  are typical values), and  $B$  is a material constant (Fig. 7). The value of  $B$  does not correspond strictly to  $E_g/2k$  because many other factors influence the resistivity of the semiconductor, and additionally other mechanisms of conduction exist in the thermistor structure. Therefore the value of that constant depends on the material and manufacturing technology of the thermistor, and normally we obtain  $B \approx 3000$  K to 5000 K. By describing the relative sensitivity of the thermistor in the same way as for the metal sensors, one obtains

$$S = \frac{1}{R} \frac{dR}{dT} = -\frac{B}{T^2} \quad (10)$$

from which the value of  $S \approx -0.03$  K $^{-1}$  at 25 C is approximately 10 times greater than that for metal sensors.

There are two principal types of thermistors commercially available. The first is the bead type, where the sintered material is formed into a bead of 0.1 mm to 1 mm diameter and sealed in glass stick or epoxy sheath together with two platinum connecting leads. The second and cheaper type is a disk thermistor, where the metal oxides are pressed at 1000 C into the forms of disks, tablets, and bars. Disk type thermistors are less stable and are used for temperature compensation in electronic circuits. Only bead type thermistors may be used as temperature sensors because their stability is much better, and after annealing at 60 C a repeatability level of  $\pm 10$  mK may be achieved. Unfortunately, the interchangeability of thermistors is rather poor. Both parameters  $R_{25}$  and  $B$  differ for individual thermistors, even taken from the same batch (3). This is of particular importance because of the nonlinearity of the thermistors calibration curves. The methods of matching the  $R(t)$  characteristics are much more complicated for nonlinear characteristics than for the linear ones. The International Standardization Organization (ISO) has attempted to unify thermal

calibration curves by introducing the so-called ISO curve thermistors. The standardization concerned the  $R_{25}$  value (0.5, 1, 2, 4, 15, 100 k $\Omega$ ), the shape of the calibration curve, and the admissible limits of interchangeability (from  $\pm 0.25\%$  to  $\pm 5\%$  of resistance span). Such thermistors are much more expensive than the ordinary ones.

Positive temperature coefficient (PTC) thermistors are used as switching elements rather than as temperature measuring sensors, because of their bistable calibration curve.

### Bulk Silicon Sensors

Extremely low doped bulk silicon material shows a different mechanism of conductivity. At temperatures above 100 K, all free electrons become ionized and the temperature influences only the relaxation times, which decrease with the increase of temperature. As a consequence, the resistivity of doped silicon increases, creating a positive slope of the calibration curve of a respective sensor. At higher temperatures, however, the process of thermally excited electrons dislocating from the valence band to the conductivity band becomes more evident and stops the increase in resistivity. The mechanism described above may be practically used only when a  $p-n$  junction is created in the bulk material. This is why a special technique of resistance measurement has to be used. The technique is based on a great difference in size between the electrodes used for the measurement. One electrode is only micrometers in diameter, while the other covers the whole counter surface of the semiconductor piece (Fig. 8). This creates a great difference in the electric field density at both electrodes, and therefore only the part of the material with high field density is responsible for the measured resistance  $R$ . The relationship between the resistance and the semiconductor resistivity is given in the form

$$\rho = \frac{R}{\pi d} \quad (11)$$

where  $d$  is the fine electrode diameter.

The commercially available temperature sensors, which work according to the described principle, are known as KTY linear sensors. In fact there are two small electrodes of 22  $\mu$ m diameter, and the "back side" of the bulk silicon material is coated by a conductive layer. Such a design makes the sensor insensitive to the current direction. The very precise doping control of the material is realized by means of neutron implantation in which the silicon atoms are replaced by phosphorus atoms with excellent homogeneity over the whole bulk material. The resistance of KTY sensors at ambient temperatures is about 1 k $\Omega$  to 2 k $\Omega$ , their sensitivity is about 1%/K, and the operation temperature ranges from  $-50$  C to  $+125$  C (Fig. 7).

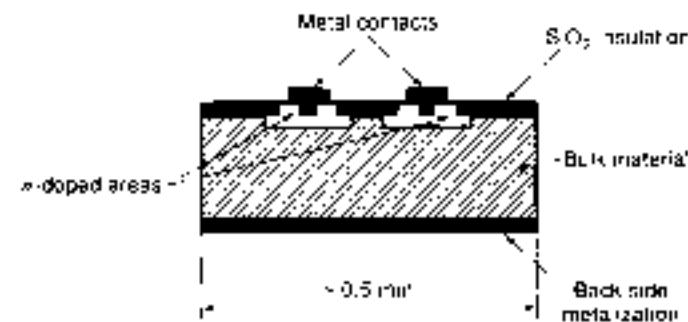


Figure 8. The design of a KTY sensor.

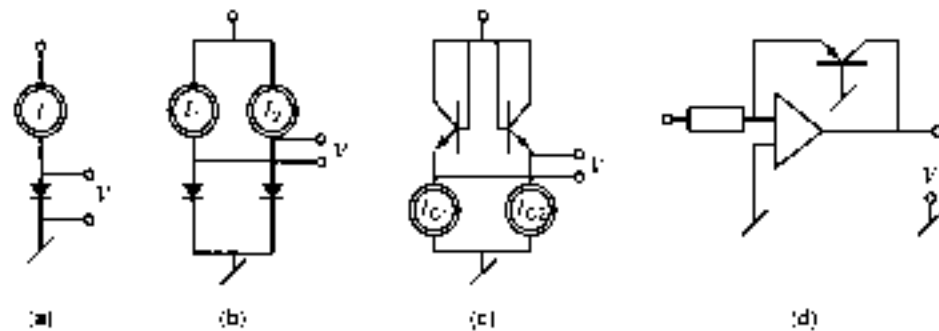


Figure 9. Diode and transistor temperature sensors: simple diode sensor (a), double diode sensor (b), double transistor sensor (c), and simple integrated circuit (d).

### Active Semiconductor Sensors

Active semiconductor sensors are those whose operating principle relies on the potential barrier between the conducting layer and the valence layer in a semiconductor, as in diodes and transistors. The simplest semiconductor device is a diode. According to the Shockley theory, the relationship between current  $I$  and voltage  $V$  in the forward polarized diode may be expressed as

$$V = V_0 + \frac{kT}{nq} \ln \frac{I + I_S(T)}{I_S(T)} \quad (12)$$

where  $V_0$  is a barrier voltage,  $n$  is a coefficient (not considered in the simplest form of the theory), and  $I_S(T)$  is reverse saturation current, many times smaller than forward current  $I$ , but very strongly dependent on the temperature. Every temperature increase by 10 K results in its value doubling, and this behavior determines the temperature properties of the diode. In the temperature range of 150 K to 400 K the thermometric calibration curve of the diode is nearly linear, with the negative slope of approximately 2 mV/K to 2.3 mV/K. Unfortunately the  $I_S$  value depends not only on the temperature but also on many factors hardly controllable in the manufacturing process. Therefore, the diode sensor's interchangeability is poor. A single diode is the cheapest temperature sensor, but each one has to be individually calibrated at one or better at two points. Even in that case the uncertainty, including nonlinearity and hysteresis, is at the level of 1.2 K. In order to improve the properties of diode temperature sensors, two integrated diodes fed from two different current sources  $I_1$  and  $I_2$  should be used (Fig. 9). The difference in voltage drop over these diodes is

$$\Delta V = V_1 - V_2 = \frac{kT}{nq} \ln \frac{I_1}{I_2} \quad (13)$$

and does not depend on the reverse saturation currents because of their similarity due to the integration. Furthermore, the output voltage bias is significantly reduced, which leads to simpler measuring circuits.

Transistors are often used instead of diodes, and the base-emitter voltage difference is the output signal of the sensor. Integrated circuit (IC) technology allows us not only to produce temperature-sensitive pairs of transistors but also to include amplifiers and signal conditioning circuits on the same chip. In this way, integrated sensors with precisely trimmed output signal can be produced. The most popular are IC temperature sensors with  $1 \mu\text{A}/\text{K}$  output (e.g., Analog Devices AD592), but sensors with voltage output of 1.5 mV/K or even 10 mV/K are manufactured too. The LM75 temperature IC sensor, produced by National Semiconductor, has a silicon band gap sensing element and is equipped with a sigma-delta A/D converter,

programmable alarms, and a two-wire  $I^2C$  interface. The operating temperature span of diode- or transistor-based IC sensors ranges from  $-55^\circ\text{C}$  to  $125^\circ\text{C}$ .

### THE MEASUREMENT OF SENSOR RESISTANCE

#### Common Problems

The first stage of every transducer consists of the measuring circuit, directly connected to a temperature sensor. A few different measuring circuits are used to transform the resistance changes  $R_s(T)$  to the output signal  $V$ .  $R_s(T)$  forms that part of the sensor resistance which changes its value according to the measured temperature:  $R_s(T) = R(T) - R_0$ , where  $R_0$  is a constant value of the sensor resistance corresponding to the reference temperature. Three problems seem to be common for all these circuits. These are (1) sensor self-heating, (2) lead resistance, and (3) linearity.

Self-heating of resistance sensors is unavoidable because the flow of the current creating the output signal causes automatic heat dissipation in the sensor, subsequent increase of its temperature, and consequent measurement error  $\Delta t$ .

$$\Delta t = P/k_s = I^2 R_s / k_s \quad (14)$$

where  $k_s$  is a dissipation factor. The dissipation factor depends on the design of the sensor and on its materials, dimensions, and shape, but it depends primarily on the environment of the sensor. Its magnitude changes dramatically with the kind of medium surrounding the sensor and with the velocity of that medium, as presented in Table I. Theoretically, the error due to self-heating can be removed from the measurement result by means of a correction procedure, but our knowledge about the value  $k_s$  is insufficient to calculate the correction value because of the instability of the environmental conditions. It is sometimes possible to correct for self-heating effects by measurement at two currents and extrapolating to zero current. The best way, however, is to limit the error due to self-heating

Table I. Dissipation Factors of Resistance Sensors Without Protective Sheath

| Sensor                   | Environment | $k_s$ (mW/K) |
|--------------------------|-------------|--------------|
| Wire-wound RTD*          | Still air   | 3-5          |
|                          | Air 1 m/s   | 10-20        |
| Thin-film RTD            | Still air   | 2            |
|                          | Still water | 75           |
| NTC bead-type thermistor | Still air   | 1            |
|                          | Stirred oil | 6            |

\* RTD, resistance temperature detector

† NTC, negative temperature coefficient

by keeping the current at the allowable level, but it results in lowering of the output signal.

The second problem is the change of lead resistances with temperature. The problem becomes serious when the distance between the sensor and the transducer reaches up to hundreds of meters and the long leads are exposed to great temperature differences (e.g., outdoor temperatures in summer and winter). For a 10  $\Omega$  copper lead the temperature change of 60 K (from 30 °C to 130 °C) causes a 2.4  $\Omega$  resistance change which is observed as a 6 K temperature error if a Pt 100 sensor is used.

The best way to avoid this kind of error is to feed the sensor from a current source by one pair of leads and to sense the voltage from the sensor by another pair of leads. This solution is called a four-wire line and is commonly used in transducers with standard analog 4–20 mA output and in all high accuracy transducers. A three-wire line instead of a four-wire line is also used, especially in bridgelike measuring circuits. Three-wire installation cancels the additive errors caused by the thermal change of lead resistance, but the multiplicative part of the error remains. The higher the sensor resistance, the lower the influence of the lead resistance. There is no need to use four- or three-wire lines for thermistors or KTY sensors.

The third problem, linearity, is common for all transducers working with more or less nonlinear sensors. While most transducers are equipped with microprocessor-controlled systems, the linearity corrections are commonly performed numerically. The look-up table method is preferred. In that method, appropriate corrected values or the values of corrections which have to be added to the directly measured uncorrected results are written in memory. The linearization algorithm consists of a simple readout from the memory. At 0.1% resolution the method requires only 1 kB of memory. Some other methods of numerical linearity correction, utilizing the reduced tables containing only node point correction values, are also used. The correction data for all the results falling in between the node points are calculated by linear interpolation.

When the numerical linearization is performed together with other numerical corrections realized by the microprocessor system, the procedure becomes complicated and needs multi-dimensional tables. In that case the possibility of analog linearization ought to be considered. Two extremely simple methods may be used for that purpose. The methods are different in their principles, but both lead to the same results because of the same mathematical model.

The first method is based on the use of voltage-controlled current sources in the structure presented in Table 2, part E. The supply currents depend linearly on the output voltage ( $= I_{sp} + \alpha V$ ). Linearization is performed by the proper selection of the coefficient  $\alpha$  according to the sensor nonlinearity. The relation between  $V$  and  $R_s$  is then nonlinear and enables a successful linearization of NiCr thermistors and nickel RTD sensors while choosing a positive  $\alpha$  coefficient and of platinum RTD sensors while choosing it negative. The magnitude of  $\alpha$  coefficient depends on the kind of the sensor and on the temperature range.

The second method is even simpler, because it does not require the voltage-controlled sources. It depends on the proper choice of the Thevenin equivalent resistance  $R_{th}$ , seen from the sensor terminals, that is from the virtual terminals of the effective part of the sensor  $R_s(t)$ . According to such a definition of  $R_{th}$ , the  $R_s$  component of the sensor belongs to  $R_{th}$ . The method is based on the theorem that any voltage between two points of

each linear electrical circuit containing one variable parameter  $R_s(t)$  may be presented in a form

$$U = \frac{U_0 R_{th} - U_\infty R_s(t)}{R_{th} + R_s(t)} \quad (15)$$

where  $U_0$  is the voltage at  $R_s = 0$ ,  $U_\infty$  is the voltage at the open circuit  $R_s = \infty$ . Some examples of linearizing circuits are presented in Table 2, A to D.

The above described methods do not lead to the canceling of the nonlinearity but only to its reduction [11]. In practice more complicated circuits are used in integrated temperature measuring transducers like AD1841 or B-B IXR 100, but the idea of linearization procedure depends on the above given principles.

### The Most Popular Measuring Circuits for Resistance Temperature Sensors

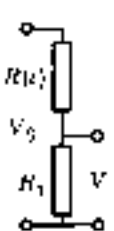
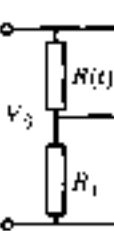
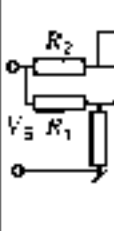
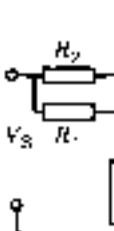
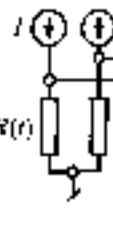
Generally the resistance sensors are manufactured with high accuracy. Transducers have to be matched to the sensors in order not to increase the total uncertainty. It is evident that for SPRTs the most accurate methods and instruments should be used for measuring the resistance and more exactly the resistance ratio. Costs and compactness are less important. The uncertainty of modern precise resistance ratio measurements is as low as a few parts per million, but only with very special apparatus used in advanced well-equipped laboratories. Such measurement circuits will be not presented here. However, with conventional temperature measuring transducers the accuracy of resistance measurements has to be high too. Let us note that according to IEC 751 standard, the Pt-100 uncertainty at 0 °C is only 10.1 K, which means  $\pm 0.04 \Omega$ . In order to protect the sensor accuracy, the uncertainty of the transducer might to be less than, say,  $\pm 0.01 \Omega$ , which gives 0.01% with respect to a 100  $\Omega$  sensor. For resistance-measuring instruments in common use, it is a rather high requirement and a bridge circuit is therefore the one which has to be primarily considered as the input stage of the transducer.

Balanced bridges are contemporarily used almost only in self-balancing chart recorders or  $x$ - $y$ - $t$  recorders. The complicated mechanical design of such instruments together with the need for the precise potentiometer, makes these instruments rather expensive. Such instruments have been formerly widely used not only in industrial applications, but also in laboratories and research. High cost and the absence of an electrical output signal (which may eventually be obtained from an additional potentiometer) make these instruments not very suitable for modern instrumentation systems.

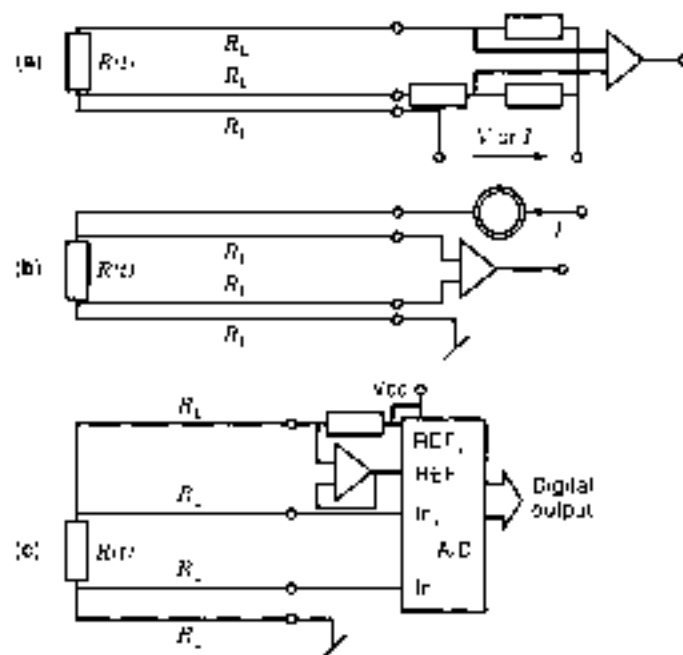
To the contrary, unbalanced bridges are very often used as the first stages of contemporary transducers working with resistance temperature sensors (Fig. 10x4). The differential structure of any unbalanced bridge circuit enables easy adjustment to the desired temperature range. The output voltage is not strictly proportional to the resistance, because the unbalanced bridge belongs to the class of the circuits described by Eq. (15) and presented in Table 2 as cases B and C. Therefore, an unbalanced bridge may also be used as a linearizing circuit for some types of sensors. To do that, an appropriate matching of the bridge branches have to be performed. Unbalanced bridges are supplied either from a voltage source or from current sources. The constant current supply is preferred



**Table 2.** Some simplified examples of linearizing circuits for resistance sensors. Circuit A (potentiometric), B (bridge circuit) and C (active bridge circuit) are for NTC thermistors or RTD nickel sensors. Circuit D (positive feedback circuit) is for KTD platinum sensor. Circuit E (with controlled current) can be used for all sensors regarding the proper sign of a coefficient. The balance state of bridge circuits is assumed at  $R_x(t) = 0$ , ( $R_1 = R_2 = R_3 = R_0$ ).

|   |   |   |   |   |
|---|---|---|---|---|
|  |  |  |               |  |
| $R_x(t) = R_0 \cdot R_2(t)$ $R_{T0} = R_0 + R_1$                                  |   |   |   |   |
| $V = \frac{V_0 R_1}{R_1 + R_0 + R_x(t)}$  | $V = \frac{V_0}{2} \frac{R_x(t)}{2R_0 + R_x(t)}$                                  | $V = V_0 \frac{R_x(t)}{2R_0 + R_x(t)}$  | $V = V_0 \frac{R_1 R_2 R_4 + (R_1 R_3 - R_2 R_4) R(t)}{R_1 R_2 R_4 + (R_1 R_3 - R_2 R_4) R(t)}$ | $V = I_0 (1 + \alpha V)$  |
|   |   |   |   | $V = \frac{I_0 R_x}{1 + \alpha I_0 R_x}$  |

especially for low-resistance sensors, as Pt-100 or Ni-100, where the three-wire connection between the sensor and the transducer is needed in order to reduce the line temperature error. The reduction is twice better using a current source than using a voltage source.



**Figure 10.** The measuring circuits which reduce the influence of the leads resistances. (a) Three-wire bridge circuit makes it possible to connect one lead to the sensor and the second to the resistor in opposite bridge arm and then to reduce their influence. (b) Four-wire circuit with the current source enables canceling of the lead resistance influence. (c) The implementation of an A/D converter in the four-wire circuit provides direct conversion of the analog signal to the digital one.

The output voltage from a bridge is led to a direct-current (dc) differential amplifier. The signal is usually high enough for a conventional low noise operational amplifier with a proper compensation of bias currents. In some extremely precise instruments the switched-capacitor-based instrumentation amplifiers are used (i.e., Linear Technology LTC 1043). The aim of the amplifier is not only to increase the signal but also to allow the transition from differential to a single-ended signal. It is a general requirement, especially in the industrial measuring system that the sensor must be grounded. When the three-wire configuration is used, two of them are close to ground potential, but according to the principles of noise protection they must not to be grounded at two points (at sensor's side and at transducer's side) in order to avoid the ground loop which introduces additional unknown voltages. The circuit with floating voltage supply and grounded one-amplifier input is less convenient because of the limitations in scaling the circuit parameters. The greatest comfort in circuit parameters scaling is provided by a four-wire installation because it consists of two almost separated circuits (Fig. 10b). The only problem to solve is the subtraction of that part of voltage which corresponds to the low limit of the measured temperature. It may be done either by a bias voltage or by another differential structure containing a sufficiently stable voltage source. Integrated circuits, which incorporate a controlled gain amplifier, a linearization circuit, and isolated output (i.e., Analog Devices LB41), facilitate the design of the measuring system. A/D converters with reference input may be used for direct four-wire connection to the sensor supplied from the voltage source instead of a more complicated current source (Fig. 10c).

Some completely different temperature measuring circuits—that is, circuits with frequency output, where the sensor resistance influences either the oscillator frequency or the duty cycle of square-wave output voltage—are also known. The practical implementation of such circuits are limited mostly to those in a form of integrated circuits—for example, the SMT 160-30 produced by Smartec.

## RESISTIVE SENSORS FOR LOW TEMPERATURE RANGES

The range of temperatures below 20 K becomes more and more interesting not only for the researchers but also for the technologists. The practical use of the superconductivity requires the precise temperature measurements and control of temperatures as low as 4 K. In some cryogenic technologies the high magnetic fields and nuclear radiation are simultaneously present. Temperature sensors destined for low-temperature applications have to be resistant to those environmental conditions too. It is reasonable to distinguish a special group of sensors working at low temperatures in spite of their different principles of operation and design.

As stated before, a platinum resistance thermometer does not work properly at temperatures below 10 K. For that range a different alloy has been developed, namely rhodium with 0.5% iron (2,3). The technology of preparing this, 0.05 mm diameter rhodium-iron wires is complicated. It includes chemical iron deposition on powdered rhodium and then a series of metallurgical processes. The helically wound sensor is hermetically encapsulated similarly to SPRT sensors. The most useful operating range is 0.3 K to 30 K; but due to its relatively low slope of resistance versus temperature, it may be used up to the normal ambient temperatures too. The stability of an Rh-Fe sensor is relatively good, much better than that of low-temperature semiconductor sensors. Semiconductor sensors, however, are much simpler and smaller, and for that reason they are used too.

Some specially prepared and composed thermistors, usually made from iron oxide, are able to measure temperatures as low as 5 K. According to Eq. (5), thermistor sensitivity and nonlinearity increases dramatically at lower temperatures, creating problems with covering a wider range of temperatures. This is a common problem of all low-temperature sensors related to the "wring" representation of temperatures on the linear scale, as discussed in the first section of this article. The greatest advantage of thermistors is their low sensitivity to the magnetic fields. Germanium was previously used in electronic semiconductor devices to the same degree as silicon. The physical principles of its conductivity remain the same as those described for silicon. Germanium temperature sensors have been used for measurement purposes for a much longer time than bulk silicon sensors. Their operating area, however, is shifted toward the very low, cryogenic temperature range. The bulk germanium with a very small amount of added impurities forms a low-temperature sensor which may be used down to 1.6 K, but due to the very strong dependence of its properties on the amount of the impurities introduced, the individual calibration of each sensor is necessary. The calibration process at extremely low temperatures is always a complicated and expensive one. Other problems related to the use of the germanium sensor are (1) rapid and unexpected changes in resistance of the magnitude corresponding to a few millikelvins, (2) high sensitivity to the mechanical stress, and (3) a strong Peltier effect causing an additional heat distribution along the sensor. The instability exhibited by many low-temperature Ge resistance thermometers is thought to be due to the difficulty of attaching leads in a way that defines the resistor without subjecting it to strain effects. A long-lasting experience with this kind of sensor (since 1960) has resulted in gathering a considerable amount of information, which enables us to reduce the effects of all inconveniences listed above. Specially doped germanium resistors are insensitive to magnetic fields (4).

Individually calibrated diode sensors may also be used in the very low temperature region, down to 10 K. Sensitivity is not the same as at medium temperatures, and it increases rapidly below a certain temperature (approximately 25 K for silicon diodes), but the sensor calibration curve remains repeatable with the uncertainty of  $\pm 10$  mK. Commercially available diode sensors are produced with a wider uncertainty span, exceeding  $\pm 0.25$  K but with quite good reproducibility of  $\pm 50$  mK (4).

## THERMOCOUPLE SENSORS

## Physical Principles

A temperature difference between two points of a conducting wire forces free electron diffusion from the point of higher temperature to the point of lower temperature. Such dislocation of electrons produces a voltage difference, which furms the electron flow in the opposite direction. In the state of dynamic equilibrium, both processes are in balance. A voltage difference caused by the temperature difference is known as thermal electromotive force (emf), and it provides a measure of temperature difference between any two points of the wire. Thermal conductivity of the metal wire causes temperature distribution along the wire, and hence the thermal emf may be considered as continuously distributed along the wire too. The problem of how to measure the thermal emf arises because each electrical contact of the connecting leads with the heated wire is also a thermal contact and generates subsequent thermal emf corresponding to the temperature difference at the ends of the connecting leads. If the materials of the heated wire and connecting leads are the same, two emfs appear in the loop with the opposite signs and are reduced to zero. However, the magnitude of the thermal emf depends on the material; and it is possible to find a pair of materials (A and B) with essentially different emfs, thereby generating relatively high voltages in the loop (Fig. 11).

By denoting the thermal emf sensitivities to the temperature of both wires as  $S_{A}(t)$  and  $S_{B}(t)$ , one obtains

$$E_{AB} = \int_{t_1}^{t_2} [S_{A}(t) - S_{B}(t)] dt = \int_{t_1}^{t_2} S_{AB}(t) dt \quad (16)$$

where  $S_{AB}(t)$  is the sensitivity of a thermocouple known as a Seebeck coefficient or "thermoelectric power." In other words, a temperature difference produces thermal emf, and the inhomogeneity of the materials in the loop allows us to detect it. The greatest inhomogeneity appears at the contact points of two different materials. These points constitute "hot junction" and "cold junction," or more properly, a measuring junction and reference junction. Any other material (C in Fig. 11) included in the thermocouple loop but not exposed to the temperature difference does not produce any additional emf in the loop. Otherwise, each unexpected inhomogeneity, caused not only by inclusion of any additional material D but also by chemical reactions or annealing processes in the material under temperature difference, is a source of additional thermal emf. Therefore the whole of the thermocouple loop ought to be considered as the sensor not just the tip and this makes the consideration of sources of uncertainty much different to most other temperature sensors.

The thermal emf effect discovered by Seebeck in 1821 is superposed by two other effects related to the current flow in the

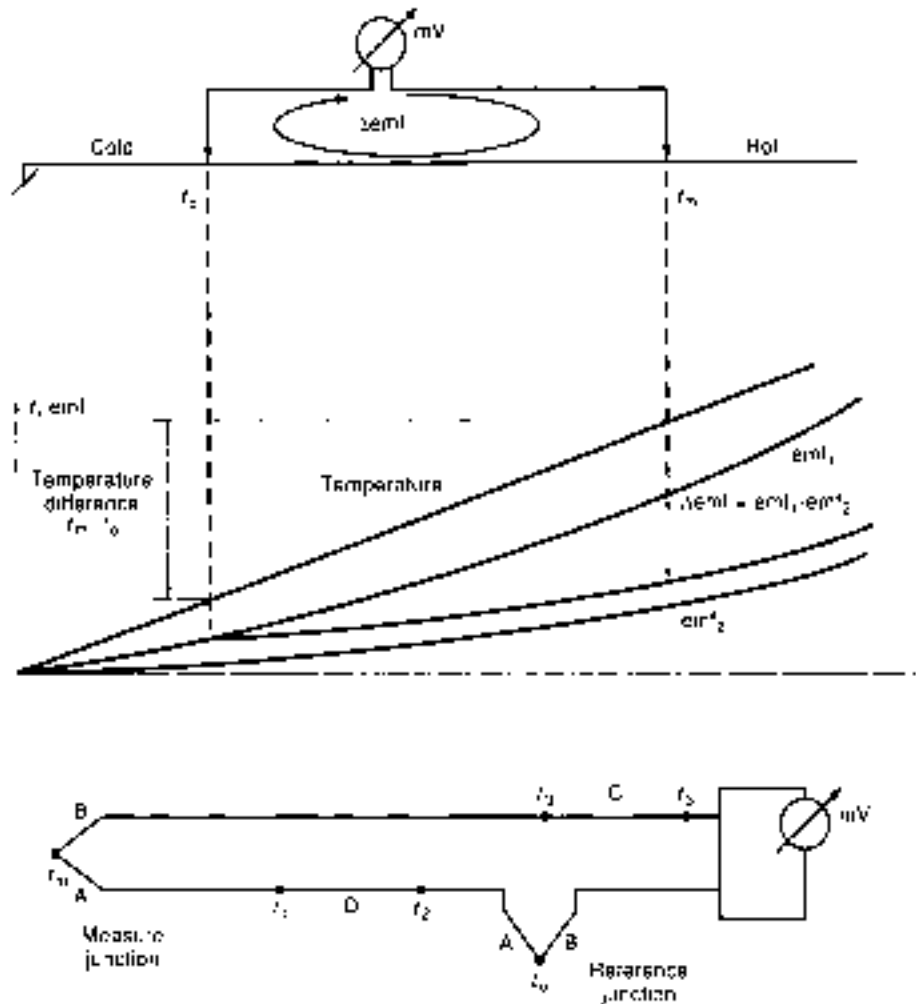


Figure 11. Temperature difference measurement by a thermocouple circuit. Temperature difference  $t_m - t_0$  corresponds to the emf difference  $\Delta emf$ .

loop: (1) Thomson effect and (2) Peltier effect. In the Thomson effect an additional emf is induced by current flow in the presence of a temperature difference in the conductor, involving heat liberation or absorption by the conductor at the rate:

$$Q = \int_{t_1}^{t_2} S_T i dt \quad (17)$$

where  $S_T$  is a Thomson coefficient of the particular material  $A$ . For the same reason as with Seebeck effect, the Thomson emf in the whole loop is different from zero only in a case when two different materials make two branches of the loop. The Peltier effect reveals in the additional heat absorption or heat generation forced by the current flowing through a junction of different materials  $Q_P = V_{PAB} i$ , where  $V_P$  is called the Peltier coefficient (Peltier emf).

Three effects described above are interrelated according to the equation

$$E_{AB} = V_{PAB}(t_0) - V_{PAB}(t_m) + \int_{t_1}^{t_2} S_{W1} i dt + \int_{t_2}^{t_1} S_{B1} i dt \quad (18)$$

Because of the low currents flowing in the temperature-measuring thermocouple loops, the effects of additional heat emitted in the Thomson and Peltier effects may be normally neglected.

### Thermocouples

It is evident that a thermocouple has to be composed of thermowires, with very different thermal emfs. In measurement practice, however, a lot of additional requirements are of a great importance. Most of them stem from the fact that thermocouples are usually used at high temperatures. Therefore the wires themselves have to be resistant to high temperatures. Their melting points must be at least by 150 K higher than the upper temperature of the measuring range. At high temperatures the atmosphere is usually corrosive and hence much more aggressive than at low temperatures. The degeneration of the lattice structure and the chemical changes observed at the surface of the wires cause the inhomogeneity of the wire materials and lead to the successively arising measurement errors. This is a reason why thermocouple wires are rather thick, and the upper temperature limit for continuous work case depends on their diameters. The cost of wires, especially those made of noble metals, is important too. Very special alloys have been developed in order to meet the above-mentioned requirements. The work in this area is still going on, and the result has been continuous improvements and modifications of the thermocouple wires. Table 3 presents the most popular thermocouples and their essential properties.

The calibration curves of some thermocouples are subject to standardization in the form of reference tables, similar to the corresponding tables for resistance sensors. Worse stability of

Table 3. Thermocouple Data

| Type | + Wire          | - Wire             | Temp. Range<br>Temperature<br>Span<br>(°C) | Short Time<br>Allowable<br>Temperature<br>(°C) | Thermal emf<br>Span at 0°C<br>Reference<br>Junction<br>(mV) | Allowable Uncertainty (IEC) |  |                             |  | Properties  |
|------|-----------------|--------------------|--|--|---|-----------------------------|--|-----------------------------|--|---|
|      |                 |                    |  |  |   | Class 1<br>at<br>0°C<br>(%) | Class 1<br>at<br>Upper<br>Limit<br>(%) | Class 2<br>at<br>0°C<br>(%) | Class 2<br>at<br>Upper<br>Limit<br>(%) |   |
| S    | PtRh10          | Pt                 | 0 to +1600                                 | 1760   | 0.23 to -19.6   | 1                           | 2.5                                    | 1.5                         | 4                                      | Stable. For oxidizing and inert atmospheres. Sensitive to contamination.                            |
| R    | PtRh13          | Pt                 | 0 to +1600                                 | 1760   | 0.25 to -21   | 1                           | 2.5                                    | 1.5                         | 4                                      |   |
| B    | PtRh30          | PtRh5              | +100 to +1600                              | 1900   | - to +13.8  |                             |  |                             | 5/27                                   | Negligible thermal emf up to 60°C.  |
| K    | NiCr<br>Chromel | NiAl<br>Alumel     | -270 to +1000                              | 1250   | -6.46 to -50.6  | 1.5                         | 4                                      | 2.5                         | 9                                      | Most frequently used. Linear for $t > 0^\circ\text{C}$ . For clean oxidizing and inert atmospheres. |
| N    | NiCrSi          | NiSi               | 270 to +1000                               | 1300   | 4.3 to +47.5  | 1.5                         | 4                                      | 2.5                         | 9                                      | More stable than type K at high temperatures.   |
| T    | Cu              | CuNi<br>Constantan | 270 to +350                                | 400  | 6.26 to -20.9   | 0.5                         | 1.4                                    | 1                           | 2.6                                    | For oxidizing and reducing atmosphere at low temperatures. Moisture-resistant.                      |
| J    | Fe              | CuNi               | 270 to +750                                | 1200   | 6.1 to +69.5  | 1.5                         | 3                                      | 2.5                         | 6.6                                    | For oxidizing and inert atmospheres.  |
| E    | NiCr            | CuNi               | 270 to +700                                | 1900   | 9.6 to +75.4  | 1.5                         | 3.2                                    | 2.5                         | 6.75                                   | High thermal emf, also at low temperatures.   |
| ...  | WRe5            | WRe26              | 0 to +2400                                 | 2700   | 0 to +40.7  |                             |  |                             |  | For highest temperatures. For neutral and reducing atmospheres only.                                |
| —    | NiCr            | Al                 | -273 to 0                                  | -  | -5.3 to 0   |                             |  |                             |  | Linear for low temperatures, also below 4.2 K.  |

Thermocouples results in their much greater uncertainties as compared to resistance sensors (see Fig. 5). When a lower uncertainty is required, the individual calibration of thermocouples is not recommended because the validity of the results is rather short-lived.

### Thermocouple Sensors

A great variety of thermocouple sensors with different sizes and designs are destined for a wide range of applications, such as power plants, nuclear plants, metallurgy, chemical reactors, and the glass industry, as well as laboratories, research works, and communal applications (5).

Industrial-immersion-type thermocouple sensors look like thick resistance sensors because they are manufactured in a similar form, with a long protective pipe and a head on one end. The sensor element (wires with the junction) are mounted inside and may be replaced if necessary. However, the materials used for shields differ considerably. For high-temperature sensors a conventional carbon steel protective pipe is insufficient. Either stainless steel (18% Cr, 8% Ni), Inconel (NiCr, 15% Fe), hastelloy, or bronze has to be used depending on the environmental conditions. Sometimes there is a need for molybdenum or tungsten sheath (for highest temperatures). Noble metal thermocouples have to be additionally protected by means of an internal ceramic (alumina  $\text{Al}_2\text{O}_3$ ) coating, against the contamination of the thermocouple wires by the particles of pipe material that occurs at high temperatures. Some outer porous ceramic protection tubes are used with the sensors for open fire furnaces.

A special type of sensor is produced in the form of a metal shielded double thermowire cable with  $\text{MgO}$  or  $\text{Al}_2\text{O}_3$  insulation. These are called shielded thermocouples or mineral insulated metal sheathed (MIMS) sensors. The same type of insula-

tion is used in resistance heaters. The thermocouple junction is formed by connecting the both wires. The external diameter of the MIMS may be as low as 0.25 mm (more commonly 0.6 mm to 3 mm), and the bonding radius allowed is normally twice that of the diameter. This constitutes a great advantage of the sensor, being an ability to penetrate hardly accessible spots. This kind of sensor is now obtainable in lengths of up to a tenth of a meter, with the sensing junction, as well as the plug on the opposite end, formed and sealed by the manufacturer. Former MIMS were produced in a form simply cut from one piece of cable, but the hygroscopic properties of the insulation made it very hard for the user to provide the proper hermetic sealing.

The next group of thermocouple sensors are those designated for the measurements of moving surface temperatures and designed as free thermowires or thermostrips, either suspended on elastic arms or shaped into a form of an elastic arch. The measuring junction is situated in the middle of the free part of the thermostrips and should be pressed to the surface during measurement. The smoothness of the thermojunction allows the measurement of the moving or rotating elements without heat generating by friction. The elasticity of the sensor ensures a good thermal contact with the surfaces of different shapes (e.g., with rollers of different diameters).

In metallurgy, two kinds of immersion-type sensors are commonly used for measurement of the molten metals temperature. Both kinds work under the transient state conditions. The construction must be strong enough to pierce the layer of the blast furnace slag. In the first design, two sharp cut thick bars from thermocouple materials are placed near each other at the end of a long handle. The stick is immersed in the molten metal, thereby creating a junction. In the second design the exchangeable cap with very thin thermocouple wires (50  $\mu\text{m}$  to 100  $\mu\text{m}$ ) and a protective cover is placed at the end of the handle. The

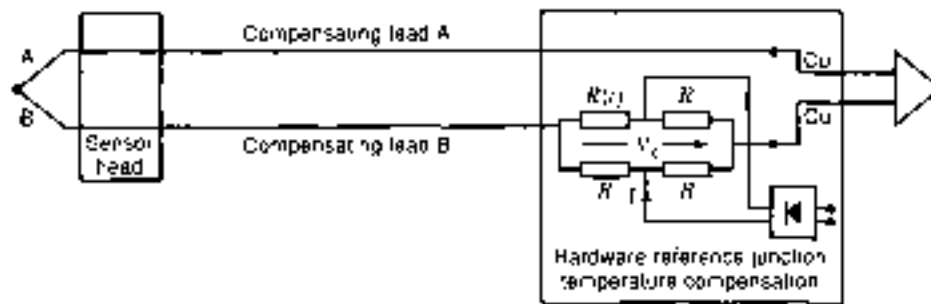


Figure 12. Basic thermocouple circuit with compensating leads which eliminate the influence of the lead temperature variations and with reference junction temperature compensation circuit. Resistances  $R$  and supply current  $I$  are matched according to the type of the thermocouple.

cover is damaged when immersed in the molten metal, and after each measurement the cup is replaced by the new one.

Many other thermocouple sensors are present on the market or are custom-designed for particular purposes. The use of a thermocouple sensor instead of another sensor type is recommended in a case when most important are (1) high temperature range, (2) the small dimensions of the sensor, and (3) relatively low cost.

#### Thermocouple Measuring Circuits

At a first glance, the task seems to be easy: Create a reference junction and measure a dc voltage. However, some problems arise particularly in the industrial environment and at high temperatures. In the large area plants the distance between the sensors and the transducers is long, and at high temperatures the sensor head temperature is usually unstable to such a degree that it is impossible to treat it as a reference junction. Therefore, the thermocouple wires have to be brought to the place where the temperature is constant, say in the transducer (6.7). For evident reasons it is not a good solution, especially if noble metal wires are used. In such a case, extension wires are used as a connection between the sensor head and the reference junction. These are special compensation leads having the same thermal emf as the thermocouple wires, but with a much lower temperature range, namely that expected to occur at the sensor head (Fig. 12). Compensation leads have to be matched to the thermocouple; and in order to avoid misconnections, the colors of their insulation are subject to standardization. Compensating wires are much cheaper than thermocouple wires. A special noble metal thermocouple has been developed (Type B, Table 3), which does not require any compensation leads because its thermal emf at temperatures up to 50°C is practically equal to zero and with temperatures up to 120°C it is very low. For that thermocouple, neither a reference junction nor the compensating leads are needed, assuming that the ambient temperature of the transducer and the sensor head temperature do not exceed 50°C and 120°C, respectively.

For all other thermocouples, however, the reference junction is necessary. In laboratory practice ice-water baths, and in industrial measurements, thermostats may be used. Both are unpractical. Instead of stabilizing the temperature of a reference junction, it is more convenient to measure it and to introduce a compensating voltage into an emf measurement loop. Such a method is now used in almost all instruments and transducers. The most common compensating circuit is shown in Fig. 12. At nominal temperature of the reference junction (say 25°C),  $R_{RT} = R$  and  $V_c = 0$ . As the reference temperature increases,  $R_{RT}$  increases accordingly, producing a compensating voltage  $V_c$ , equal to the change of the thermal emf corresponding to the dif-

ference between the actual temperature and 25°C. The supply current  $I$  is matched according to the sensitivity of particular thermocouple. There exist also a great number of integrated circuits for compensation of the reference junction temperature where a diode sensor is used instead of the temperature-sensitive resistor (i.e., Linear Technology 1025). The amplification of a thermocouple signal, together with the reference junction compensation and with some additional functions, is performed by integrated circuits such as Analog Devices AD594, or Linear Technology LTK001.

A method of reference junction voltage correction instead of compensation is also used. It is based on the measurement of the reference junction temperature (for example, by means of a semiconductor sensor), followed by a numerical calculation of the correction value. There is also a common need for the numerical correction of a result in all instruments working with thermocouples, because of the nonlinearity of these sensors. The correction is usually performed by the look-up table method described before.

Another problem caused by the noise superposing on a relatively weak dc signal transmitted over long compensating leads. In order to avoid the electromagnetically induced voltages, the wires in the compensating leads should be twisted. The protection against the common mode noise is provided by shielding the wires and connecting the shield to a guard terminal of the instrument or transducer. In this way the current flowing through the stray capacitance between the leads and the supply power lines or induced by any source in the grounding loop is shunted and does not affect the measured voltage. The noise voltage may also be suppressed by filtering of the output signal.

A very successful method eliminating all problems due to long compensating leads is to place the whole transducer in the thermometer's head. The current developments in electronic components technology enables design of the compact and temperature-resistant transducers comprising all compensating and linearizing elements and delivering the standard 4–20 mA output signal. Many companies offer such a solution now, and this design seems to be very promising for all immersion-type thermocouple thermometers.

#### QUARTZ TEMPERATURE SENSORS

The piezoelectric properties of quartz crystals ( $\text{SiO}_2$ ) are applied in the design of extremely precise, stable, and relatively cheap oscillators. The applications of these oscillators are very widespread, from counters and frequency meters to precise timers in clocks and computers. The most important requirement for all these purposes is temperature stability. It

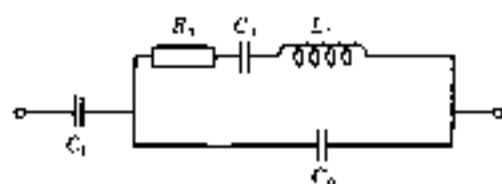


Figure 13. An equivalent circuit of a piezoelectric resonator.  $C_0$  is a geometrical capacity between two electrodes and  $L_1$ ,  $C_1$ ,  $R_1$  are quartz plate parameters which depend on its mechanical properties.  $C_1$  is a load capacity for tuning of the resonance frequency.

is achieved by appropriate cut of the oscillator plate from the quartz crystal. For temperature-invariant oscillators the so-called AT cut is used with the cutting plane inclined to a  $x$  axis (optical axis) of the crystal at  $-36^\circ$ . Any other cut results in a smaller or greater dependence of the oscillator frequency on the temperature. This very property is used in quartz temperature sensors. A plate obtained by a Y cut with the inclination of  $-4^\circ$  called HT cut has the highest value of temperature coefficient. The relation between the frequency and temperature may be written as

$$f(t) = f_0(t) + 90 \times 10^{-6} \Delta t + 60 \times 10^{-8} \Delta t^2 + 30 \times 10^{-10} \Delta t^3 \quad (19)$$

where  $f_0$  is frequency at temperature  $t = t_0$  and  $\Delta t = t - t_0$ . The third and the fourth terms in parentheses introduce non-linearity, which in this particular case is a great disadvantage because a conventional frequency meter cannot be used as a temperature-indicating instrument. Therefore, the LC cut, instead of the AT cut, is used for quartz temperature sensors. The LC cut of a quartz crystal with the cut plane inclined at  $+11^\circ 10'$  to the  $x$  axis and at  $+9^\circ 24'$  to the  $z$  axis forms an oscillator with frequency linearly depending on the temperature but with a lower sensitivity ( $35 \times 10^{-8}$  1/K instead of about  $100 \times 10^{-8}$  1/K with HT cut).

A quartz plate with two electrodes forms a resonator which may be presented in the simplest form as an equivalent electrical circuit, as shown in Fig. 13. In the circuit,  $C_0$  is a geometrical capacity between two electrodes and  $L_1$ ,  $C_1$ , and  $R_1$  are quartz plate parameters, which depend on its mechanical properties. Two resonance frequencies exist for this circuit: Serial resonance frequency  $f_s = 1/2\pi \sqrt{L_1 C_1}$  and parallel resonance frequency  $f_p = 1/2\pi \sqrt{L_1 C_E}$ , where  $C_E = C_0 C_1 / (C_0 + C_1)$ . Both frequencies have close values, because the capacities  $C_1$  and  $C_E$  are of the same order. Using of an additional load capacity  $C_L$ , the resonance frequency of the plate may be tuned in a limited range between  $f_s$  and  $f_p$ . Two opposite surfaces of the resonator are coated with thin gold layers. The resonator is mounted in a hermetical case protecting it from atmospheric

air moisture. Spring contact elements ensure a low level of mechanical damping. The oscillator frequency  $f_0$  depends on the dimensions of the resonator. It is very convenient to adjust that frequency to such a value that the relationship between the frequency and the temperature may be obtained by simply shifting a decimal point on the frequency meter. Hence if the sensitivity coefficient  $\alpha$  is equal to  $35 \times 10^{-8}$  1/K, the condition is fulfilled for  $f_0 \approx 28.6$  MHz, because in that case  $f_{0\alpha} = 1000$  Hz/K. The tuning feature of the oscillator allows us to meet the above requirement in spite of some differences in individual plates parameters (3). The connection between the sensor and the oscillator must be performed by high-frequency coaxial cable. By the use of frequency meters with high resolution, high resolution of temperature measurements may be achieved too. Much better solution, however, is application of a differential structure of the measuring circuit (Fig. 14) where a mixer forms a low-frequency signal  $\Delta f$ , which corresponds to the difference between the measured temperature  $t$  and a reference temperature  $t_0$ :  $\Delta f = f - f_0 = f_{0\alpha}(t - t_0)$ . In such a state the resolution of the frequency meter may be much more effectively used. Taking as an example the temperature range from 0°C to 200°C, the value of  $f_{0\alpha}$  equal to 1000 Hz/K, and the 61/2 digit resolution of the frequency meter (which is a common practice), a 0.1 mK temperature resolution is achieved. This extremely high resolution is a reason why quartz thermometers are commonly equipped with two sensors allowing the measurement of temperature difference. In many practical cases, it is not the absolute value of temperature but the difference of temperatures that has to be known with a great accuracy. A double quartz thermometer is an excellent instrument for this purpose. Please note that the meaning of the term "absolute value" is used here differently than "absolute temperature scale." The uncertainty of the quartz thermometers depends primarily on aging and relaxation. Single-point recalibration from time to time and avoidance rapid temperature shocks are therefore highly recommended. With these conditions met, the uncertainty of  $\pm 50$  mK may be sustained for a long time.

## RADIATION THERMOMETRY

The principle of radiation thermometry is the measurement of the thermal energy emitted by radiation from the object under inspection. It is not the whole energy that is measured but only its very little part, corresponding to the radiation focused on the radiation-sensitive element placed in the thermometer. The essential difference between the radiation thermometry and all other methods of temperature measurements is the lack of the thermal equilibrium between the object and the sensor. With radiation thermometry the only way of thermal energy transfer from the object to a sensor would be electromagnetic wave

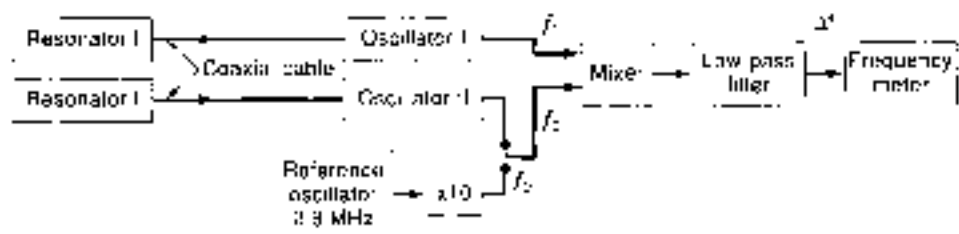
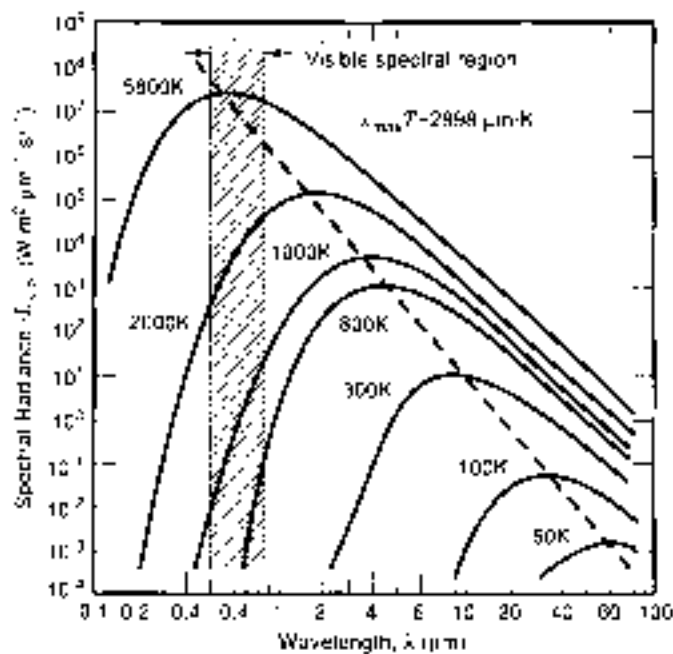


Figure 14. The differential structure of a quartz thermometer.



**Figure 16.** Spectral heat radiation density. Dotted line indicates the Wien's law. After Ref. 8. The radiance energy declines dramatically with the temperature. It is the main reason why low temperature radiation thermometers were developed many years later than high temperature ones.

propagation in the range from ultra-violet radiation (0.2  $\mu\text{m}$  wavelength) through visible spectrum (0.48  $\mu\text{m}$  to 0.78  $\mu\text{m}$ ) up to far-infrared radiation (50  $\mu\text{m}$ ). Thermal energy transfer does not tend to equalizing of the temperatures but only excites the sensor. The excitation level, and consequently the output signal of the sensor, depends on the portion of the received energy. Proper design, together with the proper use of a radiation thermometer, ensures a strictly defined dependence of that portion of energy and the temperature of the object. All other object properties and radiation parameters such as emissivity, dimensions, the distance to the sensor, atmosphere composition and temperature, and radiation of other objects including the thermometer body and many others, either have to be kept at the same level as during the thermometer calibration or must have no influence on the sensor. These requirements seem to be more serious than in conventional thermometers since they are related to the environment in which the measurement is performed, rather than to the sensor and the instrument themselves.

In order to answer the question of how to meet the above requirements, some essential properties of thermal radiation have to be considered. The spectral density of heat energy emitted by a black body is given by Planck's law

$$M(\lambda, T) = C_1 \lambda^{-5} \frac{1}{e^{C_2/\lambda T} - 1} \quad (20)$$

or in a simplified (but sufficient for our discussion) form of Wien's law

$$M(\lambda, T) = C_3 \lambda^{-5} e^{-C_4/\lambda T} \quad (21)$$

In both equations,  $C_1 = 37.4 \text{ mW } \mu\text{m}^{-2} / \text{m}^2$  and  $C_2 = 14,388 \text{ K} \cdot \mu\text{m}$ . The lower the temperature of the radiation source, the longer the wavelength of the emitted radiation (Fig. 15). The product of the temperature and the wavelength corresponding to the maximum of spectral density remains constant, according to Wien's law of displacement,  $T\lambda_{\text{max}} = 2898 \text{ K} \cdot \mu\text{m}$ . The measurement of low temperatures requires the use of the sensors and of the methods which detect the radiation in far-infrared spectrum.

The total thermal energy emitted by a surface of a black body with the area  $F$  is an integral of the Wien's equation:

$$E_{\text{th}} = F \int_0^\infty M(\lambda, T) d\lambda \quad (22)$$

The black body is defined as an object which does not reflect thermal radiation and therefore, according to Kirchhoff's law of radiation (absorptive power = emissive power), emits the whole radiant energy relative to its temperature. The emissivity factor  $\epsilon$  of a black body is equal to 1. To simulate a black body (e.g. for the calibration of radiation thermometer), a closed cavity with a relatively small aperture may be used. The inner surface of the cavity has to be specially shaped. In the radiation thermometry practice, only some objects (such as a hearth of a furnace) may be treated as black cavities. All other objects, and especially all objects in the air, have the emissivity  $\epsilon$  smaller than 1, and their radiation density has to be multiplied by  $\epsilon$ . The magnitude of emissivity depends on the material, surface finish (polishing, oxidization, roughening, deadening), its temperature, incident and viewed angles of heat flux direction, and polarization of radiation. Furthermore, the emissivity depends on the wavelength too. Therefore, the whole heat energy emitted by a uniform surface  $F$  observed by the radiation thermometer may be expressed as

$$K = F \int_0^\infty \epsilon(\lambda) M(\lambda, T) d\lambda \quad (23)$$

where  $\epsilon(\lambda)$  is usually known with a very poor accuracy.

Next we take into consideration the properties of the atmosphere that the radiation is passing through. Application of radiation thermometers with high-temperature objects is always disturbed by the presence of smoke and dust particles absorbing the radiation. The blow of purging air is used to clear the optical path between the object and the thermometer and to protect the optical parts from contamination with dust. Nevertheless, the outer surface of the instrument optics has to be cleaned from time to time.

In measurements performed in open air the concentration of absorbing gases is considerably lower, but the distance between the object and the instrument is much greater so the absorption plays a significant role too. The contents of CO, CO<sub>2</sub>, and water vapor in the air are most significant. The spectral distribution of the absorption caused by these gases is not uniform and shows extremely great variations. Only two bands of thermal radiation wavelength may be indicated as almost free from absorption. These are 3.5  $\mu\text{m}$  to 4.5  $\mu\text{m}$  (near atmospheric window) and 8  $\mu\text{m}$  to 13  $\mu\text{m}$  (far atmospheric window). Hot gases disturb the measurement process by their own radiation too. According to Kirchhoff's law, the spectral distribution of emissivity is the same as the distribution of absorption, and it declines in the spectral ranges of atmospheric window.

For these two reasons, it is highly recommended that radiation thermometers operate within one of those windows.

Considering the spectral disturbances, two additional aspects must not be overlooked: (1) the spectral transmittance of the materials used for optical parts of the thermometers (lenses, windows, and filters) and (2) the spectral sensitivity of radiation detectors. Glass optic elements may be used in the range of visible and near-infrared radiation up to 2  $\mu\text{m}$  wavelength that corresponds to measured temperatures over above 500 C. Quartz lenses enable slight widening of that range (up to 3.5  $\mu\text{m}$ ). For these thermometers, however, which cover a much lower temperature range, very special materials have to be used. These are ZnSe (0.6  $\mu\text{m}$  to 16  $\mu\text{m}$ ), GaAs (1.2  $\mu\text{m}$  to 12  $\mu\text{m}$ ), or CdTe (1.7  $\mu\text{m}$  to 25  $\mu\text{m}$ ) and some other materials, such as Chalcogenid, KRS-5, or specially prepared ZnS. All of them (except ZnSe) are not transparent to visible light, and therefore the optical path for visual radiation must be simultaneously used to aim at the target.

### Infrared Radiation Detectors

Two groups of infrared (IR) detectors are presently used in radiation thermometry. These are thermal detectors with low but spectrum-independent sensitivity and semiconductor photon detectors (IR diodes and IR photovoltaic sensors), much more sensitive but working in the limited spectral zones. Thermopiles, bolometers, and pyroelectric detectors are the thermal detectors.

Thermopiles consists of a large number (of up to 66) of thermocouples with hot junctions concentrated on a small surface and exposed to the thermal radiation flux, along with reference junctions kept at a temperature close to the ambient temperature (Fig. 16). The thermocouple materials are Bi-Sb or Ag-

poly Si, with a hot junction deposited on a very thin (0.7  $\mu\text{m}$  to 1  $\mu\text{m}$ ) silicon membrane isolated with an  $\text{SiO}_2$  layer and with a reference junction deposited on bulk silicon material, which forms a frame around the membrane. A low thermal conductivity of the thin membrane secures proper thermal insulation between hot and reference junctions. The detector is fabricated by Micro Systems Technology (MST) and may be integrated with signal conditioning elements or even A/D converters on one chip with dimensions not exceeding a few millimeters. These thermocouples possess a very high thermoelectric power of approximately 100  $\mu\text{V}/\text{K}$ , comparing with a few  $\mu\text{V}/\text{K}$  for conventional metal thermocouples. This ensures high sensitivity of the detector expressed in volts per watt of thermal energy.

$$S_{\text{th}} = \frac{V_{\text{out}}}{E} \quad (24)$$

Thermal noise is the factor limiting the possibilities of thermal energy measurement. The Johnson noise equivalent voltage is given by

$$V_N = 2\sqrt{kTR\Delta f} \quad (25)$$

where  $k$  is Boltzmann's constant,  $R$  is the detector resistance,  $T$  is the detector temperature, and  $\Delta f$  is the bandwidth of the associated amplifier, determined either by the chopper frequency or by the detector speed. Substituting Eq. (24) in Eq. (25), one achieves the noise equivalent power (NEP) in the form

$$\text{NEP} = E_{\text{Nmin}} = \frac{2\sqrt{kTR\Delta f}}{S_{\text{th}}} \quad (26)$$

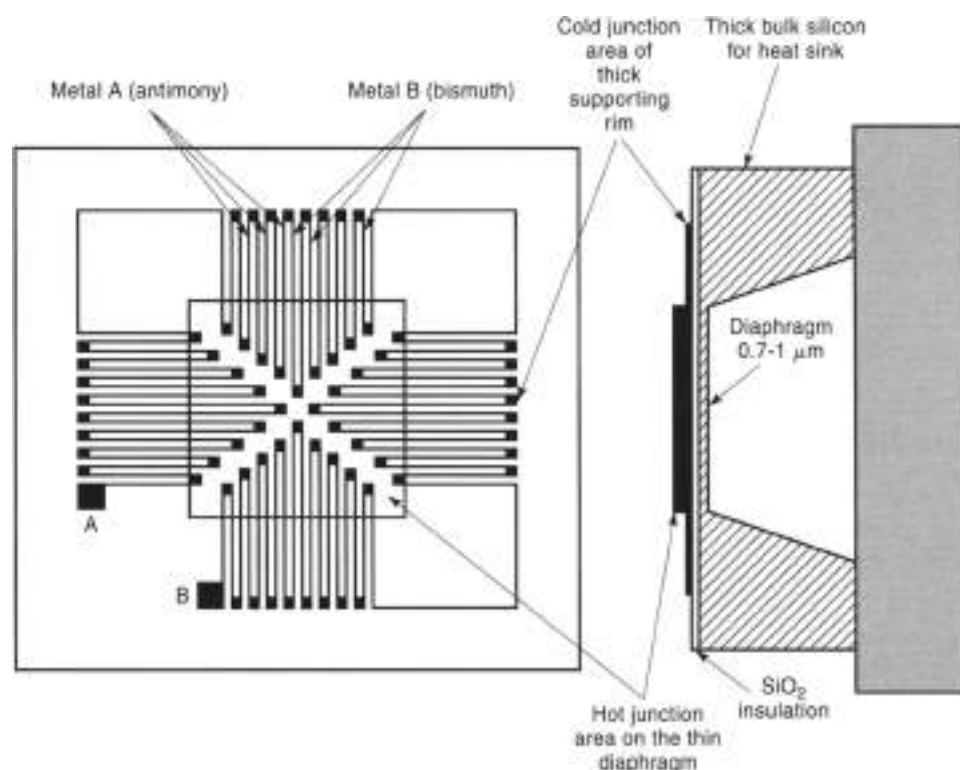


Figure 16. Thermopile manufactured in Micro Machining Technology.



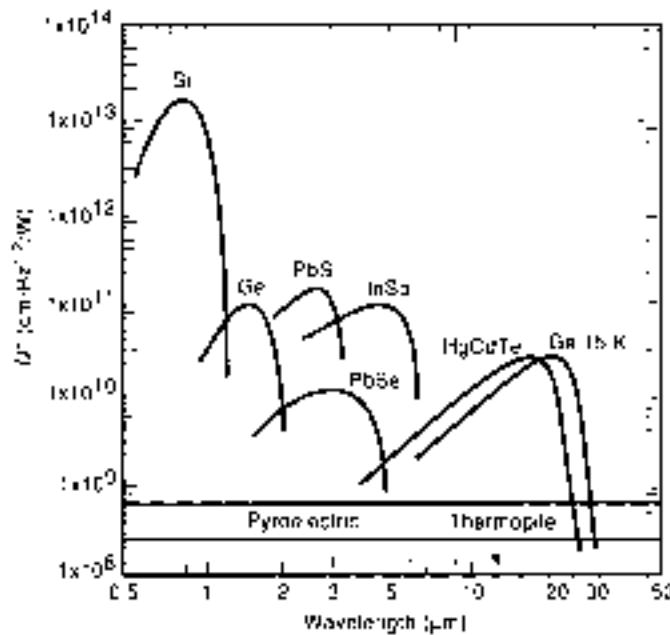


Figure 17. Specific spectral detectivity of thermal and photon infrared detectors. The design of low-temperature radiation thermometers is more complicated not only because of the lower radiance energy but by the lower sensitivity of the photon detectors working in the far IR region.

The reciprocal of NEP is called detectivity  $D^*$ . Hence the sensitivity  $S_V$  is proportional to the square root of detector area  $A$ , and the frequency band  $\Delta f$  is determined by the amplifier rather than by the detector itself. The properties of the detector are better described by specific spectral detectivity  $D^*(\lambda)$ , defined as

$$D^*(\lambda) = (A \cdot \lambda \cdot \Delta f)^{-1/2} \quad (27)$$

Thermopile specific spectral detectivity is usually about  $10^7 \text{ cm}^2 \cdot \text{Hz}^{1/2} \cdot \text{W}^{-1}$  (Fig. 17).

Bolometers with thermistors used as temperature sensing elements have the same detectivity but are less frequently used. Pyroelectric detectors where the electrical charge evoked on the piezoelectric plate follows the temperature changes are a few times more sensitive, but they do not work at steady-state conditions and the thermal flux must be mechanically modulated by means of a rotating disk with slots or holes. Mechanical chopping enables the use of alternating-current (ac) amplifiers, this is more convenient than using direct-current (dc) ones and additionally cuts off a large portion of low-frequency noise (red noise), thus increasing the specific spectral detectivity. For this reason, the chopping technique is in common use now.

Photon detectors are much more sensitive than thermal detectors and are mostly used when the limited range of wavelength is preferred. They are manufactured as semiconductor photodiodes or photovoltaic elements. The most sensitive Si detectors [ $D^*(\lambda) = 10^{13} \text{ cm}^2 \cdot \text{Hz}^{1/2} \cdot \text{W}^{-1}$ ] are suitable for visual spectra only; PbS, InAs, and InSb detectors [ $D^*(\lambda) = 10^{11} \text{ cm}^2 \cdot \text{Hz}^{1/2} \cdot \text{W}^{-1}$ ] cover the range up to 3  $\mu\text{m}$  to 5  $\mu\text{m}$ ; and only an Hg-Cd-Te detector may be used up to the 20  $\mu\text{m}$  range, achieving its maximum detectivity [ $D^*(\lambda) = 5 \cdot 10^{11} \text{ cm}^2 \cdot \text{Hz}^{1/2} \cdot \text{W}^{-1}$ ] in the far-infrared radiation region. Cooling of detectors in order to

improve their detectivity by noise reduction results sometimes in shifting of the spectral characteristics toward longer wavelengths. This effect is observed in particular for Ge detectors. Therefore, in comparing the spectral detectivities of various detectors, precise specification of their operating temperatures must be taken into account (5). Detector cooling is frequently performed by single-stage or multistage (4) cooler systems, which are compact and easy to use. The only problem to solve in their design is providing quick dissipation of the heat to enable the continuous cooling of the detector. Stirling coolers which require pumps are seldom used. It is quite possible, however, that in the near future, micropumps for that purpose will be produced by Micro Machining Technology.

### Wide-Band Radiation Thermometers

Conceptually the simplest method of radiation thermometry is to detect the thermal energy emitted by an object toward a sensor regardless of its spectral distribution. This energy depends on the temperature according to the Stefan-Boltzmann law:

$$E = \epsilon_0 \sigma T^4 - \epsilon_1 \sigma T_1^4 \quad (28)$$

where  $\sigma$  is the Stefan-Boltzmann constant,  $T$  and  $T_1$  are the object and thermometer temperatures, respectively, and  $\epsilon_0$  and  $\epsilon_1$  are the emissivities of the object and the thermometer, respectively. The exponent 4 in Eq. (28) makes the calibration curve nonlinear and increases the uncertainty at the lower range of measured temperatures where the sensitivity is also low. This form of relationship, however, decreases the influence of the object emissivity  $\epsilon_0$ . The thermometer calibrated with the use of a black body ( $\epsilon_0 = 1$ ) always indicates lower temperature  $T_0$  than is actually existent at the observed surface with the emittance  $\epsilon_0$ . By the comparison of equal states of the detector output signals during calibration and measurement, and assuming that all other factors are the same, one obtains

$$T = \frac{1}{4\sqrt{\epsilon_0}} T_0 \quad (29)$$

(The second term in Eq. (28) has been neglected, which is allowed in case of higher object temperatures.) For example, taking  $\epsilon_0 = 0.5$ , the real temperature is not twice the indicated temperature but only 19% higher, as expressed in kelvins. Nevertheless, such a great difference is not negligible and has to be corrected. The simplest and the most common way is to adjust the emissivity value in the instrument which calculates and introduces the appropriate result correction. It would be a good method if the values of  $\epsilon_0$  were known with sufficient accuracy, but usually this does not take place. Furthermore,  $\epsilon_0$  is used here in the meaning of average emissivity over the whole spectral range. It has been shown before that for many reasons the heat flux incident at the sensor is spectrally disturbed, and hence the averaging should be weighted with regard to all these disturbances. Calculations become complicated to such a degree that their practical usefulness becomes questionable. The only way is then to calibrate the thermometer at the same environmental conditions as its operational conditions. The calibration is more effective when influence variables are stable. Higher stability may be achieved by narrowing the wavelength band. For this reason the class of wide-band thermometers includes not only total radiation thermometers but also thermometers with intentionally limited spectral bands.

Total radiation thermometers are commonly equipped with thermal detectors and are used for measurements inside closed cavities as tanks or ovens where the emissivity is equal to 1 and the distance to the target is relatively low and stable. For applications in open air the thermometers with narrow band and adjustable emissivity are used. They work in the first atmospheric window (for higher temperatures) or in the second atmospheric window (for lower temperatures). For lower temperature ranges the compensation of ambient temperature ( $T_a$  in Eq. (26)) is commonly performed. Some of the devices are referred to as "near touch" radiation thermometers, in which an infrared detector is placed immediately over the surface, and the mirror optic is used to focus the radiant energy at the detector.

The majority of wide-band radiation thermometers are produced as hand-held or tripod-based instruments, commonly equipped with optical or laser aiming facility. Because of the progress in noiseless amplifications of weak electric signals, the lowest temperature range of the discussed instruments has been pushed down to the level of  $-60$  C or even  $-100$  C. It is no longer true that those radiation thermometers are destined for high temperatures only, as it was in the past. It still remains true, however, that accurate measurements may be obtained only in the case of sufficiently stable conditions. The responsibility for ensuring repeatable conditions is with the user. The accuracy of the measurement depends rather on his skills and experience than on the quality of an instrument.

#### Monochromatic Radiation Thermometers

The name for the device derives from the times when radiation thermometers were used in the visual band only and indicates that a very narrow spectral band  $\Delta\lambda$  of emitted heat flux is used for measurement purposes. In optical pyrometers the narrow band was being filtered using colored windows. Now the infrared interference filters are used for that purpose. Interference filters consist of a number of thin transparent layers deposited on a substrate. The optical properties and the thicknesses of the layers are specially matched to transmit through only a desired wavelength band  $\Delta\lambda$  with the middle wavelength equal to  $\lambda_1$ , as well as to reflect all higher and lower band frequencies. The filter remains cool because the undesired heat energy is reflected and not absorbed. The filtered bands are matched first of all with respect to the temperature range of the thermometer, but also with respect to the atmospheric windows. The part of energy emitted in wavelength band  $\Delta\lambda$  may be described as

$$E_{\lambda} = \epsilon_{\lambda} C_1 \lambda^{-5} e^{-C_2/\lambda T} \Delta\lambda \quad (30)$$

where  $\epsilon_{\lambda}$  is the emissivity at wavelength  $\lambda$ . The volume of energy incident at the detector is much lower than that in wide-band thermometers, but narrow-band photon detectors with much higher detectivity may be used in this case. Assuming the same calibration procedure as described before for wide-band thermometers, one obtains the relationship between the real temperature  $T$  and indicated temperature  $T_m$  in the form of

$$\frac{1}{T} = \frac{1}{T_m} + \frac{\lambda_1}{C_2} \ln \epsilon_1 \quad (31)$$

and for the small differences between  $T$  and  $T_m$  in the form of

$$T = T_m \left( 1 + \frac{T_m \lambda_1}{C_2} \ln \epsilon_1 \right) \quad (32)$$

The differences between real and indicated temperatures are lower than in wide-band thermometers [Eq. (29)]. However, it is more important that the value of  $\epsilon_1$  be better defined and more stable than an average emissivity  $\epsilon_a$  which has been used for calculations in the case of wide-band thermometers. Furthermore, all disturbances such as atmospheric absorption, hot gas radiation, contamination of optical parts by dust particles, and variations in detector properties have no influence on the output signal so far because they do not concern directly the filtered band  $\Delta\lambda$ . The maintenance of the monochromatic thermometers is therefore simpler and their stability is much better, but the problem of emissivity correction remains.

#### Radiation Ratio Thermometers

The idea of radiation ratio thermometers is based on simultaneous measurement of radiation in two narrow wavelength bands followed by calculating the temperature from the ratio of the results. Both bands have to be chosen on the "increasing" part of the Wien's spectral energy distribution curve (see Fig. 15), and they differ for each temperature range. Denoting the energy ratio at wavelength  $\lambda_2$  and  $\lambda_1$  by  $R$ , one obtains

$$T = \frac{\lambda_2 - \lambda_1}{\lambda_2 \lambda_1} C_2 \left[ \ln R \left( \frac{\lambda_2}{\lambda_1} \right)^5 \right]^{-1} \quad (33)$$

for all gray bodies independent from their emissivity. For nongray bodies, like all metals, in which the emissivity is a function of wavelength, a very small correction is needed:

$$T = T_m \left( 1 + T_m \frac{\lambda_2 \lambda_1}{C_2 (\lambda_2 - \lambda_1)} \ln \frac{\epsilon_1}{\epsilon_2} \right) \quad (34)$$

where  $\epsilon_1$  and  $\epsilon_2$  are emissivities for wavelength  $\lambda_1$  and  $\lambda_2$ , respectively. The difference between  $\epsilon_1$  and  $\epsilon_2$  is very low, even in metals, which causes, according to Eq. (34), a very little difference between  $T$  and  $T_m$ . The ratio thermometers may be considered emissivity-independent instruments. The design of radiation ratio thermometers is much more complicated than other radiation thermometers, and special requirements for low chrominance optic elements have to be fulfilled. Therefore this kind of instruments is rather expensive.

In summary, it is worthwhile to point out that in the last decade an intensive progress in radiation thermometry has been achieved. The basic ideas remain unchanged, an optical pyrometer with disappearing filament remains a prototype of narrow band thermometers, and a two-color pyrometer is an ancestor of the ratio thermometers, but the progress in detector technology, the introduction of interference filters, the possibilities of low noise signal amplification, and numerical result corrections make the use of these instruments simpler and allow moving the operational temperature range toward lower temperatures. The use of fiber optics to transmit the thermal energy from hardly accessible places or through the areas with nuclear radiation or strong electromagnetic fields creates the next step of development which cannot be underestimated. Up

to now the optical fibers work at short wavelengths only (up to 2.5  $\mu\text{m}$ ), but with low distances the temperatures as low as 100 C may be measured. Further progress in that field is expected.

## ERRORS AND UNCERTAINTIES

First of all we have to distinguish between errors and uncertainties. According to its definition, an error is a difference between a result of the measurement and a measured true value. Errors always have certain values and a sign, but they remain unknown. In order to determine an error, we can try to calculate it or to measure it. Errors are measured in the calibration process. A determined error added to the result with the opposite sign as a correction makes the result free from that particular error. Many errors exist which are very hard to determine, even if their sources are theoretically known. In radiation thermometry we know exactly how the emissivity influences the result, but we do not know the particular value of the target emissivity. We are only able to estimate the range of results which is affected by more or less unknown sources of errors. It is a range of uncertainty, and the most probable value of the result is commonly (but not always) placed in the middle of this range. Therefore, uncertainty is denoted as  $\pm \Delta$ , where  $\Delta$  is half of the estimated range.

All predictable errors in sensors or transducers are normally accounted by the producer by means of compensations and corrections. The accuracy data which are found in transducer specifications or in standards are uncertainty limits. ISO Guide For Expression of Uncertainties (10) distinguishes two kinds of uncertainties: type A and type B. Type A uncertainties are easy to detect and estimate by repetition of measurements performed in the same conditions and then by calculating the standard deviation of the results (classical random uncertainties). Type A uncertainties dominate in high accuracy measurements. Random effects observed in industrial measurements are caused by measured instability and environmental variations rather than by measuring instruments themselves. Type B uncertainties are those which remain constant by repetition of measurements. They may be caused by residuals remaining after nonlinearity compensation, hysteresis effects, or they may be caused by the influence of variables such as pressure, nuclear radiation, electromagnetic fields, humidity, dust, lead temperature variations, ambient temperature variations, velocity of the medium under measurement, aging, and many others. Type B uncertainties dominate in medium accuracy thermometers. Estimation of type B uncertainties needs some investigations performed either by the manufacturer of the thermometer or by the user, and some amount of experience too. It is a more complicated task than a type A estimation but the uncertainty data given in thermometer specification may be adequately used here. The calibration of a particular sensor or transducer transforms some uncertainties into errors, allowing us to account for them in a form of correction. All the rest remains as a part of a type B uncertainty. Laboratory calibration of temperature sensors is expensive, but calibration performed at the site of thermometer installation is more expensive and sometimes even impossible. On the other hand, it is impossible to restore all measurement conditions during laboratory calibration, and thus the improvement of accuracy by the recalibration is always limited.

Two remarks dealing with specification sheets ought to be pointed out here. In thermometry the uncertainty data referred to us "% of reading" or "% of full scale" (FS) have no sense because temperature scales have no physical zero, and  $\pm 2\%$  of 273 K gives  $\pm 6$  K but  $\pm 2\%$  of 0 C gives  $\pm 0$  C. All uncertainties have to be referred to in temperature units.

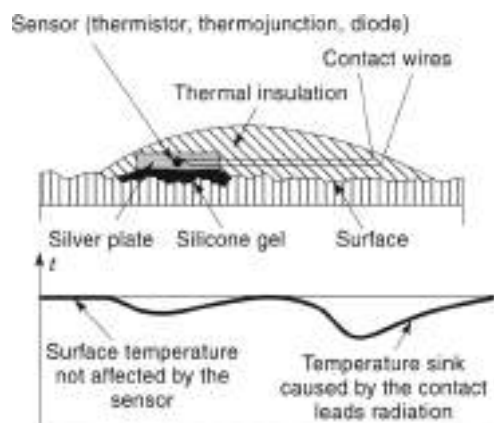
The next remark is concerned with resolution. Resolution is not a measure of accuracy and normally has nothing to do with uncertainty. It is only the ability to indicate small differences in temperature. High resolution may sometimes be very useful, but we ought to have in our minds that by the use of numerical result display the resolution is commonly 10 times better than the uncertainty involved.

In the sensor-transducer pair the uncertainty of sensor usually dominates. Excluding standards from our considerations, it may be stated that of all temperature sensors, metal resistance thermometers are the ones with the lowest uncertainty. In the temperature range of -50 C to +150 C, quartz thermometers are much more accurate and stable than all semiconductor sensors (the fact is reflected in their price). Traditional NTC thermistors seem to be less attractive in comparison with modern linear semiconductor sensors and more and more cheap integrated sensors which have the advantage of standardized calibrated output signal. However, the uncertainty of all semiconductor sensors is not better than  $\pm 0.5$  C to 1 C. Thermocouples are the less accurate sensors. Their uncertainty is rapidly increasing, with the measured temperature reaching up to +10 K at temperatures above 1000 C. This is a much higher value than for radiation thermometers where the uncertainty is rather optimistically referred to as  $\pm 1$  K to 2 K. The advantage of thermocouple sensors is their lower cost, but the costs of transducers are comparable now. Some radiation sensors are equipped with such a signal conditioning system that their output signal fits to conventional thermocouple transducers. The thermocouples, however, are the unique sensors which have to be immersed in the high-temperature medium.

The discussion presented above dealt with the instrument uncertainty, but more serious sources of uncertainty are related to the methods of temperature measurement. Generally, the problem is in the difference between the temperature of a sensor and the temperature which ought to be measured. This problem will be considered separately for steady-state conditions and for dynamic conditions.

### Steady-State Errors of Temperature Measurement

In order to provide the same temperature of an object and of a sensor, the conditions for easy heat transfer from the object to the sensor and simultaneously for the difficult heat transfer from the sensor to the ambient environment should be created. For immersion-type sensors the length of the immersed part should be as large as possible because the heat runs away from the sensor through the protective tube, and a higher temperature gradient along the shield facilitates the transfer. Thermal insulation of the tank or pipe line where the temperature is measured lowers the gradient, and therefore the error too. The problem is much more serious in the case of gas media because of their low heat transmission coefficient and the possibility of heat escape to the pipe walls by radiation. The problem weight increases with the measured temperature, not only because of the higher temperature gradient but also due to the need for



**Figure 18.** The measurement of surface temperature. The heat is accumulated in the silver plate with high thermal conductivity in order to equalize the surface and sensor temperature. The air gap between the surface and the plate might be filled with silicone gel. The connection leads act as radiators and therefore run along the surface to a distance far enough to avoid the surface temperature disturbance near the sensor.

increasing shield thickness and for better thermal insulation of the sensor by protective ceramics.

The measurement of surface temperature is the best case for obtaining a false result. The sources of errors are the same, but the available remedies are very limited. The typical situation is schematically presented in Fig. 18. The sensor ought to have the best thermal contact with the surface but simultaneously should be insulated from the environment. The thermal insulation provides protection from the temperature decrease at the point of measurement, but perhaps it may cause a local increase of the temperature especially when the surface was intensively cooled. Touching of the surface by the sensor involves thermal effects which are different in each particular case and practically unpredictable. From this point of view, the advantage of radiation thermometry for surface temperature measurements is evident.

#### Dynamic Errors

Dynamic errors are caused by the thermal inertia of the sensors and become important while transient temperatures have to be measured. The dynamic properties of a sensor are described by the response time of the output signal after a rapid change of sensor temperature, or by the time constant. The response time is usually defined as the time elapsing between 10% and 90% of the output signal change, but other definitions are also used. The response time depends on the sensor properties such as its material, shape, and dimensions, but depends first of all on the environment surrounding the sensor, characterized by the heat transmission coefficient. That coefficient is a few times greater in liquids than in gases, and it increases with the velocity of the medium. With any comparison of the sensor dynamic properties, exactly the same conditions have to be secured.

Assuming the linearity of the thermometer the idea of Laplace transformation and transfer function may be used for describing the thermometer dynamic properties. The simplest model of the dynamic behavior of a sensor may be presented in

a form of the first-order transfer function:

$$K(s) = \frac{T_1(s)}{T_0(s)} = \frac{1}{1 + \tau_1 s} \quad (35)$$

where  $T_1(s)$  is the Laplace transformation of the thermometer temperature,  $T_0(s)$  is the Laplace transformation of the object temperature,  $s$  is the Laplace operator and  $\tau_1$  is a time constant. The model is valid for nonembedded sensors only, which are rather seldom used. Two other models are also used. The second-order model

$$K_2(s) = \frac{1}{(1 + \tau_1 s)(1 + \tau_2 s)} \quad (36)$$

accounts for the delay of the output signal and better describes the thermometer with the shielded sensors. A model with the differential action

$$K_0(s) = \frac{1 - \tau_1 s}{(1 + \tau_1 s)(1 + \tau_2 s)} \quad (37)$$

may in turn be used for surface temperature thermometers.

All these models may be treated as rough approximations because in fact the dynamic properties of temperature sensors are nonlinear. The experiment has shown that even for a nonembedded thermocouple sensor the time constant at 400 °C has been three times larger than at that 1200 °C (5). Also the warming curves and quenching curves always differ from one another.

For dynamically nonlinear sensors, it is not so easy to transform the results from one operating condition to other. One of the methods for estimation of dynamic behaviors of a sensor under its operating conditions uses an intentionally generated self-heating impulse in order to record the thermometer answer and then to calculate its dynamic parameters.

The dynamic behavior of the radiation thermometers depends on the dynamic properties of the infrared sensors which are rather slow, with the response time varying from 0.2 s to 2 s, and which are dynamically nonlinear too.

#### BIBLIOGRAPHY

1. H. Preston-Torrens, The international temperature scale of 1990 (ITS-90), *Metrologia*, 27: 3-10, 1990.
2. T. J. Quinn, *Temperature*, 2nd ed., New York: Academic Press, 1990. Deep study of ITS and high accuracy temperature measurements.
3. W. Gopel, J. Hesse, and J. N. Zemel (eds.), *Sensors: A Comprehensive Survey*, Vol. 4, *Thermal Sensors*, T. Rienfi, J. Scholz (eds.), Weinheim: VCH, 1989.
4. *The Temperature Handbook*, Stanford, CA: Omega Engineering, 1992. A comprehensive review of the market of thermometers and thermal sensors.
5. I. Von Karolyi-Gyossy, *Thermometer Praxis*, 2nd ed., Essen: Vulkan Verlag, 1987. In German. Comprehensive work with a lot of practical information.
6. J. F. Schooley, *Thermometry*, New York: ORU Press, 1988.
7. L. Michalski, K. Fekorskarf, and J. McGhee, *Temperature Measurement*, New York: Wiley, 1999.
8. D. P. De Witt and G. D. Nutter (eds.), *Theory and Practice of Radiation Thermometry*, New York: Wiley, 1988.

9. *The Infrared Temperature Handbook*, Stanford, CA: Omega Engineering, 1991.
10. *Guide to the Expression of Uncertainty in Measurement*, ISO, IEC, OIML, BIPM, Paris, in Switzerland 1993.
11. J. Zakrzewski, "Temperature Sensors," in J. G. Webster (ed.), *Wiley Encyclopedia of Electrical and Electronics Engineering*, Vol. 21, New York: Wiley, 1999, p. 1005.

JAN ZAKRZEWSKI  
Silesia Technology University of  
Poland

## THERMOELECTRIC CONVERSION

Thermoelectric effects offer a means of converting heat to electric power using only solid-state components. By supplying electrical power to a thermoelectric material, it may also be used to produce refrigeration. Because a thermoelectric system has no moving parts, it may be reliable and compact. However, currently available power generators and refrigerators based upon thermoelectric effects are too expensive and inefficient for large-scale applications. Today, thermoelectric generators are used to produce electric power in satellites and thermoelectric refrigerators are used to cool small heat loads or in situations requiring portability.

In this article, we review the physics of thermoelectric effects and survey some present and potential applications of thermoelectric technology. Then an explanation of the thermoelectric effects and the operation of thermoelectric generators and refrigerators is given. There is no attempt to be comprehensive here as the detailed theory of thermoelectric effects has been well described elsewhere [1,2].

Following this, we briefly discuss the present status of thermoelectric refrigeration and power generation. These technologies are commercially active, with thermoelectric options dominating several niche markets. We also describe several lines of research that attempt to improve the efficiency of thermoelectric materials.

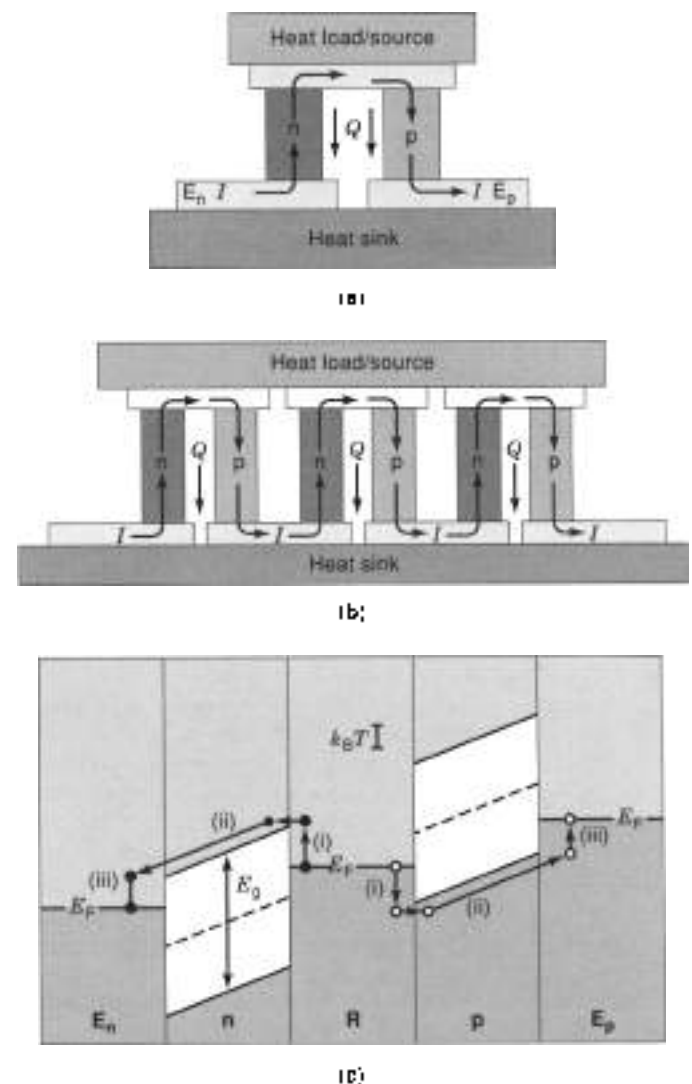
Finally, a new research field, cryogenic thermoelectric refrigeration, is described. Here, the efficiency can be quite high, so that the temperature may be reduced by a large factor. Also, we describe proposed designs for thermoelectric refrigeration using quantum-scale devices as well as recent experimental verification of these concepts in superconducting tunnel junctions.

## THERMOELECTRIC EFFECTS

Thermoelectric effects occur because electrons and holes (which are vacancies in the electron sea in conductive materials) carry energy. Thermal energy at a temperature  $T$  (in K or Kelvin, measured from absolute zero) is distributed among all parts of a conductive material—the atoms making up the crystal lattice, the electrons, and the holes. As a result, electrons and holes have a range of energies; the electronic "energy spread" in a metal is a few times  $k_B T$ , where  $k_B$  is Boltzmann's constant. A thermoelectric effect occurs when the most energetic (or "hot") electrons and holes are transported preferentially during current flow in a conductive material. In a sense, the hot electrons

and holes "evaporate" from one part of a material to another, cooling the region that they leave in the same way as the evaporation of water has a cooling effect.

There are several distinct but related physical effects that arise from this heat capacity of electrons and holes. When an electric current is driven through a thermoelectric material (for instance, by applying a voltage across the thermoelectric material), a heat flow occurs as well. This is called the Peltier effect, which is the basis of thermoelectric refrigeration. When one end of a piece of thermoelectric material is hotter than the other, a voltage develops across the material. This is called the Seebeck effect, which is used in thermoelectric power generation. Finally, when a temperature gradient exists inside a thermoelectric material, a gradient in the electric potential develops due to the Thomson effect. We will not discuss the Thomson effect here, since it is beyond the scope of this article, but more details may be found in Refs. 1–3.



**Figure 1.** Diagram of a thermoelectric device. (a) Heat flows vertically and electric current flows from left to right. (b) Placing multiple thermoelements electrically in series and thermally in parallel increases the cooling power. (c) Representation of the microscopic mechanism of the thermoelectric effects in terms of electron energy states in each material. Vertical scale is electron energy.

These effects are known in the science of thermodynamics as reversible processes (1), in the sense that they may operate in the forward or in reverse directions. For instance, by reversing the direction of current flow in a thermoelectric refrigerator, the direction of heat flow is reversed, thereby causing heating instead of cooling. The efficiency of thermoelectric refrigerators and generators is limited by certain irreversible phenomena (1). The two main irreversible effects are thermal conduction and electrical resistance.

When a temperature gradient exists in any material, heat flows through the material to minimize the temperature gradient. This is clearly an irreversible process, otherwise objects would spontaneously develop cold and hot spots until they cracked or evaporated. Thermal conduction in a thermoelectric refrigerator represents a heat leak. In a thermoelectric generator, it is manifest as a power loss.

The passage of electricity through a conducting material is accompanied by the generation of heat, as the electrons and holes lose energy to various mechanisms such as the crystal lattice. Similar to thermal conduction, this resistive process is irreversible. If it were reversible, then a warm object would spontaneously emit electric current and cool itself to absolute zero. Also similar to thermal conduction, resistive losses represent an irreversible heat leak in a thermoelectric refrigerator and a power loss in a thermoelectric generator.

Under some conditions, thermoelectric effects take on an irreversible nature when they are driven past the "small signal" limit (4-6). In the following, we will discuss situations under which these extreme conditions are of benefit to thermoelectric refrigeration.

A generic thermoelectric circuit is shown in Fig. 1(a). The thermoelements *n* and *p* embody the thermoelectric effects in this generic circuit. In the *positive* thermoelement *p*, electric current *I* and heat *Q* flow in the *same* direction. In the *negative* thermoelement *n*, heat is conducted in the direction *opposite* to current flow. In the circuit of Fig. 1(a), the thermoelements *n* and *p* are said to be thermally in parallel, as heat flows through each in parallel, and electrically in series, because electric current must traverse them sequentially. To enhance the thermoelectric performance of this circuit, it can be seen from Fig. 1(b) that many more such elements may be added so that they are electrically in series, and thermally in parallel.

The circuit is completed by metallic elements  $R$ ,  $E_n$  and  $E_p$ , which serve as conductors and heat sinks.  $R$  is a "reservoir" that is attached to a heat source (in the case of a thermoelectric generator) or to a device to be cooled (in a thermoelectric refrigerator).  $E_n$  and  $E_p$  are "electrodes" that are used to electrically connect the thermoelectric circuit to an external electric circuit (a current source to drive thermoelectric refrigeration or an electric load in the case of thermoelectric generation).  $E_n$  and  $E_p$  also are thermally sunk to a heat sink to maintain their temperature at  $T$ .

When the thermoelectric circuit of Fig. 1(a) is in operation, electric current flows from  $E_n$ , through *n*, through *R*, through *p*, to  $E_p$ . Heat *Q* flows from *R*, through *n* and *p*, to  $E_n$  and  $E_p$ . In the case of a thermoelectric refrigerator, a voltage is applied between  $E_n$  and  $E_p$  to drive an electric current which causes the flow of heat *Q* from the heat load to the heat sink by the thermopower of *n* and *p*. For thermoelectric power generation, a source of heat is applied to *R*, which causes heat *Q* to flow down. Due to the thermopower of *n* and *p*, electric current *I* flows up through *n* and down through *p*, the net result of which

is to drive electric current from  $E_n$  to  $E_p$ , which may be used to drive an electric load.

Any thermoelectric effect is caused by differences in the energy levels electrons occupy in a material, and hence may be visualized by tracing the path of electrons through these different energy levels. Such a diagram is shown in Fig. 1(c), which represents the energy levels (vertical scale) that electrons occupy when they are traveling through the circuit components of Fig. 1(a).

Several aspects of solid-state physics must be understood to interpret Fig. 1(c): First, electrons are thermally excited so that they are continually absorbing and emitting energy from other electrons and the crystal lattice. The typical amount of energy exchanged is about  $k_B T$ , as mentioned earlier. The magnitude of  $k_B T$  is shown in Fig. 1(c) for reference. If we could employ "Maxwell's demon" (3) to capture the electrons when they were excited to a high energy and use their excess energy to perform work, then heat would be extracted from the metal, resulting in a cooling effect. This is how a thermoelectric refrigerator works.

Consider an electron in *R*, at an energy called the "Fermi level"  $E_F$ , which is essentially the average energy for the electrons that participate in electrical transport (3). Because electrons belong to the particle class called Fermions, they fill energy states one at a time until  $E_F$  is reached. Thermal excitations occur for energies around  $E_F$ , with a range of  $k_B T$ . Some states below  $E_F$  are vacated by an electron, which has absorbed energy to be excited to an empty state. These vacant states are called holes, and can move about in the same way as electrons. One can imagine this as a percolation process, although in terms of quantum mechanics electrons and holes are described similarly (3). The excitation energy of a hole is opposite to that of an electron; thermally excited holes lie below  $E_F$ .

To understand the microscopic operation of a thermoelectric circuit, we can trace the path of an electron through Fig. 1(c): (i) First, the electron in *R* is thermally excited to above  $E_F$ . Then the electron may move to the conduction band of the semiconductor *n*. This band is shown with a slope because the electrical resistance of *n* causes a voltage drop with distance (ii). Finally, the electron is deposited in the electrode  $E_n$ , where it gives up its "extra" energy relative to  $E_F$ . An analogous path for a hole is shown in Fig. 1(c) as well.

One thing to note is that the direction of current flow for both the electron and hole paths in Fig. 1(c) is to the right because the charge of the electron is negative and the charge of a hole is positive. An analogy between electrostatic potential energy (as is shown in this diagram) and gravity might be drawn here: it is as though electrons slide down slopes in the energy bands and holes percolate up like bubbles. Thus, an "extra" current source must be connected to *R* for a current to traverse the circuit. It is also important to note where the different energy-exchange mechanisms are taking place. In Step (i), the electron and hole absorb thermal energy from the other electrons or the crystal lattice of *R*, which in turn absorbs heat from the heat source in the case of thermoelectric generation or the heat load in the case of thermoelectric refrigeration. In Step (ii), the electron and hole give up heat by Ohmic heating due to the electrical resistance of *n* and *p*. Finally, in Step (iii), the electron and hole release their excitation energy to the electrodes  $E_n$  and  $E_p$ , which then release the heat to the heat sink. The heat leak due to the thermal conductivity of *n* and *p* is not shown here, since it is mostly due to the crystal lattice rather than the electrons.

## THERMOELECTRIC REFRIGERATION

By applying electric current to the thermoelectric circuit of Fig. 1(a), heat is extracted from  $R$ . This refrigerator is driven by the Peltier effect. Specifically, if a current  $I$  is driven through a thermoelectric material, a flow of heat

$$Q = \Pi I \quad (1)$$

accompanies the flow of current;  $\Pi$  is known as the Peltier coefficient of the thermoelectric material.

For a negative thermoelement such as  $n$  in Fig. 1(a),  $\Pi$  is negative. Thus, the electric current and heat current flow in opposite directions in  $n$ , so that while the current flows up, heat flows down. For a positive thermoelement such as  $p$  in Fig. 1(a),  $\Pi$  is positive. In this case, the electric and heat currents flow in the same direction; in Fig. 1(a), both heat and electric current flow downward through  $p$ . Thus, heat flows downward through both  $n$  and  $p$  so that both contribute to cooling  $R$ .

In a real thermoelectric refrigerator, there are other sources and flows of heat that must be considered when trying to understand the refrigerating behavior. If the thermoelectric refrigerator is cooling a heat load (e.g., a semiconductor diode laser or a solid state infrared detector), which is producing a heat flow  $Q_L$ , then this heat flows into  $R$  and adds to the heat load that the thermoelements must carry away to accomplish a given amount of refrigeration.

In an ideal thermoelectric material, there would be no electrical resistance, and hence no Ohmic heating. However, this is not the case for real materials. If the electrical resistance of  $n$  is  $R_n$  and that of  $p$  is  $R_p$ , then Ohm's law (3) dictates that the amount of heat  $I^2 R_n$  is generated in  $n$  and  $I^2 R_p$  is generated in  $p$ . Since this heat is generated uniformly throughout the thermoelectric materials, it is generally assumed that half the heat is transmitted up to  $R$ , and half goes down to  $E_n$  and  $E_p$ .

As the thermoelectric refrigerator operates, the flow of heat  $Q$  causes a temperature difference  $\Delta T$  to develop between  $R$  and  $E_n$  and  $E_p$ . In an ideal thermoelectric material, the thermal conductivity would be zero, so that no heat leaked back across to the cold side. However, thermoelectric materials always have a significant thermal conductivity. Assume that the values of thermal conductance for  $n$  and  $p$  are  $K_n$  and  $K_p$ , respectively. Then a heat leak of  $K_n \Delta T$  flows upward through  $n$  and  $K_p \Delta T$  flows through  $p$ .

Without a heat load, the net heat that flows from  $R$  to  $E_n$  is

$$Q_n = \Pi_n I - K_n \Delta T - \frac{1}{2} I^2 R_n \quad (2)$$

and that flowing from  $R$  to  $E_p$  is

$$Q_p = \Pi_p I - K_p \Delta T - \frac{1}{2} I^2 R_p \quad (3)$$

The cooling power of the thermoelectric refrigerator is the sum of  $Q_n$ ,  $Q_p$ , and the heat load  $-Q_L$ .

$$Q_{\text{cooling}} = Q_n + Q_p - Q_L = (\Pi_p + \Pi_n) I - (K_n + K_p) \Delta T - \frac{1}{2} I^2 (R_n + R_p) - Q_L \quad (4)$$

$Q_{\text{cooling}}$  represents the heat flow that the thermoelectric refrigerator draws from  $R$  and the heat load. If  $Q_{\text{cooling}}$  is positive at any temperature, then  $R$  will drop in temperature due to

the net cooling action. If  $Q_{\text{cooling}}$  is positive in the absence of a temperature difference  $\Delta T$ , then  $R$  will cool until it sits at a temperature below the "ambient" temperature  $T$ .

The coefficient of performance  $\psi$  of a thermoelectric refrigerator is defined to be the ratio of the cooling power to the electrical power  $P$  required to drive the current  $I$  through the refrigerator

$$\psi = \frac{Q_{\text{cooling}}}{P} \quad (5)$$

$P$  is computed by the relationship  $P = IV$ , where  $V$  is the voltage drop across the device. To compute  $V$ , we need to describe the other major thermoelectric effect—the Seebeck effect.

A contact potential  $V_{\text{contact}}$  generally exists at a junction of dissimilar materials. In thermoelectric devices, this contact potential and its temperature dependence are utilized. In particular, the contact potential between a thermoelectric material (e.g.,  $n$  or  $p$ ) and an ordinary metal (e.g.,  $R$ ,  $E_n$ , or  $E_p$ ) is given by

$$V_{\text{contact (thermoelectric, metal)}} = \alpha T + V_0 \quad (6)$$

where  $V_0$  depends upon the particular materials chosen and the text in curly brackets indicates the direction of positive contact potential; in this case from thermoelectric material to metal. The linear temperature coefficient  $\alpha$  in the contact potential is called the Seebeck coefficient. Because the junctions at  $R$  and those at  $E_n$  and  $E_p$  sit at different temperatures, a net potential drop occurs across each thermoelement

$$\begin{aligned} V_n &= V_{\text{contact (} R, n)} - V_{\text{contact (} n, R)} \\ &= V_{\text{contact (} n, R)} - V_{\text{contact (} n, R)} \\ &= (\alpha_n T + V_0) - (\alpha_n T - \Delta T) - V_0 \\ &= -\alpha_n \Delta T \\ V_p &= V_{\text{contact (} R, p)} + V_{\text{contact (} p, E_p)} \\ &= -V_{\text{contact (} p, R)} + V_{\text{contact (} p, E_p)} \\ &= -(\alpha_p T - \Delta T) - V_0 + (\alpha_p T + V_0) \\ &= \alpha_p \Delta T \end{aligned} \quad (7)$$

Due to the electrical resistance of the thermoelements, there is also an Ohmic voltage drop  $I(R_n + R_p)$  in the circuit; thus the total power supplied to the thermoelectric refrigerator is

$$\begin{aligned} P &= IV \\ &= I(\alpha_p - \alpha_n) \Delta T - I^2 (R_n + R_p) \end{aligned} \quad (8)$$

The coefficient of performance becomes

$$\begin{aligned} \psi &= \frac{Q_{\text{cooling}}}{P} \\ &= \frac{(\Pi_p + \Pi_n) I - (K_n + K_p) \Delta T - \frac{1}{2} I^2 (R_n + R_p) - Q_L}{I(\alpha_p - \alpha_n) \Delta T - I^2 (R_n + R_p)} \end{aligned} \quad (9)$$

Generally, one would like to know the best coefficient of performance that a given material system can supply. For this case,

we consider the limit with no load ( $Q_L = 0$ ). Also, we consider the case when the thermoelements have similar properties, namely

$$\begin{aligned} R_n &= R_p = R; & K_n &= K_p = K; & -\alpha_n &= \alpha_p = \alpha; \\ -\Pi_n &= \Pi_p = \Pi \end{aligned} \quad (9)$$

so that the coefficient of performance simplifies to

$$\eta' = \frac{(11I - K\Delta T) - (1I^2R)}{I\alpha\Delta T + I^2R} \quad (15')$$

We can further simplify Eq. (15') using the Kelvin relationship

$$\Pi = T\alpha \quad (10)$$

which may be derived from the Onsager relation in the theory of irreversible thermodynamics. Equation (10) is only valid for small  $\nabla T$  and  $\Delta T$ . This topic is beyond the scope of this article; more details may be found in Refs. 2-4.

We can find the value of  $I$  at which Eq. (15') reaches its maximum value by taking the derivative of  $\eta'$  with respect to  $I$  and setting it equal to zero. After performing this operation and substituting this current value into Eq. (15'), we arrive at the optimal value of the coefficient of performance

$$\begin{aligned} \eta'_{\text{max}} &= \frac{T - \Delta T}{\Delta T} \left( \frac{\sqrt{1 + ZT} - \frac{T}{T - \Delta T}}{1 + ZT + 1} \right) \\ &= \eta'_{\text{Carnot}} \left( \frac{\sqrt{1 + ZT} - \eta'_{\text{Carnot}}}{\sqrt{1 + ZT} + 1} \right) \end{aligned} \quad (15'')$$

where  $T = 1/2(T + (T - \Delta T))$  is the mean temperature in the thermoelements and  $\eta'_{\text{Carnot}} = (T - \Delta T)/\Delta T$  is the Carnot thermodynamic maximum efficiency of a refrigerator (1-3). The quantity

$$Z = \frac{\alpha^2}{RK} \quad (11)$$

is known as the thermoelectric figure of merit, and has the dimensions of inverse temperature. Thus, the quantity

$$ZT = \frac{\alpha^2}{RK} T \quad (12)$$

is known as the dimensionless figure of merit, which summarizes the thermoelectric performance of a material at temperature  $T$ . More analytical details of the efficiency of a thermoelectric refrigerator may be obtained from Refs. 1 and 2. Here, we will restrict ourselves to general comments regarding the status of thermoelectric technology.

The material parameters of the thermoelements enter Eq. (15') only through  $Z$ . In this sense, the figure of merit summarizes the thermoelectric performance of a thermoelectric material. Although Eq. (11) includes the electrical parameters  $R$  and  $K$ , we can see that it may also be defined in terms of intensive material parameters. For instance, suppose that the thermoelements are of uniform cross-sectional area  $A$  and length

$L$ . Then the thermal conductance and electrical resistance may be written as

$$K = \kappa \frac{A}{L}; \quad R = \rho \frac{L}{A} \quad (13)$$

where  $\kappa$  is the thermal conductivity and  $\rho$  is the electrical resistivity of the thermoelements. The figure of merit then becomes

$$Z = \frac{\alpha^2}{\rho\kappa} \quad (14)$$

which contains only intensive quantities, and thus the thermoelectric figure of merit is itself an intensive material parameter. The dimensionless figure of merit becomes

$$ZT = \frac{\alpha^2}{\rho\kappa} T \quad (12')$$

Maximizing  $ZT$  is thus seen as the central goal of thermoelectric materials research. In practice, it has been difficult to exceed  $ZT = 1$  (at room temperature). There has even been speculation that  $ZT = 1$  represents a fundamental physical limit (1,2). In any case,  $ZT$  must be in the range of 3 to 5 for thermoelectric refrigeration to compete with conventional vapor compression refrigeration (1). Efficiency is important in a refrigerator, both in terms of the economic and environmental costs of power consumption and in terms of the amount of heat rejected for a given amount of cooling. Thus, thermoelectric materials have not yet offered a strong competitor in the area of large-scale refrigeration or air conditioning.

The practical difficulties in maximizing  $ZT$  are in a compromise between the material parameters that comprise  $ZT$ . For instance, metals have a low electrical resistivity but also a low thermopower and a high thermal conductivity. The least thermally conductive materials are often electrically insulating.

As it turns out, the best thermoelectric materials have been heavily doped semiconductors. Figure 2 shows a summary of the thermal and electrical properties of conductive materials as a function of electron concentration. It can be seen that electrical and thermal conductivity tend to increase with the electron concentration, whereas the thermopower tends to decrease. These factors lead to a peak in thermoelectric figure of merit in the semiconductor range.

The best materials for thermoelectric refrigeration are compound semiconductors. The most common material is bismuth telluride ( $\text{Bi}_2\text{Te}_3$ ). This material is used in commercially available thermoelectric refrigerators (such as the one shown in Fig. 3), and may attain  $ZT$  values approaching 1. In fact, thermoelectric refrigerators (such as the one depicted) have a growing market in cooling solid-state electronic devices such as diode lasers and infrared radiation detectors. They are also dominant in areas where compactness or portability are required, such as in portable refrigerated food containers or biological-specimen transport vessels. Also, thermoelectric refrigerators can be operated in reverse to heat a sample; thus active temperature control is possible, as utilized in biological applications.

Another practical aspect of thermoelectric technology is the use of multiple stages in commercial devices. These can offer improved efficiency because the design and materials



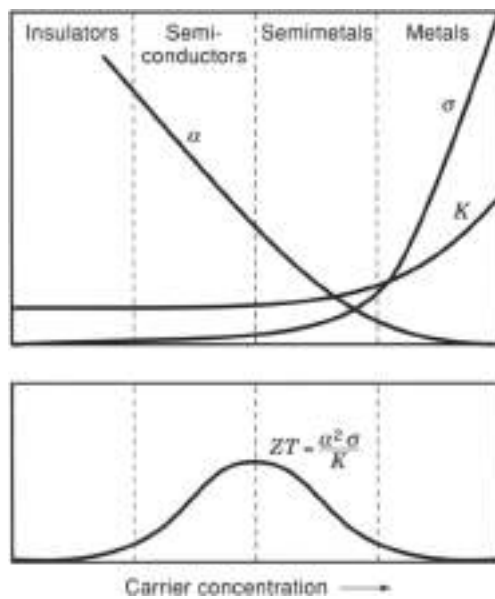


Figure 2. Schematic graph of the properties of thermoelectric materials as a function of electron concentration (1).

properties can be optimized for the operating temperature of each stage. Lower temperatures can be achieved this way in thermoelectric refrigerators.

#### THERMOELECTRIC POWER GENERATION

The circuit of Fig. 1(a) may be used as a thermoelectric generator by heating  $R$ . Heat  $Q_h$  is supplied from a heat source to the hot electrode, which causes a flow of thermoelectric current  $I$ . The thermal conductance  $K$  of the thermoelements allows heat to leak from the hot electrode to the ambient-temperature heat sink, and Ohmic heating of the thermoelements dissipates some of the generated power. The efficiency  $\eta$  of a thermoelectric generator is defined to be

$$\eta = \frac{P}{Q_h} = \frac{IV}{Q_h} \quad (14)$$

where  $P$  is the electrical power supplied to a load through the ambient-temperature electrodes and  $V$  is the total voltage drop

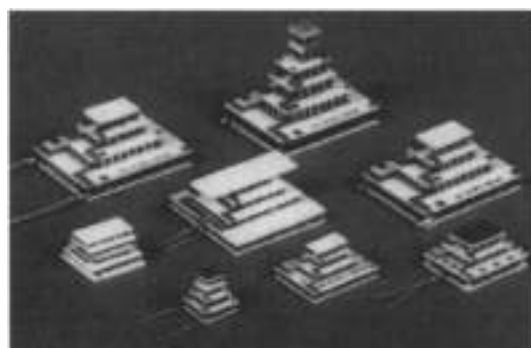


Figure 3. Photograph of a commercially available thermoelectric refrigerator using  $\text{Bi}_2\text{Te}_3$  semiconductors as thermoelements. Product literature from Marlow Industries, Inc.

across the generator. By following a line of reasoning analogous to the forementioned one for thermoelectric refrigeration, the efficiency of a thermoelectric generator may be expressed in terms of the dimensionless figure of merit of the thermoelement materials

$$\eta = \frac{\Delta T}{T + \Delta T} \left( \frac{\sqrt{1 + ZT} - 1}{\sqrt{1 + ZT} + \left( \frac{T}{T + \Delta T} \right)} \right) \quad (14')$$

$$= \eta_{\text{Carnot}} \left( \frac{\sqrt{1 + ZT} - 1}{\sqrt{1 + ZT} + (1 - \eta_{\text{Carnot}})} \right)$$

where  $T$  and  $Z$  are defined in the foregoing and  $\eta_{\text{Carnot}} = \Delta T / (T + \Delta T)$  is the Carnot thermodynamic limit on the efficiency of a generator (1-3). As discussed in the foregoing,  $ZT$  is around 1 at best. For this value,  $\eta$  ranges from  $0.2\eta_{\text{Carnot}}$  to  $0.3\eta_{\text{Carnot}}$ . For a temperature difference of 100 C with  $T = 300$  K,  $\eta_{\text{Carnot}} = 0.25$  and the thermoelectric efficiency with  $ZT = 1$  would be under 6%. Thus, thermoelectric power generation is not competitive with more conventional means. For comparison, a solar cell may have a generating efficiency of 8%, an automobile engine may have 15% efficiency, and a steam-generating plant may achieve 20 to 40% efficiency (1).

What would it take for thermoelectrics to compete in power generation? Suppose for a moment that  $ZT$  for a given material was optimized over the temperature range 300 to 600 K, and that  $T = 300$  K and the heat-source temperature is 600 K. Then  $\eta_{\text{Carnot}} = 0.5$  and, for  $ZT = 1$ , the generator becomes 11% efficient. For  $ZT = 3$ , the generator becomes 20% efficient, the point at which it begins to compete. Thus, if thermoelectric materials are to be competitive with more conventional systems for mainstream power-generation applications,  $ZT$  must be improved by at least a factor of three. Furthermore, the materials must be inexpensive, whereas semiconductors tend to be fairly expensive. Thus, the prospects for large-scale power generation using thermoelectrics are not great at present.

As in refrigeration, the present applications of thermoelectric power generation take advantage of the compactness and reliability of solid-state components. The most prominent application is in power generation for satellites, in which a radiation source such as a radioactive isotope is used to supply heat to the generator. The material system of choice is the alloy of silicon and germanium, which has a lower thermal conductivity than the crystalline forms due to enhanced phonon scattering, and is thermally stable enough to operate at a high hot-electrode temperature. These generators have been found to be extremely reliable, and nuclear power provides sufficient heat to overcome the relatively low-generating efficiency (1).

#### EFFORTS TO IMPROVE THERMOELECTRIC EFFICIENCY

The traditional approach to optimizing the efficiency of thermoelectric conversion has been incremental. For instance, commercial devices today are still based upon  $\text{Bi}_2\text{Te}_3$  and related compounds, developed in the 1950s and 1960s, with advances mostly coming in the form of material purity or means of attaching thermoelements.

Recently, a number of workers have become interested in revisiting the basic assumptions of thermoelectric conversion (1-20). Later, we will discuss the new field of cryogenic

thermoelectric refrigeration (4,5,7,8). Here, we briefly summarize recent work in thermoelectric conversion at higher temperatures, which could impact the broadest range of applications.

A notable improvement in materials synthesis over the last few decades has been thin-film deposition with high crystalline quality and purity by using methods such as molecular beam epitaxy (11). With this capability, traditional semiconducting materials may be formed into structures not available to the original designers of thermoelectric devices in the 1950s and 1960s.

Accordingly, several groups have begun exploring the thermoelectric properties of multilayer structures composed of layers of the  $\text{Bi}_2\text{Te}_3$  and related compounds traditionally used as thermoelectric elements (9). This activity was initially of a theoretical nature, and it was observed that the thermoelectric figure of merit of these materials in a properly designed multilayer may be increased from its bulk value. This was confirmed experimentally, but the figure of merit of the entire structure, including nonthermoelectric layers, has not been demonstrated yet to exceed the bulk value of the thermoelectric materials.

It has recently been proposed (6) that thermionic effects might be used to overcome limitations imposed by the "linear-response" regime implicitly assumed in the traditional theoretical understanding of thermoelectric effects as described in earlier. In a thermionic effect, electrons or holes are thermally excited across a large potential barrier and then transported away. In terms of the microscopic picture of Fig. 4(a), the conduction would become thermionic if the barrier between  $R$  and  $n$  or  $p$  were much higher than  $k_B T$ . To overcome the potential barrier, each electron or hole must absorb a large amount of energy from the crystal lattice; thus a large amount of heat is extracted for a relatively smaller amount of electric current than is the case for a traditional thermoelement. With thermionic effects, there is also the possibility of introducing a vacuum barrier or insulating material between electrodes to reduce heat leaks.

Another exciting area of materials research is along the traditional lines of bulk thermoelectric devices, but in new material systems. The skutterudites are a new thermoelectric system proposed recently (10). These materials have a complex crystal structure which frustrates heat flow, and can be doped to provide electrical conductivity. Initial measurements of the dimensionless figure of merit in this system have exceeded 1, with values of 1.4 to 1.5 being reported by several groups. Because the skutterudites family is unexplored and fairly extensive, there may exist a skutterudites thermoelectric material with a sufficiently high figure of merit and adequate thermal and mechanical stability to compete in more large-scale applications.

### CRYOGENIC THERMOELECTRIC REFRIGERATION

At temperatures near absolute zero, the reduction in ambient thermal energy allows quantum-mechanical effects to dominate the properties of some types of matter, giving rise to extraordinary phenomena. Some metals lose all electrical resistance and fall into the superconducting state (3). The viscosity of liquid helium vanishes, causing it to flow over the sides of a container (3). This superfluidity of liquid helium gives rise to thermodynamic properties, which are exploited in the dilution refrigerator, the most common means of performing experimental work at temperatures between 0.01 and 1 K.

The low-temperature properties of matter of interest here are related to phonons, the quantum-mechanical manifestation of vibrations in the crystalline lattice. Phonons are the predominant repository of heat in solids at temperatures more than a few kelvins above absolute zero. Near absolute zero, however, electrons in a metallic solid hold most of the heat, and hence thermal energy. Thus, cooling the electrons in a metal at these low temperatures should provide an effective means of refrigerating the entire metal.

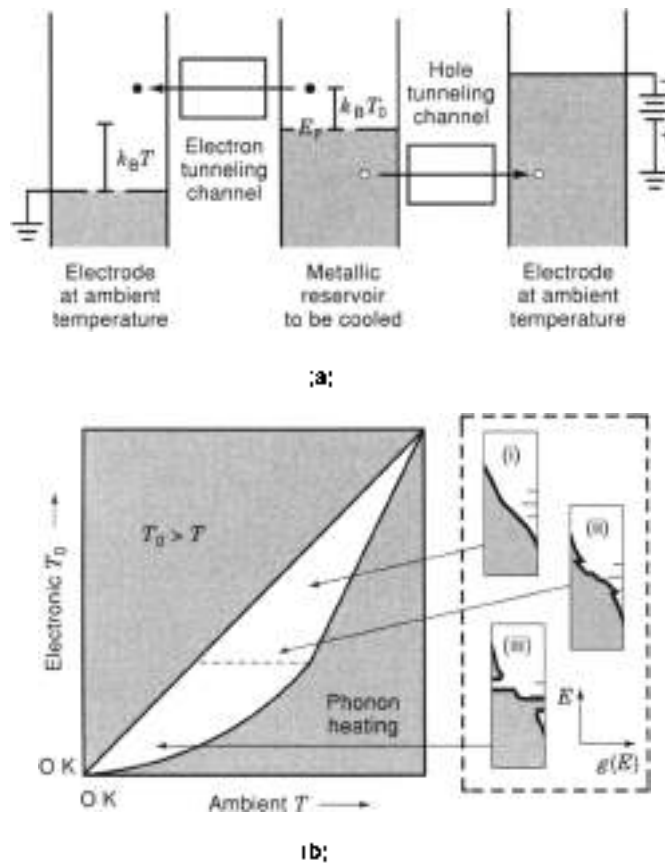
Because it is a vibration or deformation wave in the crystalline lattice, a phonon has a wavelength. At higher temperatures, the predominant wavelength—that of phonons that carry the thermal energy  $k_B T$ —is shorter. Near absolute zero, this phonon wavelength becomes so long that it barely interacts with electrons in the solid, analogous to the way that broad swells on the sea do not overturn a small boat as do abrupt breakers.

This "decoupling" of the electrons and the phonons in a solid enables the operation of certain devices based upon "ballistic" electron transport (12). It also allows the temperature of the electrons in small devices to be controlled independently of the phonon temperature. This is the basis of the first experimental demonstrations of cryogenic thermoelectric refrigeration, to be discussed in the following.

The structure and operation regime of a cryogenic thermoelectric refrigerator are depicted in Fig. 4(b). In Fig. 4(a), the metallic reservoir at the center is cooled by the removal of hot electrons and holes and their deposition on electrodes that reside at the "ambient" temperature, which here may refer to the phonon temperature or the true ambient temperature of the medium surrounding the refrigerator structure. Physically, such a refrigerator would probably be comprised of thin metal films or semiconductor layers on a substrate; the diagram of Fig. 4(a) shows electron energies, rather than physical structure.

In this light, the vertical position of an electron in Fig. 4(a) corresponds to its energy. The "sea" of electrons in the various structures corresponds to the Fermi sea of electrons in a metal. It is at the surface  $E_F$  of this Fermi sea that thermal excitations—in the form of thermally excited electrons and holes—are contained. Thus, it can be seen that a hole below the "sea level" of  $E_F$  carries the same amount of thermal energy as an electron excited an equal amount above  $E_F$ . If this Fermi sea of electrons is at temperature  $T$ , the average excitation energy of its electrons and holes from  $E_F$  is roughly  $k_B T$  (4). In addition, we use the term "hole" hereafter to refer to a vacant electronic state in a metal, which is more general than the definition used for  $p$ -type semiconductor material (3). This nomenclature simplifies the understanding of a thermoelectric device.

The voltage differences between different pieces of metal, such as the electrodes and the reservoir of Fig. 4(a), are reflected in differences between their Fermi levels. By applying a voltage to this structure, an electric current flows. The structures marked "electron tunneling channel" and "hole tunneling channel" will be explained in more detail in what follows, but the basic idea is the following. Because the temperature of an electron sea is defined by the energies to which electrons and holes are excited from the Fermi levels, removing the excited electrons and holes should cool the remaining electrons. The alternative to this selective process is to allow both electrons and holes to move arbitrarily through both interfaces, resulting in a loss of control over the direction of heat flow. The preferential



**Figure 4.** (a) Energy-level diagram and (b) temperature regime of operation of a cryogenic thermoelectric refrigerator. Reproduced from Ref. 5.

control over the direction of heat flow during electrical transport is the basis of thermoelectric refrigeration.

Thus, for thermoelectric refrigeration to be possible, it is required to preferentially remove excited charge carriers from near the Fermi level. The conventional way of doing this is to use a negative thermoelement, which carries thermally excited electrons away from the reservoir, and a positive thermoelement that removes thermally excited holes. The terms "negative" and "positive" arise from the sign of the charge carriers. Another interpretation of these terms is that, for a positive thermoelement, heat and electric current flow in the same direction, whereas the opposite is true for a negative thermoelement.

As was mentioned in the foregoing, an important difference between the thermal properties of matter at room temperature and cryogenic temperatures is the fact that the electrons hold most of the heat within 1 K of absolute zero. Combined with the fact that electrons and phonons and other degrees of freedom of a metal tend to decouple at cryogenic temperatures, this gives rise to some exotic possibilities for thermoelectric refrigeration. Figure 4(b) shows a diagram of the ambient temperature  $T$  versus the electronic temperature  $T_0$  for a "generic" thermoelectric refrigerator as shown in Fig. 4(a), in which we assume that the electrons are cooled independently of the phonons. There are several temperature regimes of interest.

In the upper left-half of the  $(T, T_0)$  diagram of Fig. 4(b), the electronic temperature exceeds the ambient; thus this region is not of interest in a discussion of refrigeration. In the lower

right-half, the electrons are cooler than the ambient, and there are situations one can imagine. First, if the ambient temperature  $T$  is too high, the electrons will absorb more energy from phonons than can be removed by the electron and hole tunneling channels, and hence the electrons will be heated. This situation is depicted for the hashed region in the lower right corner of the diagram.

The unhashed area to the center of the  $(T, T_0)$  diagram of Fig. 4(b) is thus the region in which the refrigerator is operational. In this region, there are several additional types of behavior, which come into play according to the rate at which the electrons can "thermalize," in other words, refill states emptied by the electron and hole tunneling channels. In regime (i), the electron distribution  $g(E)$  (shown as a function of electron energy  $E$  on the right-hand side of Fig. 4(b)), is continuous, and has a form expected for the thermally excited electron distribution near the Fermi level (8). However, as the temperature decreases (regime (ii)), the hot electrons and holes are removed at a rate approaching that at which the electrons scatter among themselves to smooth their energy distribution. Finally, at the lowest temperatures (regime (iii)), the bands of energy that are aligned with the electron and hole tunneling channels will be fully saturated by the cooling process. In essence, the electrons are not only decoupled from the phonons, but from each other, as far as the refrigeration process is concerned; and individual sets of energy states may be selectively cooled even when "hot" electrons and holes exist at other energies.

Figure 4(b) describes the case when the electrons and phonons are decoupled. To apply such a refrigerator, the "metallic reservoir to be cooled" would have to be able to exchange electrons with any heat load, which requires integral design and fabrication. Cryogenic thermoelectric cooling would be most useful when the electrons and phonons are coupled and the cooling power is strong enough to achieve bulk cooling. In the following, we will show two examples of cryogenic thermoelectric refrigerators and illustrate the temperature regime over which these limits apply.

#### THERMOELECTRIC REFRIGERATION USING QUANTUM-SCALE DEVICES

The promising nature of cryogenic thermoelectric refrigeration was suggested in 1993 (4), when it was proposed that quantum dots be used as thermoelements. Although this quantum-dot refrigerator has not yet been realized in practice, it is a good example because its efficiency can be shown to approach the Carnot limit (5). Figure 5(a) shows the physical structure and Fig. 5(b) shows the energy-level diagram of one such quantum-dot thermoelectric refrigerator. The individual electronic states of the quantum dots are separated by  $\Delta$ , but each is lifetime-broadened by  $\gamma$ . These are the only important properties of quantum dots for this device—the quantum dots may actually be larger or smaller than the reservoir  $R$  to be cooled, depending upon the material system and the temperature range of interest (6).

A quantum dot is a piece of metal so small that the individual quantum states, which electrons occupy, are separated by an energy comparable to or in excess of the thermal energy  $k_B T$ . In Fig. 5(b), the individual quantum states in the quantum dots are denoted by horizontal lines. One property of such states is that electrons and holes may travel through them from states



below 100 mK was demonstrated in an initial experiment (7). An extra NIS tunneling junction (not shown) was used as a thermocouple to measure electron temperature. In this experiment, the metal thin films comprising the refrigerator were deposited on a bulk substrate, and were patterned on the micron scale; thus it was assumed that the electrons were cooled but the phonons remained at 100 mK, due to the decoupling of electron and phonon degrees of freedom in small, cold metals.

More recently, cooling from 300 to 100 mK was achieved by an improved NIS refrigerator (8). This device was fabricated on a membrane (as depicted in Fig. 6a). Thus, there was not a strong thermal coupling to the ambient-temperature substrate, and the crystal lattice in the metal films comprising the refrigerator could be cooled along with the electrons, thereby reducing the heat leak due to phonon absorption. The feasibility of this experiment was theoretically predicted in Ref. (5). These devices are being developed to cool advanced radiation detectors for use in astronomy, but any electronic device that must operate at temperatures below 1 K could benefit from this type of refrigerator.

The significance of this result is that thermoelectric cooling has been achieved in which the kelvin temperature was reduced by a factor of 3. To accomplish this with a bulk material, the figure of merit would have to be in the range of  $ZT = 2$  to 4

(1,2). Thus, previous speculations (1,2) that  $ZT = 1$  represents a fundamental limit of physics may now be laid to rest.

## PROSPECTS AND APPLICATIONS FOR CRYOGENIC THERMOELECTRIC REFRIGERATION

The success of the NIS refrigerator indicates the promise that thermoelectric refrigeration holds for the sub-1 K regime. Any material interface or device structure, which conducts electric current but has a significant energy dependence, should be capable of thermoelectric refrigeration at some temperature scale (5). One particularly interesting material system is the III-V semiconductors, which may be formed into heterostructures with custom-tailored energy band structures and high carrier mobilities. These structures have been used for power devices, and thus can carry large electric currents, and various structures could be envisioned that would produce a significant cooling effect.

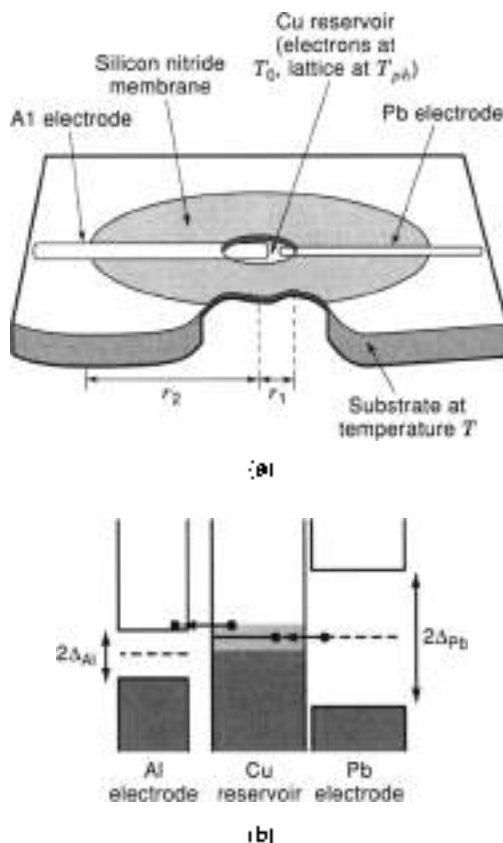
## CONCLUDING REMARKS

Widespread application of thermoelectric energy conversion remains a goal of materials science. While certain niche applications support a healthy, small Peltier-cooler industry, thermoelectric energy conversion will not compete economically with conventional methods until inexpensive, robust materials with a higher figure of merit are available. However, a good thermoelectric material would be very useful. For instance, as high-performance microelectronics continues to pack more transistors on a chip, heat loads are expected to exceed 100 W in several generations (14). This will require some form of direct cooling, and an efficient, low-cost thermoelectric alternative is attractive compared to water cooling, which may be hard to maintain reliably in a consumer end product.

One observation that leads to cautious optimism is that the present commercially available thermoelectric systems are based upon decades-old materials science developments. Many advances in computational materials modeling, materials synthesis (such as molecular-beam epitaxy), and materials characterization (such as high-resolution transmission electron microscopy and scanning-probe microscopy) have occurred in the meantime, enabling the exploration of complex material systems such as the high-temperature superconductors and colossal magnetoresistive materials. The skutterudites discussed in this article are an example of the fruits of modern materials science. As further complex material systems are explored, the right combination of low thermal conductivity, high electrical conductivity, and large thermopower may yet be found in an inexpensive, robust material.

New device concepts in thermoelectric conversion may overcome some traditional barriers as well. For instance, by using thermionic effects instead of reversible thermoelectric effects (6), enhanced efficiency might be achieved by high energy barriers and a departure from the traditional linear-response regime of thermoelectric behavior. Therefore, the cooling or power generation process would be far more efficient than would be expected from examination of the material figure of merit.

Cryogenic thermoelectric refrigeration represents a research frontier and, accordingly, applications are still in the research phase. However, although commercial applications remain distant, cryogenic thermoelectric refrigeration may



**Figure 6.** (a) Schematic and (b) energy level diagram of the normal-insulator superconductor (NIS) thermoelectric refrigerator. Hot electrons and holes in the Cu reservoir tunnel into quasi-particle states in the Al superconducting thermoelement, cooling  $R$ . The Pb contact electrode is used to inject charge into  $R$ , but has no net heating or cooling effect. The structure is shown fabricated on a silicon nitride membrane to minimize the heat leak due to substrate thermal conduction. Reproduced from Ref. (5).

provide an enabling technology for scientific work such as the mK-range radiation detector used in astronomy as discussed in this article. And, of course, if sufficient advantages are realized by cryogenic thermoelectric refrigeration, it may give rise to new commercial applications. For instance, liquid nitrogen is used to cool components in cellular-telephone base stations. If a thermoelectric refrigerator could be used to cool from 77 K to lower temperatures, new classes of electrical components could be utilized in widespread applications such as this.

With even greater extrapolation, one could imagine a cascaded series of thermoelectric refrigerators, for which the "hot side" is at room temperature and the "cold side" is at a cryogenic temperature. An example of an application where such a device might receive widespread commercial application in the future is in quantum computers, which should be capable of massively parallel computations but may have to be kept cold to preserve quantum phase coherence.

Cryogenic thermoelectric refrigeration may also lead to unexpected scientific discoveries; for instance, it is not known whether there exists a class of metals incapable of supporting a superconducting state or if all metals will go superconducting at a low enough temperature. At present, experiments have only been conducted to the range of  $10^{-5}$  K (15). Direct thermoelectric cooling may be able to cool metals further, perhaps answering this fundamental question in our understanding of the electronic structure of matter. Perhaps the most fundamental contribution of this field so far is that  $ZT = 1$  is not a universal physical limit, and that there may still be a highly efficient thermoelectric material system waiting to be discovered.

#### ACKNOWLEDGMENTS

We would like to acknowledge useful technical discussions with John Martinis (NIST, Boulder), Michael Nahum (Harvard University), and Paul Warburton (Kings College).

#### BIBLIOGRAPHY

1. R. B. Heikes and R. W. Ure, Jr., *Thermoelectricity: Science and Engineering*, New York: Interscience Publishers, 1961.
2. H. J. Goldsmid, *Thermoelectric Refrigeration*, New York: Plenum, 1964.
3. N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, Philadelphia: Saunders College, 1976.
4. H. L. Edwards, Q. Niu, and A. L. de Lozanne, A quantum-dot refrigerator, *Appl. Phys. Lett.*, **63**(13): 1815-1817, 27, 1993.
5. H. L. Edwards et al., Cryogenic cooling using tunneling structures with sharp energy features, *Physical Rev. B*, **52**(19): 5714-5736, 1995.
6. A. Shukcari and J. E.owers, Heterostructure integrated thermionic coolers, *Appl. Phys. Lett.*, **71**(9): 1234-1236, 1997.
7. M. Nahum, T. M. Eilev, and J. M. Martinis, Electronic microrefrigerator based on a normal insulator-superconductor tunnel junction, *Appl. Phys. Lett.*, **66**(24): 3123-3125, 1994.
8. M. M. Leivo, J. P. Pekola, and H. V. Averin, Efficient Peltier refrigeration by a pair of normal metal, insulator, superconductor junctions, *Appl. Phys. Lett.*, **68**(14): 1996-1998, 1996.
9. P. J. Lin-Chung and T. L. Rennerke, Thermoelectric figure of merit of cuprate superlattice systems, *Physical Rev. B*, **61**(19): 13244-13248, 1995; J. G. Solo and G. D. Mahan, Thermoelectric figure of merit of superlattices, *Appl. Phys. Lett.*, **65**(21): 2690-2692, 1994; L. D. Hicks, T. C. Harman, X. Sun, and M. S. Dresselhaus, Experimental study of the effect of quantum-well structures on the thermoelectric figure of merit, *Physical Rev. B*, **53**(16): R10493-R10496, 1996; L. D. Hicks and M. S. Dresselhaus, Effect of quantum-well structures on the thermoelectric figure of merit, *Physical Rev. B*, **47**: 12727, 1992.
10. B. C. Sales, H. Munzras, and R. K. Williams, Filled Skutterdite antimonides: A new class of thermoelectric materials, *Science*, **272**: 1325-1326, 1996.
11. A. Y. Chiu and J. R. Arthur, Molecular beam epitaxy, *Progress in Solid State Chemistry*, **10**: 157, 1979.
12. H. Ehrenreich and D. Turnbull (eds.), *Solid State Physics*, **44: Semiconductor Heterostructures and Nanostructures, New York: Academic Press, 1991; Supriyo Datta, *Electronic Transport in Mesoscopic Systems* (Cambridge Studies in Semiconductor Physics and Microelectronic Engineering, 1), Cambridge: Cambridge University Press, 1997.**
13. M. A. Reed et al., Observation of discrete electronic states in a zero-dimensional semiconductor nanostructure, *Physical Rev. Lett.*, **80**(10): 535-537, 1988.
14. *The National Technology Roadmap for Semiconductors*, to be published by the Semiconductor Industry Association, San Jose, CA.
15. Frank Paoli, Solid state physics at microkelvin temperatures: Is anything left to learn, *Phys Today*, **48**(1): 34-43, 1993.

#### Reading List

- R. B. Heikes and Roland W. Ure, Jr., *Thermoelectricity, Science and Engineering*, New York: Interscience Publishers, 1961
- H. J. Goldsmid, *Thermoelectric Refrigeration*, New York: Plenum, 1964
- N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, Philadelphia: Saunders College, 1976
- H. L. Edwards et al., Cryogenic cooling using tunneling structures with sharp energy features, *Physical Rev. B*, **52**(19): 5714-5736, 1995
- H. Ehrenreich and D. Turnbull (eds.), *Solid State Physics*, **44: Semiconductor Heterostructures and Nanostructures, New York: Academic Press, 1991**
- C. Wu, A Silent Cool, *Science News*, **162**(10): 152-153, September 6, 1997.
- C. Wood, Materials for thermoelectric energy conversion, *Rep. Prog. Phys.*, **15**(4): April 1986
- L. Kouwenhoven, Coupled quantum dots as artificial molecules, *Science*, **268**: 1440, 1995.
- A. Wapdarski, A superconducting siphon, *Nature*, **373**: 106, 1995.

HAR EDWARDS  
Texas Instruments, Inc.  
QIAN NIU  
ALEX de LOZANNE  
The University of Texas at Austin

# INDEX

- A-15 superconductors. *See also*  $\beta$ -Tungsten superconductors metallurgy, 188–190
- Abrikosov theory, type I and II superconductors, repulsion between vortices, 183–184
- Abrikosov vortex lattice, high-temperature superconductor magnetic properties, 261–262
- Absorption parameters, capacitor energy storage, 333–334
- Accelerating field measurements, superconducting cavity resonators, 529
- Ac losses, hysteresis in superconductors, theoretical background, 138
- Ac magnetic field, magnetic resonance and, 59–61
- Active magnetic bearings (AMBs), magnetic levitation in, 28–29
- Active microwave circuits:  
analog-to-digital converters, 161  
detectors and mixers, 161  
future applications, 161–162  
precision voltage and frequency sources, 161  
properties of, 159–160  
superconducting digital logic circuits, 160–161  
theoretical background, 152–153
- Active semiconductor sensors, temperature measurement, 609
- Adiabatic conservation, charged-particle motion, magnetic flux, 16
- Adiabatic demagnetization refrigerators (ADR):  
components of, 54–55  
historical background, 55  
properties of, 53–54  
refrigerant construction (salt pill), 56  
temperature regulation, 57–58
- Adiabatic (potted) magnets, superconducting magnet quenching propagation, 77–78
- Adiabatic protection criterion:  
electromagnetic superconductors, 111  
superconducting magnet quenching detection, 76–77
- Adiabatic stability:  
electromagnetic superconductors, 110  
flux jump stabilization, 319–320
- Air capacitors, properties of, 336
- Alnico, hard (permanent) magnets, 39
- Alternating current (ac) circuit:  
capacitor energy storage, 335  
magnetic flux, 13–15  
rotating magnetic field, 14  
sinusoidally wound stators, 14–15  
superconductor performance, 21–22  
voltage induction, 13–14
- magnetic levitation, 24
- superconducting quantum interference device (SQUID) for, susceptibility assessment, 525
- Alternating current (ac) Josephson effect:  
defined, 117  
high-temperature superconductors, properties of, 250–251  
international volt standard, 118  
tunneling devices, 164–166
- Alternating current (ac) loss:  
 $\mu$ SMES (superconducting magnetic energy storage) device, excess storage costs *vs.*, 515–516  
superconducting transformer design, 599–600  
superconductor metallurgy, critical current density ( $J_c$ ), 196
- Aluminum conduits, cable-in-conduit coil (CICC) systems, 233
- Alzheimer's disease, diagnostic imaging, 354–355
- Ambient pressure of oxygen (APO), YBCO superconductors, metal-doping effects, 270–271
- Ammeter, superconducting quantum interference device (SQUID) as, 523–524
- Amorphous magnetic materials, sensing properties, 447–448
- Ampère-Laplace law, bioelectromagnetic field calculations, 468
- Ampere's law, critical current density, magnetic flux line lattice, 95–97
- Ampere turns, thin detector solenoid design, 414–416
- Amplifiers:  
gyrotron technology, 399  
pyroelectric IR detectors, noise, 431
- Amplitude of magnetization, hysteresis loss, filament diameter, 141
- Analog-to-digital converter (ADC):  
active microwave circuits, 161  
Josephson junction, tunneling devices, 169–170  
Josephson junctions, 119  
rapid single-flux-quantum (RFQ) logic:  
advanced applications, 135–136  
theoretical background, 121–122
- Andreev reflection, Peltier refrigeration, 504–506
- Angiography:  
brain imaging, 354  
cardiac angiography, 357–358
- Angular momentum, electron paramagnetic resonance (EPR), 62
- Anisotropic magnetoresistance effect (AMR), read recording heads, 38–39
- Anisotropy:  
ferromagnetic materials, 31–32  
high-temperature superconductors, 263–264  
constituents in, 281  
hysteresis loss, 140–141
- Anomalous energy diffusion, Tokamak fusion reactor, 388
- Antenna, passive microwave circuits, 158–159
- Antiferromagnetism, principles, 34
- APC process, ductile alloys:  
characteristics of, 304–305  
research background, 295–296
- Arc-free current interrupter,  
superconducting magnet quenching,  
coil protection circuits, 67
- Arrhenius factor, flux creep and FLL melting, 105–106
- Auditory evoked fields, magnetic source imaging (MSI), 471–472
- Austenitic steel, cable-in-conduit coil (CICC) systems, 232–233
- Axial gradometers, neuromagnetic fields, 466
- Axons, magnetic source imaging (MSI), 467

- Azimuthal magnetization:**  
 hysteresis loss, transport currents, 143  
 superconducting magnetics, coil size mismatch, 585
- Azimuthal precompression, superconducting magnetics, 588**
- Bandpass filters, high-temperature superconductors (HTSs), microwave technology, 557–560**
- Bardeen-Cooper-Schreiffer (BCS) theory:**  
 critical current density, 94  
 high-temperature superconductors (HTSs), 260  
 Josephson junctions, 248–249  
 low-temperature superconductors (LTSs), 260  
 microwave technology materials, 153  
 origins of superconductivity, 3–4  
 Peltier refrigeration, 504–506  
 superconducting electronics development, 117  
 type I and II superconductors:  
 Cooper pairing, 181  
 pair-pair correlations, 181–182  
 phase locking, 182  
 physical characteristics, 180–181
- Barium ferrite, hard (permanent) magnets, 39**
- Barium strontium titanate pyroelectric detectors:**  
 ferroelectric materials, microwave technology, 491–492  
 properties of, 432
- Barkhausen effect:**  
 Barkhausen jump statistics, 49–50  
 domain wall-defect interaction, 44–47  
 equilibrium magnetization distribution, 44  
 experimental techniques, 50  
 flux pinning and losses in superconductors, 51–52  
 irreversible magnetization, 47–48  
 nondestructive testing, 50–51  
 principles of, 41–43  
 stress effects, 48  
 temperature dependence, 48–49
- Barkhausen jump:**  
 defined, 48–49  
 domain wall motion, 43  
 statistics of, 48–49
- Barkhausen noise:**  
 defined, 41  
 stress effects, 48  
 temperature dependence, 48–49
- Barriers, superconductor metallurgy, 195–200**  
 cabling degradation, 197  
 stress and strain, 197–200
- Batch-type magnetic separator, properties of, 453–454**
- Bead thermistors, semiconductor temperature sensors, 608**
- Beam compression, gyrotron systems, 397**
- Beam tunnel, gyrotron systems, 397–398**
- Bean critical state model:**  
 bulk superconductor structural characterization, 267  
 critical current density, critical state magnetization, 103–105  
 hysteresis loss:  
 accuracy of, 144  
 dc magnetization curve, 139  
 superconducting levitation, 566
- Bean-London model, Barkhausen effect, flux pinning and losses in superconductors, 52**
- Belted cables, energy storage capacitors, 338**
- Bending magnets, particle accelerators:**  
 coordinate system definitions, 577  
 normal dipole magnet, 577  
 normal quadrupole magnet, 577–578
- Beta decay, tritium fuel, 378**
- Biaxial texturing, high-temperature superconductors, wire and tape applications, 277**
- Bicrystal grain boundary junctions, high-temperature superconductors, 251–252**
- Bicrystal substrates, Josephson junctions, tunneling devices, SQUID applications, 168–169**
- Biepitaxial grain boundary junctions, high-temperature superconductors (HTSs), 252**
- Bioelectromagnetic studies:**  
 bioelectric *vs.* biomagnetic signals, 470–471  
 magnetic source imaging (MSI):  
 cellular sources, 467  
 electromagnetic calculations, 468  
 myocardium, 467–468  
 neurons, 467  
 numerical approaches, 468–469  
 primary current, 467  
 quasi-static approximation, 467  
 regularization and constraints, 470  
 superconducting quantum interference device (SQUID) for, 526–527
- Biopsy, breast cancer diagnostic imaging, 356–357**
- Biot-Savart law, electromagnetic superconductors, winding topology, 112**
- Bismuth high-temperature superconductors (HTSs):**  
 constituents in, 280–281  
 properties of, 110  
 structural properties, 284–285  
 substituents in, 293–294  
 thermoelectric refrigeration, 626–627
- Bismuth strontium calcium copper oxide (BSCCO) superconductor:**  
 cryogenic stabilization, 206  
 Josephson junctions, high-temperature superconductors (HTSs), 249  
 structural properties, 263  
 transformer design, 597–599
- Black body, radiation thermometry, 617**
- Bloch equations, magnetic resonance, 61**  
 diagnostic imaging, 349–350
- Bohr magneton, magnetic resonance, 60–61**
- BOLD (blood oxygen level dependent) process, functional magnetic resonance imaging (fMRI), 350–351**
- Bolometric detection:**  
 frequency converters and mixers, 372  
 hot electron bolometer mixer, 374–375  
 infrared radiation detector, 619  
 superconducting electronics and, 117  
 uncooled infrared detector arrays, 424–430  
 operating principles, 424–426  
 vanadium oxide microbolometer arrays, 424–426  
 YBaCuO microbolometer arrays, 428–430
- Boltzmann constant:**  
 direct current (dc) superconducting quantum interference device (dc SQUID), 519  
 Peltier effect, 500
- Bone tumors, diagnostic imaging, 361**
- Bootstrap current, Tokamak fusion reactor, 388**  
 plasma heating, 390–391
- Boulders, high-temperature superconductors, thin film development, 245**
- Boundary-element method (BEM), bioelectromagnetic field calculations, 468–469**
- Boundary-induced coupling currents (BICCs), coupling loss, 147**
- Boundary layers, cable-in-conduit conductor stability:**  
 supercritical helium, 309  
 transient heat transfer, 315–316
- Braided bundles, cable-in-conduit coil (CICC) systems, 230–232**
- Brain imaging:**  
 magnetic source imaging (MSI):  
 auditory evoked fields, 471–472  
 cortical rhythms and reactivity, 472–473  
 elliptical foci, 473  
 principles and techniques, 353–355  
 angiography, 354  
 cerebrovascular disease, 353  
 demyelinating disorders, 355  
 neurodegenerative disease, 354–355  
 tumors, 354
- Brandt model, flux line lattice, summation theory, 100**
- Breast cancer, diagnostic imaging, 356–357**
- Bridge circuits, temperature sensor resistance, 610–611**
- Brittleness:**  
 $\beta$ -tungsten superconductor metallurgy, 189–190  
 type II superconductors, cryogenic stabilization, 206
- Bronze manufacturing process,  $\beta$ -tungsten superconductors, 191**
- Bronze superconductors, properties of, 194–195**
- Bubble memory, intermediate (soft/hard) magnetic materials for, 40**
- Bulk flux pinning, critical current density, 102**
- Bulk silicon sensors, semiconductor temperature sensors, 608**
- Bulk superconductors:**  
 applications, 277–279  
 magnetic properties, 267



- TBCCO 2212 superconductors, 268–269  
 TI 1223 superconductors, 269–270  
 YBCO (123) superconductors, 267–268  
 metal-doping effects, 270–271  
 processing techniques, 264–266  
 TBCCO2212 and 1223 superconductors, 265–266  
 YBaCuO 123 conductors, 265  
 YBCO 123 bulk conductors, 265  
 structural properties, 266–267  
 TBCCO 2212 superconductors, 268–269  
 TI 1223 superconductors, 269–270  
 Bulk thermoelectric devices, thermoelectric conversion, 628
- Cable-in-conduit conductors (CICC):  
 comparison with other systems, 238  
 coupling current loss:  
 interstrand resistance, 148–149  
 transport currents, 147  
 electromagnetic superconductors, 109  
 forced flow design, 228–238  
 aluminum-based conduits, 233  
 braided bundles, 230–232  
 chromium electroplating, 229  
 coextrusion jacketing, 237–238  
 conduit materials, 232–233  
 double conduit jacketing, 238  
 historical background, 308  
 incoloy 908 conduit, 233  
 jacketing methods, 233–238  
 laser beam welding, 237  
 pullthrough jacketing, 235–237  
 stability calculations, 312–313  
 strand bundling, 229–230  
 strand coating, 228–229  
 supercritical helium stability margin, 308–311  
 superfluid helium stability margin, 311–312  
 tube mill jacketing, 234–235  
 superconducting magnet quenching  
 detection, advanced voltage sensors, 72–75  
 superconducting magnet quenching  
 propagation, 78–82  
 helium breakdown, 82–83  
 long coil solutions, 80–81  
 Shajii quench theory, 79  
 short coil solutions, 80  
 thermal-hydraulic quenchback (THQB), 81–82  
 time-dependent normal zone, 79  
 universal scaling regimes, 81  
 whole coil normal, 78–79
- Cabling degradation, superconductor metallurgy, 197
- Cabling techniques:  
 ductile superconductor metallurgy, 304  
 energy storage capacitors, 338
- Cadmium electroplating, cable-in-conduit coil (CICC) systems, 229
- Cadmium zinc telluride (CZT), digital x-ray systems and radiology, 347
- Capacitance:  
 capacitor energy storage theory, 332  
 rapid single-flux quantum (RFSQ) logic, 133–134
- Capacitor energy storage:  
 applications, 342–344  
 intermediate storage, 343  
 pulse power units, 343–344  
 transport applications, 342–343  
 capacitor classification, 336–337  
 electrical power supply systems, 330–331  
 power capacitors, 338–339  
 supercapacitors, 339–342  
 equivalent circuits, 341  
 integration process, 341–342  
 principles of, 340–341  
 theoretical background:  
 connection configuration, 334–335  
 dielectric material properties, 332–334  
 fundamental relationships, 331–332  
 intrinsic energy storage, 337–338  
 voltage, charge, and current configuration, 335–336
- Cardiac imaging, 357–358  
 coronary angiography, 357–358  
 coronary artery calcification detection, 358  
 magnetic source imaging (MSI):  
 distributed source reconstruction, 475  
 evoked fields, 475  
 ventricular preexcitation, 473–474  
 ventricular tachycardia, 474–475
- Carnot cycle:  
 Peltier refrigeration, 501–502  
 superconducting cavity resonators,  
 refrigerator requirements, 531–532  
 temperature sensors, 602–603  
 thermoelectric power generation, 627  
 thermoelectric refrigeration, 626–627
- Carousel electromagnets, separation properties, 458
- Casting process, electromagnetic levitation, 27
- Cathode-ray tubes, soft magnetic shields, 39
- Cationic compounds, high-temperature superconductors (HTS), structural properties, 282
- CBEMA curve, power quality data, 507–509
- Cellular structure, magnetic source imaging (MSI), 467
- Celsius scale, temperature sensors, 602–603
- Ceramic materials:  
 direct-current (dc) superconducting quantum interference device (dc SQUID), 127–128  
 energy storage capacitors, 337  
 high-temperature superconducting (HTS) electronics, 116  
 Josephson junctions, 249  
 rapid single-flux quantum (RFSQ) logic circuits, 122–123  
 superconducting fault current limiters (SFCL), 549
- Cerebrovascular diseases, brain imaging techniques, 353–354
- Charged coupled devices (CCDs), digital x-ray systems and radiology, 347
- Charged-particle motion, magnetic flux, 15–16
- Chebyshev filter, passive microwave circuits, 157–158
- Chemical process, magnetic separator equipment, 460
- Chest radiology, 355–356  
 lung cancer, 355  
 pulmonary embolism, 355–356
- Chevrel-phase superconductors, chemical composition, 187
- “Chirp” filters, passive microwave circuits, frequency agile devices, 158
- Circuit assembly, gyrotron systems, 398
- Circuit integration, Josephson junctions, 255–257  
 fabrication issues, 255–256  
 groundplanes, 256–257  
 multilayer circuit requirements, 255
- Circular iron yoke current-line distribution, superconducting magnetics, 581
- Closed cycle refrigeration, superconducting quantum interference device (SQUID), 522
- Closed field cycle, hysteresis loss, filament diameter, 141
- Closed-loop Hall effect current transducer, technology of, 406–407
- Closed magnetic field line confinement device, nuclear reactors, 385
- Coal purification, magnetic separator equipment, 462
- Coefficient of friction (COF),  
 superconducting levitation, 566–567
- Coercive field, ferromagnetic materials, 31, 33
- Coextrusion jacketing method, cable-in-conduit coil (CICC) systems, 237–238
- Coherence length:  
 critical current density ( $J_c$ ):  
 flux pinning, 98–99  
 properties, 92  
 high-temperature superconductors, 281  
 Josephson junctions, microbridges, 255  
 microwave technology materials, 153  
 superconductivity research, 3  
 type II superconductivity, magnetic flux, 20–21
- Coil protection circuits:  
 detector magnetics, 416–418  
 quench back, 417–418  
 quench protection dump resistor, 416–417  
 detector solenoid design, connections and leads, 419–420  
 fusion reactors, superconducting magnetics, 569–575  
 NMR magnets, 487–488  
 superconducting magnetics, 65–70  
 cold dump circuits, 68–69  
 cold dump switches, 69–70  
 dump resistors, 67–68  
 end design, 584, 590  
 external dump, 65–66  
 internal dumps, 68  
 interrupters, 66–67  
 multiple magnets, external quenching, 68  
 two-dimensional geometry, 582–584  
 thin detector solenoid design, 415–416

- Cold-worked microstructures:  
 ductile superconductor metallurgy, 300–302  
 niobium-titanium composite superconductor metallurgy, 297
- Collar pole unloading, superconducting magnetics, azimuthal precompression, 588
- Collective-pinning model, flux line lattice, summation theory, 100
- Collector device, gyrotron technology, 398
- Colorectal tumors, diagnostic imaging, 358–359
- Colossal magnetoresistance, double-exchange materials, 36
- Compact synchrotrons, design and operation, 543–544
- Complementary metal oxide semiconductor (CMOS), digital x-ray systems and radiology, 347
- Complementary output switching logic (COSL) gate, superconducting quantum interference device (SQUID), voltage-state logic circuits, 129–131
- Composite assembly, ductile superconductor metallurgy:  
 diffusion barriers, 299  
 stabilizer addition, 298–299  
 strand geometry and filament spacing, 299–300
- Composite dielectric materials, capacitor energy storage, 334
- Composite diffusion process,  $\beta$ -tungsten superconductor production, 190
- Composite superconductors:  
 cryogenic stabilization, Stekly criterion, 211–212  
 technical electromagnetic superconductors, 197
- Compositional phase diagram, high-temperature superconductors, thin film development, 242
- Compound action fields (CAF), magnetic source imaging (MSI), 475
- Computer-assisted tomography (CAT):  
 brain tumor imaging, 354  
 X-ray transmission computed tomography (X-ray CT), 346–347
- Conduction currents, capacitor energy storage, 334
- Conduit materials, cable-in-conduit coil (CICC) systems, 232–233
- Confinement technology:  
 fusion reactors, 380–381  
 dc electric fields, 380  
 gravitational fields, 380  
 inertial confinement, 380  
 magnetic field confinement, 380–381  
 RF electromagnetic fields, 380  
 tokamak fusion reactor:  
 anomalous diffusion, 388  
 enhanced confinement, 388  
 enhanced reverse shear mode, 389  
 particle collisions, 388  
 plasma confinement system, 387–388
- Constant electric field criterion, critical current density, resistive measurement, 93–94
- Constant resistivity criterion, critical current density, resistive measurement, 93–94
- Contact potential, thermoelectric refrigeration, 625–627
- Contact resistance, quantum Hall effect, 404
- Continuous electron beam accelerator facility (CEBAF), passive microwave circuit resonators, 157
- Continuous-type magnetic separator, properties of, 456–457
- Continuous wave operation, superconducting cavity resonators, 528–529
- Contrast agents, magnetic resonance imaging (MRI), 350
- Control electronics, direct-current (dc) superconducting quantum interference device (dc SQUID), 518–519
- Conversion loss, frequency converters and mixers, 366  
 Schottky diode mixers, 369
- Cooper pairing:  
 active microwave circuits, 159–160  
 frequency conversion and mixing, superconductor-insulator-superconductor (SIS) components, 373–374  
 high-temperature superconductors (HTSs):  
 microstructures, 287  
 substituents in, 287–288
- Josephson junction, 123–125  
 metal-insulator-metal (MIM) tunneling, 164  
 principles of, 248–249  
 proximity effect, 257–258  
 passive microwave circuits, transmission lines, 156
- Peltier refrigeration, 504–506
- superconducting cavity resonators:  
 radio frequency (RF) superconductivity, 529–530  
 thermal breakdown, 537–538
- type I and II superconductors:  
 BCS theory, 181  
 low-temperature alloys, 179–180  
 repulsion between vortices, 183–184
- Coplanar strip, ferroelectric materials, 493
- Coplanar waveguide, ferroelectric materials, 493
- Copper materials:  
 flux jump stabilization, 318  
 high-temperature superconductors (HTSs):  
 La-(Ba,Sr,Ca)-Cu-O and Nd(Ce)-CuO compounds, 282–283  
 structural properties, 262–263  
 pipes, flat cable soldering, strand soldering techniques, 223–224  
 YBCO superconductor substituents, 290–291
- Core interaction, critical current density ( $J_c$ ), flux pinning, 98–99
- Coronary artery calcification, diagnostic imaging, 358
- Cortical rhythms, magnetic source imaging (MSI), 472–473
- Cost optimization:  
 fusion reactors, superconducting magnetics, 575–576  
 superconducting transformer design, 600
- Coulomb barrier:  
 fusion reactors, 376  
 Hall effect transducers, 402–403
- Counterpulsed circuits, superconducting magnet quenching:  
 external dump circuits, 66  
 interruption switches, 67
- Coupling current losses in superconductors:  
 hysteresis and:  
 anisotropy and variable angle orientation, 140–141  
 current loss, 144–150  
 dc magnetization curve, 138–139  
 dc transport current, 142–143  
 estimation accuracy, 144, 150  
 filament diameter, 141  
 flat cable coupling loss, 147  
 formulas for, 139–140  
 interstrand resistance, CICC, 149  
 niobium-tin strands, 149  
 niobium-titanium strands, 149  
 saturation of coupling current, 146  
 self-field loss, 143–144  
 spatially charged magnetic field, 147–148  
 steady state coupling loss, 145  
 target interstrand resistance, CICC, 149–150  
 theoretical background, 138  
 transient coupling loss, 145–146  
 transport current coupling loss, 146–147  
 zero field crossing, 141–142  
 superconducting magnetics, 586–587
- Cowound sensors, superconducting magnet quenching detection, advanced voltage sensors, 72–75
- Critical current ( $I_c$ ):  
 active microwave circuits, 159  
 vs. critical current density, 93  
 cryogenic stabilization, 205–206  
 hysteresis loss, dc magnetization curve, 139  
 Josephson junctions:  
 high-temperature superconductors (HTSs), 257  
 metal-insulator-metal (MIM) tunneling, 164  
 principles of, 249–250  
 superconductor metallurgy, stress/strain characteristics, 197–200
- Critical current density ( $J_c$ ):  
 depairing phenomenon,  $J_c$  limits, 94  
 dissipation effects, high current density superconductors, 105–106  
 flux flow and resistive transition analysis, 105–106  
 temperature dependence, 105–106  
 ductile superconductor metallurgy, precipitation heat treatment, 302  
 flux jump stabilization:  
 small perturbation, 324–326  
 transport current equations and calculations, 317–318  
 flux pinning:  
 interactive forces, 97–99  
 scaling laws, 100–102  
 summation theory, 99–100

- high-temperature superconductors (HTSs):  
 bulk processing, 264–266  
 magnetic properties, 261–262  
 microstructures, 287  
 processing techniques, 264
- hysteresis loss:  
 anisotropy and variable angle orientation, 140–141  
 dc magnetization curve, 138–139  
 filament diameter, 141  
 magnetic flux line lattice, 94–97  
 magnetization model, 102–105  
 particle accelerators, superconductor magnetics, niobium-titanium composite superconductor, 579–580  
 properties of, 91–92  
 resistive measurement of, 92–94  
 superconducting fault current limiters (SFCL), 546–548  
 superconducting magnetics, 586  
 superconductor metallurgy:  
 cabling degradation, 197  
 stabilization and barriers, 195–200  
 $\beta$ -tungsten superconductors, 188–190  
 YBCO materials:  
 bulk superconductors, 267–268  
 superconductor substituents, 292–293
- Critical energy, synchrotron radiation, 542–543
- Critical magnetic field ( $H_c$ ):  
 critical current density ( $J_c$ ), 91–92  
 magnetic flux line lattice, 94–97  
 cryogenic stabilization, 205–206  
 depairing critical current density, 94  
 flux pinning, scaling laws, 100–102  
 high-temperature superconductors (HTSs), magnetic properties, 261–262  
 microwave technology materials, 153  
 superconducting cavity resonators, radio frequency (RF) superconductivity, 529–530  
 superconductor metallurgy, stress/strain characteristics, 198–200  
 $\beta$ -tungsten superconductors, 188–190  
 type I and II superconductors:  
 low-temperature alloys, 179–180  
 surface superconductivity, 184–185
- Critical stability curve, flux jump stabilization, 324–326
- Critical state model, superconducting magnetics, 585–586
- Critical temperature ( $T_c$ ):  
 critical current density ( $J_c$ ) and, 91–92  
 cryogenic stabilization, 205  
 high-temperature superconductors (HTSs), magnetic properties, 260–262  
 microwave technology, theoretical background, 152–153  
 superconductor magnetic cooling, 590  
 superconductor metallurgy, stress/strain characteristics, 198–200  
 thallate superconductors, 268–269  
 tungsten superconductor metallurgy, 188–190  
 YBCO bulk superconductors, 267–268
- Cross-bar matrices, Josephson junction tunneling devices, 170
- Cryogenic cooling:  
 high-temperature superconductors (HTSs), packaging design, 560–562  
 MRI magnets, 490  
 $\mu$ SMES (superconducting magnetic energy storage) device, 513–514  
 superconducting electronic components, 122–123, 123  
 superconducting fault current limiters (SFCL), 546, 549–550  
 superconducting transformers, 600  
 superconductivity research, 3  
 superconductor magnetics, 591  
 thermoelectric refrigeration:  
 applications, 631  
 future research issues, 632  
 principles of, 628–629  
 thin detector solenoid, 420–422
- Cryogenic current comparator (CCC) bridge, quantum Hall effect and, 410–411
- Cryogenic heat sink, detector solenoid design criteria, 419
- Cryogenic stabilization, superconductors:  
 advanced theory and thermal management problems, 215  
 early magnet applications, 206–207  
 electromagnetic superconductors, 110–111  
 protection criterion, 111–112  
 flux jump stabilization, 321  
 forced flow superconductors:  
 perturbation spectrum, 307–308  
 Stekly criterion, 306–307  
 historical background, 204–206  
 instability disturbances, 207–208  
 distributed heat source, 208  
 flux jumps, 207–208  
 mechanical disturbances, 208  
 safety issues, 216–217  
 techniques for, 208–215  
 vs. flux-jump stabilization, 208  
 Maddock-James-Norris equal-area theorem, 212–213  
 Stekly criterion, 209–212  
 Wipf minimum propagation zone (MPZ), 213–215  
 thermophysical and heat data, 215–216
- Crystalline anisotropy, high-temperature superconductors, 263–264
- Curie temperature:  
 ferroelectric materials, microwave technology, 491–492  
 ferromagnetic materials, 31  
 ferromagnetic resonance (FMR), 63–64  
 passive microwave circuits, frequency agile devices, 158
- Current decay, superconducting particle accelerators:  
 quench protection, 592  
 quench protection, magnet string, 592–593
- Current dipole model, bioelectromagnetic source modeling, 469
- Current drive, Tokamak fusion reactors, 390
- Current interruption switches, superconducting magnet quenching, coil protection circuits, 66–67
- Current leads:  
 $\mu$ SMES (superconducting magnetic energy storage) device, excess storage costs, 515  
 superconductor technology, 329
- Current-line distributions, superconducting magnetics, 581–582
- Current measurement, superconducting quantum interference device (SQUID) for, 523–524
- Current multipole expansion, bioelectromagnetic source modeling, 469
- Current operating margins, superconducting magnetics, 583–584
- Current sharing, cryogenic stabilization, Stekly criterion, 209–212
- Current shells:  
 flux jump stabilization and thickness of, 322–324  
 superconducting magnetics:  
 approximation techniques, 582–583  
 radial support, 589–590
- Current stability, flux jump stabilization, 324
- Current-temperature control equation, magnetic refrigeration, 57–58
- Current-voltage ( $I$ - $V$ ) characteristic:  
 active microwave circuits, 161  
 active semiconductor sensors, 609  
 critical current density, resistive measurement, 92–94  
 flux flow, resistive transition analysis, 105  
 frequency mixing:  
 Schottky diode mixers, 367–370  
 superconductor-insulator-superconductor (SIS) components, 373–374  
 hot electron bolometer (HEB) mixer, 374–375  
 Josephson effect, 251  
 Josephson junctions, 249–250  
 edge-geometry weak links, 254  
 step-edge grain boundary junctions, 252–253  
 negative resistance diode mixers, 372
- Cyclotron resonance maser (CRM)  
 instability, gyrotron technology, 396
- Cyclotrons:  
 magnetic confinement fusion, magnetized plasmas, 382  
 superconducting cyclotrons, theory and principles of, 540–542
- Cylinder model, hysteresis loss, accuracy of, 144
- Cylindrical void, superconducting magnet quenching, electrical field integrity, 85
- Damage-induced Josephson junctions, link configuration, 253
- Debye length, magnetic confinement fusion, 381
- Degeneracy factor, Peltier coefficient, 500–501
- Degradation phenomenon, cryogenic stabilization, early magnet applications, 206–207

- Delay lines:  
 high-temperature superconductors (HTSs), 560  
 passive microwave circuits, 159
- Demagnetization factor:  
 Barkhausen effect, 42  
 magnetic flux and, 9–10
- Demonstration poloidal coil (DPC-EX) coil:  
 fusion reactors, superconducting magnetics, 573–574  
 sheathed flat cable manufacturing, 227–228
- Demultiplexer (DEMUX) circuits, rapid single-flux-quantum (RSFQ) logic, advanced applications, 135–136
- Demyelinating disorders, diagnostic imaging, 355
- Depairing critical current density ( $J_D$ ), properties of, 94
- Deposition-interlayer Josephson junctions:  
 edge-geometry weak links, 254  
 planar/in-line junctions, 255  
 step-edge SNS junctions, 253–254  
 trilayer-geometry weak links, 254–255
- Detection circuits:  
 active microwave circuits, 161  
 Josephson junctions, tunneling devices, 169  
 superconducting magnet quenching detection, voltage sensors, 71–72  
 superconducting quantum interference device (SQUID), 520–521
- Detector magnetics:  
 cryogenic cooling, thin detector solenoid, 420–422  
 power supply and coil quench protection, 416–418  
 quench back, 417–418  
 quench protection dump resistor, 416–417  
 solenoid development, 412–413  
 coil electrical connections and leads, 419–420  
 cold mass support system, 418–419  
 cryogenic heat sink solenoid support, 419  
 detector solenoid end design criteria, 418–420  
 historical background, 412  
 low mass thin detector magnets, 412–413  
 thin solenoid design criteria, 413–416  
 thin solenoid parameters, 413
- Deuterium fuel, nuclear fusion, 377  
 reduced activation fuels, 379
- Diagnostic imaging:  
 brain imaging, 353–355  
 angiography, 354  
 cerebrovascular disease, 353  
 demyelinating disorders, 355  
 neurodegenerative disease, 354–355  
 tumors, 354  
 breast cancer, 356–357  
 cardiac imaging, 357–358  
 coronary angiography, 357–358  
 coronary artery calcification detection, 358  
 chest imaging, 355–356  
 lung cancer, 355  
 pulmonary embolism, 355–356  
 digital x-ray systems and radiology, 347  
 Doppler ultrasound, 348  
 electron beam computed tomography (EBCT), 347  
 endoscopy, 353  
 functional magnetic resonance imaging (fMRI), 350–351  
 gastrointestinal imaging, 358–359  
 colorectal tumors, 358–359  
 liver imaging, 358  
 genitourinary imaging, 359–360  
 female reproductive organ cancer, 359  
 fetus and pregnancy, 359  
 kidneys, 359  
 prostate, 359–360  
 interventional radiology, 353  
 magnetic resonance imaging (MRI), 348–350  
 magnetic resonance spectroscopy (MRS), 350  
 picture archiving and communication system (PACS), 362  
 positron emission tomography (PET), 352  
 radiological imaging, 353  
 radionuclide emission imaging (nuclear medicine imaging), 351–353  
 research overview, 345  
 single-photon computed tomography (SPECT), 352–353  
 skeletal system, 360–362  
 bone tumors, 361  
 joint disease, 360–361  
 osteoporosis, 361–362  
 vascular system, 362  
 vertebral spine diseases, 361  
 ultrasound, 347–348  
 X-ray imaging, 345–346  
 X-ray transmission computed tomography (X-ray CT), 346–347
- Diamagnetic currents:  
 magnetic flux:  
 conservation, 6–7  
 type II superconductivity, 21  
 magnetic levitation, 26  
 plasma physics, 383  
 YBCO superconductor substituents, 291
- Diamagnetic materials:  
 defined, 30  
 properties of, 30–31
- Dielectric constant:  
 capacitor energy storage, 333  
 thin-film superconductor substrates, 153–154
- Dielectric materials, capacitor energy storage:  
 absorption and hysteresis, 333–334  
 composite capacitors, 334  
 properties of, 332–333  
 strength, 333
- Diffusion barriers, ductile superconductor metallurgy, composite assembly, 299
- Digital circuits:  
 active microwave circuits, 160–161  
 Josephson junctions:  
 application of, 250  
 development of, 119–120  
 tunneling devices, 169–170  
 superconducting electronics:  
 advanced applications, 134–136  
 rapid single-flux-quantum (RSFQ) logic, 131–134  
 theoretical background, 119–123  
 voltage-state logic, 128–131
- Digital imaging and communications standard (DICOM), diagnostic imaging storage, 362
- Digital voltmeter, quantum Hall effect and, 411
- Digital x-ray systems and radiology, principles and applications, 347
- Dilution refrigerators, historical background, 55
- Diode mixers:  
 negative resistance diode mixers, 372  
 parametric frequency conversion, 371–372  
 properties, 363–364  
 topology and design issues:  
 double balanced mixers, 371  
 double-double balanced mixers, 371  
 image rejection/enhancement, 371  
 single balanced mixers, 370–371  
 single-ended mixer, 370  
 two-diode subharmonically pumped mixers, 371
- Diode temperature sensors, active semiconductor sensors, 609
- Dipole antennas, passive microwave circuits, 158
- Dipole magnets, particle accelerators, 577  
 structure and properties, 578–579
- Dipole symmetry current-line quadruplet, superconducting magnetics, 581
- Direct current (dc) differential amplifier, temperature sensor resistance, 611
- Direct current (dc) electric fields, fusion reactor confinement, 380
- Direct current (dc) electric generator:  
 hard (permanent) magnetic materials, 39–40  
 magnetic flux, 13
- Direct current (dc) Josephson effect:  
 defined, 117  
 high-temperature superconductors, properties of, 250–251
- Direct current (dc) magnetization curve, hysteresis loss, 138–139
- Direct current (dc) superconducting quantum interference device (dc SQUID):  
 control electronics, 518–519  
 historical and theoretical background, 117–118  
 input circuits, 520–522  
 Josephson junctions:  
 bicrystal grain boundaries, 252  
 grain boundaries, 251  
 tunneling devices, 167–169  
 magnetometer components, 126–128  
 operating principles, 518–519  
 sensitivity, 519
- Direct-current (dc) switches,  
 superconducting magnet quenching, external dump circuits, 66

- Direct current (dc) transport current:  
 capacitor energy storage, 335  
 hysteresis loss, 142–143
- Direct exchange, ferromagnetic materials, 31
- Direct/indirect drive, inertial confinement, fusion reactors, 392–393
- “Dirty limit”:  
 critical current density ( $J_c$ ), flux pinning, 98–99  
 Josephson junction, proximity effect, 257–258
- Discrete phase shifter, ferroelectric materials, 498
- Disk thermistor, semiconductor temperature sensors, 608
- Dislocations, Barkhausen effect, domain wall interaction, 46
- Dissipation effects:  
 high current density superconductors:  
 critical current density ( $J_c$ ), 105–106  
 flux flow and resistive transition analysis, 105  
 temperature dependence, 105–106  
 temperature sensor resistance, 609–611
- Distributed heat sources, cryogenic stabilization, 208
- Distributed source modeling, bioelectromagnetic fields, 469–470
- Dittus-Boelter form, cable-in-conduit conductor stability, transient heat transfer, 315–316
- Divided frequency, frequency converters and mixers, 363–364
- Domain wall, Barkhausen effect:  
 dislocation interaction, 46  
 equilibrium magnetization distribution, 44  
 inclusions and voids, pinning, 46–47  
 inhomogeneous stress fields, 45–46  
 irreversible magnetization, 47–48  
 jump statistics, 49–50  
 noise and stress effects, 48  
 pinning coercivity, 44–45  
 properties, 42–43  
 stress field inhomogeneity, 45–46  
 temperature dependence, 48–49
- Doppler radar, passive microwave circuit resonators, 157
- Doppler ultrasound:  
 breast cancer imaging, 356–357  
 principles and applications, 348  
 vascular system imaging, 362
- Double balanced mixers, topology and design, 371
- Double conduit jacketing, cable-in-conduit coil (CICC) systems, 238
- Double-double balanced mixers, topology and design, 371
- Double-exchange materials, properties of, 36
- Double sideband (DSB) noise, frequency converters and mixers, 366
- Downconversion:  
 frequency converters and mixers, 363–364  
 parametric frequency conversion, 371–372  
 two-terminal nonlinear mixer devices, 365
- Drag force, electrodynamic levitation, 25–26
- Drum and pulley magnetic separators, properties of, 457
- Ductile superconductor alloys, metallurgical research:  
 APC processing, 304–305  
 cabling procedures, 304  
 cold work processing, 300–302  
 composite assembly, 298–300  
 conventional processing, 297–298  
 extrusion, 300  
 final wire drawing, 303  
 niobium-titanium alloy system, 296–297, 304  
 precipitation heat treatment, 302  
 shaping procedures, 303  
 theoretical background, 295–296  
 twisting procedures, 303
- Dumping energy, superconducting magnet quenching, multiple magnets, 68
- Dump resistors:  
 quench protection dump resistor, 416–417  
 superconducting magnet quenching, coil protection circuits, 67–68
- Dump time constant, superconducting magnet quenching detection, adiabatic protection criterion, 76–77
- Dynamic errors, temperature sensors, 622
- Dynamic range, high-temperature superconductors (HTSs), passive microwave devices, 555–557
- Dynamic resistance, transport current coupling loss, 147
- Dynamic stability criterion, flux jump stabilization, 320–321
- Dynamo action, geomagnetism and, 10–11
- Earnshaw’s theorem:  
 diamagnetic materials, 31  
 magnetic levitation, 23  
 superconducting levitation, 564
- Eddy currents:  
 Barkhausen effect, irreversible magnetization, 48  
 magnetic flux conservation, 6–7  
 magnetic levitation, 23  
 ac levitation, 24  
 electrodynamic levitation, 24–25  
 magnetic sensors, 440  
 power transformers, 37  
 super-conducting quantum interference device (SQUID), refrigeration shielding, 523
- Edge-geometry weak links, Josephson junctions, 254
- Elastic behavior, critical current density, summation theory, flux pinning, 99–100
- Electrical current density ( $J_c$ ), Peltier effect, 500
- Electrical insulation, fusion reactors, superconducting magnetics, 575
- Electrical machines, magnetic flux, 11–15  
 ac generators, 13–15  
 dc generators, 13  
 transformers, 11–13
- Electrical metrology, quantum Hall effect, 408–411
- Electric field strength:  
 capacitor energy storage theory, 332  
 superconducting magnet quenching, 83–85
- Electric flux density, capacitor energy storage theory, 332
- Electric utility grids, superconducting fault current limiters (SFCL) in, 545–546
- Electrodynamic levitation, principles of, 24–25
- Electrodynamic systems (EDS):  
 maglev transport, 28  
 magnetic separation, 449
- Electrolytic capacitors, properties of, 337
- Electromagnetic interference, shielding and sensitivity devices, 448
- Electromagnetic systems (EMS):  
 maglev transport, 28  
 superconducting electromagnets:  
 forces and stress, 112–115  
 high-temperature superconductors, 109–110  
 properties of, 107–108  
 protection issues, 111–112  
 stability, 110–111  
 technical superconductors, 108–109  
 time-varying operation, 112
- Electromagnetic unit, magnetic materials, 30
- Electromagnetism, levitation, 23–24  
 casting applications, 27
- Electromotive force (emf):  
 Barkhausen effect, 50  
 thermocouple temperature sensors, 612–613
- Electron beam computed tomography (EBCT):  
 coronary artery calcification imaging, 358  
 principles and applications, 347
- Electron-beam-damaged weak links, Josephson junctions, 253
- Electron cyclotron current drive (ECCD), gyrotron technology, 394
- Electron cyclotron resonance heating (ECRH), gyrotron technology, 394  
 circuit assembly, 398
- Electron gun, gyrotron systems, 396–397
- Electronically conducting polymers (ECP), supercapacitor technology, 340
- Electron paramagnetic resonance (EPR):  
 experimental observations, 62–63  
 principles of, 61–62
- Electroplating assembly:  
 cable-in-conduit coil (CICC) systems, chromium electroplating, 229  
 smooth pipe conductors, 219–220
- Elemental superconductors, table, 2
- Elliptical inclusions, Barkhausen effect, domain wall pinning, 47
- Embedding circuits, frequency converters and mixers, large signal analysis, harmonic balance, 368
- Endoscopy, diagnostic applications, 353
- Energy balance:  
 cryogenic stabilization, Stekly criterion, 209–212  
 stability margin calculations in He-I, CICC conductors, 312–313
- Energy density:  
 fossil fuel vs. nuclear fuel, 377–378  
 magnetic flux and, 6

- Energy storage:  
 capacitor applications, 342–344  
   intermediate storage, 343  
   pulse power units, 343–344  
   transport applications, 342–343  
 capacitor classification, 336–337  
 electrical power supply systems, 330–331  
 $\mu$ SMES (superconducting magnetic energy storage) device:  
   components, 513–514  
   configuration, 511  
   design optimization, 516–517  
   design principles, 515  
   excess storage (cost) *vs.* current, 515  
   excess storage (cost) *vs.* power (ac losses), 515–516  
   negative attributes, 513  
   positive attributes, 512–513  
   power quality, 511–517  
   *vs.* other storage devices, 511–512  
 power capacitors, 338–339  
 supercapacitors, 339–342  
   equivalent circuits, 341  
   integration process, 341–342  
   principles of, 340–341  
 superconducting levitation, 567–568  
 superconducting magnetic energy storage (SMES) device:  
   capacitor energy storage, 331  
   high-temperature superconductor materials, 279  
   power quality, 509–517  
 theoretical background:  
   connection configuration, 334–335  
   dielectric material properties, 332–334  
   fundamental relationships, 331–332  
   intrinsic energy storage, 337–338  
   voltage, charge, and current configuration, 335–336
- Engineering critical current density ( $J_c$ ), properties of, 93
- Engineering superconductivity, origins of, 3–4
- Enhanced confinement technology, Tokamak fusion reactor, 388
- Enhanced reversed shear mode, Tokamak fusion reactor, 389
- Entropy, magnetic refrigeration, 54  
 operating principles, 56–57  
 theoretical background, 55–56
- Epileptic foci, magnetic source imaging (MSI), 473
- Epitaxy, high-temperature superconductors, thin film development, 242–243
- EPRI Distribution Power Quality (DPQ) study, power quality data, 508–509
- Equilibrium magnetization distribution, Barkhausen effect, 44  
 domain wall pinning coercivity, 44–45
- Equivalent circuits, supercapacitor technology, 341
- Equivalent current dipole (ECD), bioelectromagnetic source modeling, 469
- Equivalent noise temperature, Schottky diode mixers, 369–370
- Error detection, temperature sensors, 621–622
- dynamic errors, 621–622  
 steady-state errors, 621–622
- EU-LCT conductor, sheathed flat cable manufacturing, 224–225
- Evans hybrid capacitor, supercapacitor technology, 340
- Evaporation techniques, high-temperature superconductors, thin film development, 245–247
- Evershed bearings, permanent magnetic levitation, 23
- Evoked field imaging:  
 auditory evoked fields, 471–472  
 cardiac evoked fields, 475
- Excess storage costs,  $\mu$ SMES (superconducting magnetic energy storage) device, 515
- Ex-situ processing, HTS thin films, 271
- External diffusion process,  $\beta$ -tungsten superconductors, 192
- External dump circuits, superconducting magnet quenching, 65–66
- External quenching, superconducting magnet quenching, multiple magnets, 68
- Extrusion process, ductile superconductor metallurgy, 300
- Faraday effect:  
 Barkhausen effect, 42  
 dynamo action and geomagnetism, 11  
 magnetic levitation, 23  
 ac levitation, 24
- Faraday's law:  
 Barkhausen effect, 50  
 inductive magnetic sensors, 435–436  
 superconducting electronics theory, flux quantization, 126
- Fast switching matrices, Josephson junction tunneling devices, 170
- Fault current limiters:  
 cryogenics, 549–550  
 electric utility grid application, 545–546  
 ferromagnetic types, 548  
 quenching types, 546–548  
 semiconductor types, 548–549  
 superconductor requirements, 549  
 theoretical background, 544–545
- FDG positron emission tomography (FDG-PET):  
 breast cancer imaging, 357  
 chest imaging, lung cancer, 355
- Feedback control, electromagnetic levitation, 23–24
- Female reproductive organ cancer, diagnostic imaging, 359
- Fermi-Dirac distribution function, Peltier effect, 500–501
- Fermi energy:  
 cryogenic thermoelectric refrigeration, 628–629  
 Peltier refrigeration, 503–506  
 thermoelectric effect, 624
- Ferrimagnetic materials, 34–35  
 write heads, 38
- Ferroelectric lens antenna, microwave ferroelectric materials, 495–496
- Ferroelectric materials, microwave technology:  
 applications, 494–498  
 design models, 492  
 guiding structures, 492–494  
 high-temperature superconductors, 498  
 phased array antenna, 495–497  
 theoretical background, 491–492  
 traveling wave antenna, 497–498  
 tunable filter, 495  
 varactor device, 494–495  
 voltage-controlled oscillator, 495
- Ferrofluids, properties of, 41
- Ferromagnetic materials:  
 Barkhausen effect, 43  
 properties of, 31–32  
 shielding, 9  
 shims, magnet design, 485–486  
 superconducting fault current limiters (SFCL), 548
- Ferromagnetic resonance (FMR):  
 ferrimagnetism, 35  
 nonreciprocal materials, 40–41  
 principles of, 63–64
- Fetal imaging, techniques for, 359
- Fiberoptic temperature sensors, superconducting magnet quenching detection, 74–75
- Field-effect transistors (FET), Josephson junctions, tunneling devices, 172–173
- Field emission, superconducting cavity resonators, theoretical background, 534–536
- Field measurements:  
 magnetic resonance imaging magnets, 486–487  
 nuclear resonance magnets, 486–487
- Figure of merit:  
 Peltier refrigeration, 502–503  
 superconducting cavity resonators, 530  
 thermoelectric power generation, 627  
 thermoelectric refrigeration, 626–627
- Filament diameter:  
 ductile superconductor metallurgy, 299–300  
 hysteresis loss, 141  
 $\beta$ -tungsten superconductor metallurgy, 189–190
- Filters:  
 high-temperature superconductors (HTSs), microwave technology, 557–560  
 passive microwave circuits, 157–158  
 frequency agile devices, 158
- Finite-difference method, bioelectromagnetic field calculations, 469
- Finite-element method, bioelectromagnetic field calculations, 469
- Flash evaporation, high-temperature superconductors, thin film development, 245–246
- Flat cables:  
 coupling current loss, 147  
 strand soldering techniques, copper pipe soldering, 223–224
- Flip-flop circuits, rapid single-flux quantum (RFSQ) logic, 133–134
- Fluid dynamics, magnetic separation, 450–451

- Flux conversion, principles of, 10–11
- Flux creep:
  - dissipation effects, high current density superconductors, 105–106
  - electromagnetic superconductor, time-varying field operation, 112
  - high-temperature superconductors, magnetic properties, 262
  - superconducting electromagnets, 107–108
- Flux dams, Josephson junctions, tunneling devices, 168–169
- Flux flow regime:
  - critical current density:
    - flux line lattice (FLL) structure, 97
    - resistive measurement, 92–94
    - resistive transition analysis, 105
- Flux-gate magnetometers, properties of, 445
- Flux jump phenomenon:
  - cryogenic stabilization, 207–208
  - electromagnetic superconductors, adiabatic stability, 110
  - stability criteria:
    - adiabatic criterion, 319–320
    - cryogenic criterion, 321
    - dynamic criterion, 320–321
    - equations and calculations, 317–319
    - general stability criterion, 319
    - homogeneous composites, 321–322
    - nomenclature, 327
    - outer shell thickness and heat transfer, 322–324
    - sample determinations, 324
    - small perturbation, critical state model, 324–326
    - theoretical background, 317
- Flux line lattice (FLL):
  - critical current density, 94–97
  - critical state magnetization model, 103–105
  - summation theory, 99–100
  - critical state magnetization model, 102–105
  - melting, temperature dependence, 105–106
- Fluxoids, cryogenic stabilization, flux jumps, 207–208
- Flux pinning:
  - Barkhausen effect:
    - coercivity, 43–44
    - domain wall inclusions and voids, 46–47
    - losses in superconductors, 51–52
    - type II superconductors, 52
  - critical current density ( $J_c$ ):
    - interactive forces, 97–99
    - scaling laws, 100–102
    - summation theory, 99–100
  - superconductivity research, 4
  - superconductor metallurgy, heat-treatment principles, 200–201
- Flux quantization:
  - direct-current (dc) super-conducting quantum interference device (dc SQUID), 128
  - superconducting electronics theory, 125–126
- Flux vortex, critical current density, flux line lattice (FLL) structure, 96–97
- Flywheel energy storage, superconducting levitation, 567–568
- Focal plane array (FPA), bolometric detection, 426–428
- Focal source analysis, magnetic source imaging (MSI), 476
- Focused electron beam (FEB) writing, Josephson junctions, high-temperature superconductor (HTS) materials, 171
- Focused ion beam (FIB) junctions, weak link configuration, 253
- Forced flow superconductors:
  - cable-in-conduit conductors, 228–238
    - aluminum-based conduits, 233
    - braided bundles, 230–232
    - chromium electroplating, 229
    - coextrusion jacketing, 237–238
    - conduit materials, 232–233
    - double conduit jacketing, 238
    - historical background, 308
    - incoloy 908 conduit, 233
    - jacketing methods, 233–238
    - laser beam welding, 237
    - pulldown jacketing, 235–237
    - stability calculations, 312–313
    - strand bundling, 229–230
    - strand coating, 228–229
    - supercritical helium stability margin, 308–311
    - superfluid helium stability margin, 311–312
    - tube mill jacketing, 234–235
  - comparison of techniques, 238
  - historical background, 218
  - sheathed flat cables, 224–228
    - DPC-EX JAERI conductor, 227–228
    - ETL wind-and-react flat conductor, 226–227
    - EU-LCT conductor, 224–225
    - niobium-tin sheathed flat cables, 225–226
    - TMC-FF JAERI conductor, 227
  - smooth pipe conductors, 218–224
    - electroplating assembly, 219–220
    - strand soldering, central copper pipe, 220–224
    - superconductor pipes, 218–219
  - stability:
    - cable-in-conduit conductors, 308–314
    - energy balance, 312–313
    - He-II stability calculations, 314–315
    - He-I stability calculations, 312–313
    - one-dimensional model, 313–314
    - principles of, 306–308
    - Stekly criterion, 306–307
    - transient heat transfer, 315–316
    - vs. perturbation spectrum, 307–308
    - zero-dimensional model, 313
- Forcing magnets, particle accelerators:
  - coordinate system definitions, 577
  - normal dipole magnet, 577
  - normal quadrupole magnet, 577–578
- Fossil fuels, nuclear fuel, energy density comparisons, 377–378
- 4 x 4 packet switch, development of, 135
- Four-channel pass-band multiplexer, high-temperature superconductors (HTSs), filter design, 559–560
- Fourier conduction law, superfluid helium, 311
- Fourier transform, magnetic resonance imaging, 349–350
- Free energy, Barkhausen effect, 42–43
- Free space current-line distributions, superconducting magnetics, 581
- Frequency conversion, two-terminal nonlinear mixer devices, 364–365
- Frequency-dependent penetration depth, passive microwave circuits, 156
- Frequency mixing:
  - bolometer mixers, 372
  - conversion loss, 366
  - harmonic mixers, 371
  - high-temperature superconductors, nonlinearity, microwave interaction, 155–156
  - mixer receiver noise, 366
  - negative resistance diode mixers, 372
  - parametric frequency conversion, 371–372
  - Schottky diode mixers, 366–370
    - applications, 367–368
    - conversion loss, 369
    - equivalent noise temperature, 369–370
    - harmonic balance, large signal analysis, 368
    - port impedance, 369
    - small signal analyses, 368–369
  - self-oscillating mixers, 372
  - superconducting devices, 372–375
    - hot electron bolometer (HEB) mixer, 374–375
    - SIS mixer, 373–374
  - theoretical background, 363–364
  - topology and design issues:
    - double balanced mixers, 371
    - double-double balanced mixers, 371
    - image rejection/enhancement, 371
    - single balanced mixers, 370–371
    - single-ended mixer, 370
  - two-terminal nonlinear devices:
    - frequency conversion, 364–365
    - frequency multiplication, 364
    - harmonic mixing, 365
    - intermodulation, 365–366
    - linear mixing, 365
- Frequency sources, active microwave circuits, 161
- Fuel sources, fusion reactors, 376–380
  - deuterium fuel, 377
  - fossil fuel energy density comparisons, 377–378
  - lunar helium, 379–380
  - reduced activation fuels, 379
  - tritium fuel, 378
  - tritium production, 378–379
- Functional magnetic resonance imaging (fMRI), principles and applications, 350–351
- Fusion reactors:
  - confinement technology, 380–381
  - dc electric fields, 380
  - gravitational fields, 380
  - inertial confinement, 380

- Fusion reactors (*continued*)  
 magnetic field confinement, 380–381  
 RF electromagnetic fields, 380  
 fuel sources, 376–380  
 deuterium fuel, 377  
 fossil fuel energy density comparisons, 377–378  
 lunar helium, 379–380  
 reduced activation fuels, 379  
 tritium fuel, 378  
 tritium production, 378–379  
 future issues, 394  
 inertial confinement, 392–393  
 direct and indirect drive, 392–393  
 ion drivers, 393  
 laser drivers, 393  
 parameter regime, 392  
 plant systems, 393  
 magnetic confinement fusion, 381–387  
 nuclear fusion, 375–376  
 plasma heating with, 389  
 plasma physics, 381–387  
 closed magnetic field confinement devices, 385  
 cyclotron motion, 382  
 Debye length, 381  
 diamagnetism, 383  
 Larmor radius, 382–383  
 magnetic mirror confinement, 383–385  
 magnetic moments, 383  
 rotational transform, 385–387  
 tokamaks, 387  
 superconducting magnets:  
 coil structures, 569–575  
 cost optimization, 575–576  
 design issues, 575–576  
 electrical insulation design, 575  
 mirror devices, 569–570  
 quench protection, 575  
 risk and quality assurance, 576  
 stellarators, 574–575  
 theoretical background, 568–569  
 tokamaks, 570–574  
 tokamak fusion reactor, 387–392  
 anomalous diffusion, 388  
 bootstrap current, 390–391  
 confined particle collision, 388  
 current drive, 390  
 enhanced confinement, 388–389  
 enhanced reversed shear mode, 389  
 maintenance and materials, 392  
 neutral beam injection, 390  
 plasma confinement system, 387–388  
 plasma exhaust system, 391–392  
 plasma heating methods, 389–390  
 tritium breeding and thermal conversion blanket, 392
- Gadolinium compounds, magnetic resonance imaging (MRI), 350
- Gallium arsenide (GaAs) diodes, Schottky diode mixers, 367
- Gallium-doped tunnel barriers, Josephson junctions, ramp junctions, 171–172
- Gamma iron oxide, magnetic recording media, 40
- Garnet ferrites:  
 ferrimagnetic properties, 35  
 nonreciprocal materials, 41
- Gastrointestinal imaging:  
 colorectal tumors, 358–359  
 liver, 358
- Gay-Lussac's law, temperature sensors, 602–603
- Genitourinary imaging:  
 female reproductive organ cancer, 359  
 fetus and pregnancy, 359  
 kidney, 359  
 prostate, 359–360
- Geomagnetism, dynamo action and, 10–11
- Geometric errors, superconducting magnetics, 585
- Geophysical research, superconducting quantum interference device (SQUID) for, 526
- G function, electromagnetic superconductors, cryostable protection, 111–112
- Giant magnetoresistance (GMR):  
 properties of, 35  
 thin film magnetic materials, 36
- Gifford-McMahon/Boreas cryocooler:  
 high-temperature superconductors (HTSs), packaging design, 561–562  
 superconducting electronic components, 123  
 superconducting fault current limiters (SFCL), 550
- Ginzberg-Landau theory:  
 critical current density, 92  
 flux line lattice (FLL), 95–97  
 flux pinning, 98–99  
 superconducting cavity resonators, radio frequency (RF) superconductivity, 529–530  
 superconducting electronics, 3, 117  
 type I and II superconductors, 182–183  
 repulsion between vortices, 183–184  
 YBCO superconductor substituents, 292–293
- Graded ground planes, superconducting magnet quenching, helium breakdown, 88–89
- Gradiometers, superconducting quantum interference device (SQUID), 521–522
- Grad-Shafranov equation, magnetic flux, plasma equilibrium, 16–18
- Grain boundary Josephson junctions:  
 bicrystal junctions, 251–252  
 biepitaxial junctions, 252  
 high-temperature superconductors (HTSs), microstructures, 286–287  
 high- vs. low-temperature superconductors, 249  
 naturally occurring junctions, 251  
 step-edge junctions, 252–253  
 weak link configuration, 250  
 YBCO superconductor substituents, 292–293
- Grate-type magnetic separators, properties of, 457
- Gravitational fields, fusion reactor confinement, 380
- Groundplains, Josephson junctions, circuit integration requirements, 256–257
- Guard vacuum, superconducting magnet quenching, helium breakdown, 88–89
- Guiding structures, ferroelectric materials, 492–494  
 coplanar waveguide, strip, and slotline, 493  
 microstrip, 492–493  
 parallel plate and rectangular wave guide, 493–494
- Gyroklystron, amplifier design, 399
- Gyroscopic action, magnetic levitation, 26
- Gyrotron:  
 amplifiers, 399  
 components:  
 beam tunnel, 397–398  
 circuitry, 398  
 collector, 398  
 electron gun, 396–397  
 magnetic solenoids, 397  
 output taper, 398  
 output window, 399  
 RF launcher system, 398–399  
 schematic, 396–397  
 future research issues, 399–400  
 operating principles, 395–396  
 theory and applications, 394–395  
 Gyro-TWT device, amplifier design, 399
- Hall effect sensors:  
 applications:  
 clinical Hall effect devices, 407–408  
 linear Hall effect transducers, 405–407  
 research applications, 404–405  
 technical applications, 405–406  
 historical background, 401–402  
 multiplier applications, 407  
 principles of, 402–403  
 properties of, 447  
 quantum Hall effect, 403–404  
 contact resistance, 404  
 electrical metrology applications, 408–411  
 sample dependence, 403  
 semiconductor material dependence, 403
- Hamiltonian tunneling theory, Peltier refrigeration, 504–506
- Hard disks, intermediate (soft/hard) magnetic materials for, 40
- Hard (permanent) magnets, properties of, 39–40
- Harmonic balance, Schottky diode mixers, large signal analysis, 368
- Harmonic frequencies:  
 high-temperature superconductors, nonlinearity, microwave interaction, 155–156  
 magnet design, 481–484  
 shimming, 484–486
- Harmonic mixing:  
 two-diode subharmonically pumped mixers, 371  
 two-terminal nonlinear mixer devices, 365
- Harmonic sidebands, two-terminal nonlinear mixer devices, linear mixing, 365
- Heart imaging. *See* Cardiac imaging



- Heat current density ( $J_A$ ), Peltier effect, 500
- Heating time scale, cable-in-conduit conductor stability, supercritical helium, 310
- Heat leaks:
  - cryogenic stabilization, 208
  - $\mu$ SMES (superconducting magnetic energy storage) device, 514
- Heat sinks:
  - detector solenoid design criteria, cryogenic heat sink support system, 419
  - forced flow superconductor stability, 307–308
- Heat switch, adiabatic demagnetization refrigerators (ADR), 54–55
- Heat transfer phenomena:
  - cable-in-conduit conductor stability:
    - historical background, 308
    - supercritical helium, 310–311
    - transient heat transfer, 315–316
  - cryogenic stabilization, 215–216
  - flux jump stabilization, shell thickness, 322–324
  - forced flow superconductor stability, helium flow, 308
- Heat-treatment principles, superconductor metallurgy, 200–201
- Helium compounds. *See also* Supercritical and Superfluid helium (He-II)
  - cable-in-conduit conductor stability:
    - helium-I compounds, 312–314
    - helium-II compounds, 314–315
    - supercritical helium, 308–311
    - superfluid helium, 311–312
  - nuclear fusion:
    - lunar helium, 379–380
    - reduced activation fuels, 379
  - superconducting magnet quenching:
    - breakdown research, 88–89
    - electrical integrity, 82–83
- Hemorrhage, brain imaging techniques, 353–354
- HERA (Hadron Elektron Ring Anlage) superconducting particle accelerator:
  - cable insulation, 581
  - characteristics of, 578
- Hexagonal ferrites:
  - ferrimagnetism, 35
  - nonreciprocal materials, 41
- High current density superconductors, dissipation effects:
  - critical current density ( $J_c$ ), 105–106
  - flux flow and resistive transition analysis, 105
  - temperature dependence, 105–106
- High-current electromagnetic superconductors, 109
  - current leads, 329
- High-energy physics particle detector magnetics:
  - cryogenic cooling, thin detector solenoid, 420–422
  - power supply and coil quench protection, 416–418
  - quench back, 417–418
  - quench protection dump resistor, 416–417
- solenoid development, 412–413
  - coil electrical connections and leads, 419–420
  - cold mass support system, 418–419
  - cryogenic heat sink solenoid support, 419
  - detector solenoid end design criteria, 418–420
  - historical background, 412
  - low mass thin detector magnets, 412–413
  - thin solenoid design criteria, 413–416
  - thin solenoid parameters, 413
- High gradient magnetic separators (HGMSs):
  - historical background, 448–449
  - properties of, 452
- High-power-handling operations, high-temperature superconductors (HTSs), filter design, 560
- High pressure of oxygen (HPO), YBCO superconductors, metal-doping effects, 271
- High pulsed power RF processing (HPP), superconducting cavity resonators, 534–536
- High-temperature superconductors (HTSs):
  - anisotropy, 281
  - applications of, 277–279
  - bismuth HTSs, 110
  - coherence and London penetration depth, 281
  - constituents in, historical background, 280
  - critical current density:
    - resistive measurement, 93–94
    - scaling laws, flux pinning, 101–102
  - cryogenics and, 122–123
  - cryogenic stabilization, 206
  - techniques for, 208–209
  - thermal management problems, 215
  - thermophysical and heat data, 215–216
- direct-current (dc) superconducting quantum interference device (dc SQUID), 127–128
- electromagnetic properties, 109–110, 263–264
- ferroelectrics, 498
- film growth:
  - epitaxy, 242–243
  - flash evaporation, 245–246
  - historical background, 240–241
  - laser ablation, 244–245
  - materials, 241
  - molecular beam epitaxy, 246
  - postanneal techniques, 243–244
  - reactive coevaporation, 246
  - in situ techniques, 244, 246–247
  - sputtering, 244
  - thermodynamics, 241–242
- historical background, 260
- Josephson junctions, 125
  - ac and dc Josephson effects, 250–251
  - applications, 250
  - circuit integration, 255–257
  - configurations, 250
  - damaged-induced junctions, 253
  - deposition-interlayer junctions, 253–255
  - grain boundary junctions, 251–253
- historical background, 248
- microbridges, 255
- principles of, 248–249
- properties of, 249–250
- proximity effect, 257–258
- reproducibility limits, 259
- SNS resistance, 258–259
- SQUID applications, 168–169
- tunneling devices, 170–172
- vs.* LTS materials, 249
- YBCO materials, 249
- magnetic properties of, 260–262
- microstructures in, 285–287
- microwave technology materials, 153
  - cryogenic packaging, 540–562
  - delay lines, 560
  - film processing technology, 551–553
  - filter design, 557–560
  - future applications, 161–162
  - nonlinearity properties, 155
  - passive microwave circuits, 158, 553–557
- $\mu$ SMES (superconducting magnetic energy storage) device, 514
- particle accelerators, superconductor magnetics, 580
- processing techniques, 264
- rapid single-flux quantum (RFSQ) logic circuits, 122–123
- recent developments and applications, 116–123
- reciprocating magnetic separators, 458
- solenoid electromagnetics, 456
- structural properties, 262–263, 281–285
  - Bi-, Tl-, and Hg-based compounds, 284–285
  - La-(Ba,Sr,Ca)-Cu-O and Nd(Ce)-CuO compounds, 282–283
  - YBaCuO compounds, 283–284
- substituents in, 287–294
  - Bi-, Tl-, and Hg-based compounds, 293–294
  - La-(Ba,Sr,Ca)-Cu-O and Nd(Ce)-CuO compounds, 288–290
  - oxygen and oxygen vacancies, 288
  - YBCO compounds, 290–293
- superconducting cavity resonators, future applications, 538–539
- superconducting fault current limiters (SFCL), 544–546
  - ceramic materials, 549
  - quenching applications, 547–548
- superconducting magnet quenching propagation, 77–78
- superconducting quantum interference device (SQUID), detection coils, 520–521
- superconductivity research, 4
- technical electromagnetic superconductors, 108
- thin film magnetic materials:
  - epitaxy, 242–243
  - flash evaporation, 245–246
  - historical background, 240–241
  - laser ablation, 244–245
  - materials, 241
  - metal-organic chemical vapor deposition, 273–274

- High-temperature superconductors (HTSs)  
(*continued*)
- molecular beam epitaxy, 246
  - physical/chemical deposition techniques, 271, 273
  - postanneal techniques, 243–244
  - processed film characterization, 274
  - pulsed laser deposition, 273
  - reactive coevaporation, 246
  - in situ techniques, 244, 246–247
  - in situ vs. ex-situ* processing, 271
  - sputtering, 244, 273
  - structural, electrical, and mechanical properties, 274–276
  - thermodynamics, 241–242
  - transformer design, 596–597
    - alternating current (ac) loss, 599–600
    - cryostat requirements, 600
    - limitations of, 598–599
    - state-of-the-art issues, 600–601
  - type I and II superconductors, 180
  - wires and tapes:
    - first generation processes, 276
    - OPIT method, 276
    - second generation, YBCO conductors, 276–277
  - yttrium HTSs, 110
- High voltage direct current (HVDC)  
transmission, superconducting fault current limiters (SFCL), 545
- Hohlraum, direct/indirect drive, inertial confinement, fusion reactors, 393
- Holding magnets, hard (permanent) magnetic materials, 40
- Homogeneous composites, flux jump stabilization, 321–322
- Hooge parameter, bolometric detection, 425
- “Hoop stress,” superconducting cyclotrons, 541–542
- Hospital information systems (HIS), diagnostic imaging storage, 362
- Hot electron bolometer (HEB) mixer, components, 374–375
- Hot electron noise, Schottky diode mixers, 367–368
- Hot junctions, infrared radiation detector, 618–619
- Hot spot temperature, quench protection, superconducting particle accelerators, 592
- Huntington’s disease, diagnostic imaging, 355
- HYPRES voltage standard system, development of, 118–120
- Hysteresis:
- Barkhausen effect, 43
    - domain wall pinning connectivity, 45
    - irreversible magnetization, 47–48
    - jump statistics, 50
  - capacitor energy storage, 333–334
  - coupling losses in superconductors:
    - anisotropy and variable angle orientation, 140–141
    - current loss, 144–150
    - dc magnetization curve, 138–139
    - dc transport current, 142–143
    - estimation accuracy, 144, 150
    - filament diameter, 141
    - flat cable coupling loss, 147
    - formulas for, 139–140
    - interstrand resistance, CICC’s, 149
    - niobium-tin strands, 149
    - niobium-titanium strands, 149
    - saturation of coupling current, 146
    - self-field loss, 143–144
    - spatially charged magnetic field, 147–148
    - steady state coupling loss, 145
    - target interstrand resistance, CICC’s, 149–150
    - theoretical background, 138
    - transient coupling loss, 145–146
    - transport current couplin loss, 146–147
    - zero field crossing, 141–142
  - critical current density, critical state magnetization model, 102–105
  - cryogenic stabilization, 208
  - electromagnetic superconductor, time-varying field operation, 112
  - magnetic flux:
    - ac conditions, superconductor performance, 21–22
    - type II superconductivity, 21
  - niobium-titanium composite superconductor, 109
  - superconducting levitation, 564–565
  - temperature measurement, 605
- Hysteretic Josephson junction, digital superconducting electronic circuits, 128–131
- Ideal transformers, properties of, 12
- Ill-cooled regimes, stability margin calculations in He-I, CICC conductors, energy balance, 312–313
- Image frequency, two-terminal nonlinear mixer devices, harmonic sidebands, 365
- Image rejection and enhancement, mixer topology and design, 371
- Inclusions, Barkhausen effect, domain wall pinning, 46–47
- Incoloy 908 conduit, cable-in-conduit coil (CICC) systems, 233
- Indirect exchange, ferromagnetic materials, 31
- Induced-pole magnetic separators, properties of, 455
- Induced roll separators, properties of, 457
- Inductance:
  - magnetic flux and, 6
  - superconducting magnet quenching, external dump circuits, 65–66
- Induction potentiometer, properties, 439–440
- Inductive magnetic sensors, 435–440
  - induction potentiometer, 439–440
  - linear and rotary variable-reluctance sensors, 436
  - linear variable inductor, 439
  - microsyn, 438
  - single-coil variable-reluctance sensor, 436–437
  - synchros, 438–439
- variable-differential reluctance sensor, 437
- variable-reluctance tachogenerators, 437–438
- Inductors, soft magnetic materials, 37–38
- Industrial immersion thermocouple sensors, 615
- Industrial metal resistance thermometers, temperature measurements, 605–607
- Inertial confinement, fusion reactors, 392–393
  - direct and indirect drive, 392–393
  - ion drivers, 393
  - laser drivers, 393
  - parameter regime, 392
  - plant systems, 393
  - theory and technology, 380
- Infinite slab model, hysteresis loss, accuracy of, 144
- Infrared detector arrays:
  - radiation thermometry, 618–619
  - uncooled systems:
    - bolometric detectors, 424–430
    - pyroelectric IR detectors, 430–433
    - theoretical background, 424
- Inhomogeneous magnetic field, charged-particle motion, 16
- In-line junctions:
  - high-temperature superconductors, 255
  - superconducting fault current limiters (SFCL), quenching operations, 546–548
- Input circuits, super-conducting quantum interference device (SQUID), 520–522
- Insertion loss, high-temperature superconductors, nonlinearity, microwave interaction, 155–156
- In situ casting process:
  - HTS thin films, 271
  - $\beta$ -tungsten superconductors, 192–193
- In situ growth techniques, high-temperature superconductors, thin film development, 244
- In situ monitoring, high-temperature superconductors, thin film development, 246–247
- Instability sources, cryogenic stabilization, 207–208
  - distributed heat source, 208
  - flux jumps, 207–208
  - mechanical disturbances, 208
- Insulating surface tracking, superconducting magnet quenching, 86–88
- Insulation design:
  - fusion reactors, superconducting magnetics, 575
  - particle accelerators, superconductor magnetics, 579–580
  - cable insulation, 580–581
- Integrated circuits:
  - active semiconductor sensors, 609
  - Hall effect transducers, 401–402
- Integrated sensors (IS), temperature measurement, 604–605
- Integrated volt-seconds, superconducting magnet quenching detection, advanced voltage sensors, 74

- Intermediate energy storage, capacitor technology, 343
- Intermediate frequency (IF), frequency converters and mixers:  
conversion loss, 366  
theoretical background, 363–364
- Intermediate state:  
magnetic flux, superconductivity, 19  
type I and II superconductors, 178
- Intermodulation:  
high-temperature superconductors (HTSs), passive microwave devices, 557  
two-terminal nonlinear mixer devices, 365–366
- Internal dump resistors, superconducting magnet quenching, 68–70  
cold dump circuits, 68–69  
cold switches, 69–70
- Internally cooled superconductors (ICS), cable-in-conduit conductors (CICC) design, 308
- Internal sensor termination, superconducting magnet quenching detection, advanced voltage sensors, 73–74
- Internal tin production process,  $\beta$ -tungsten superconductors, 191–192
- International temperature scale 90 (ITS-90), temperature sensors, 603–604
- International volt standard, ac Josephson effect, 118
- Interrupters, superconducting magnet quenching, coil protection circuits, 66–67
- Interstrand coupling currents, superconducting magnetics, 586–587
- Intervalley scattering noise, Schottky diode mixers, 367–368
- Intrinsic energy storage, capacitor properties, 337–338
- Intrinsic stability property, flux jump stabilization, 317
- Inverse problems, bioelectromagnetic source modeling, 469
- Ion-beam-assisted deposition (IBAD), high-temperature superconductors:  
thin film development, 243  
wire and tape applications, 277
- Ion-damaged weak links, Josephson junctions, 253
- Ion drivers, inertial confinement, fusion reactors, 393
- Iron core structure, superconducting transformer design, 597
- Iron yoke current-line distribution, superconducting magnetics:  
coil geometry, 583  
iron saturation, 585  
principles of, 581  
radial support, 588–590
- Irreversibility field, high-temperature superconductors, magnetic properties, 262
- Irreversible magnetization, Barkhausen effect, 47–48
- Isothermal demagnetization, magnetic refrigeration, 57
- Jacketing methods:  
cable-in-conduit coil (CICC) systems, 233–238  
coextrusion method, 237–238  
double conduit method, 238  
laser beam welding, 237  
pullthrough method, 235–237  
tube mill, 234–235  
cable-in-conduit conductors (CICC), historical background, 308
- Jelly roll process,  $\beta$ -tungsten superconductors, 192
- Johnson noise:  
bolometric detection, 425  
infrared radiation detector, 618–619  
pyroelectric IR detectors, 431
- Joint disease, diagnostic imaging, 360–361
- Joint extraction, superconducting magnet quenching detection, advanced voltage sensors, 73
- Josephson effect. *See also* Ac and Dc Josephson effects  
active microwave circuits, 159  
Bardeen-Cooper-Schreiffer (BCS) theory, 260  
basic principles of, 123–125  
high-temperature superconductors, properties of, 250–251  
quantum Hall effect and, 410–411  
superconducting electronics, 117  
superconducting quantum interference device (SQUID) magnetometers, 446  
principles of, 517–518  
superconductivity research, 3–4  
tunneling devices:  
alternating current Josephson effect, 164–165  
theoretical background, 163
- Josephson junctions:  
active microwave circuits, 159–160  
detectors and mixers, 161  
superconducting digital logic circuits, 160–161  
Bardeen-Cooper-Schreiffer (BCS) theory, 260  
basic principles of, 123–125  
digital circuit development, 119–120  
direct-current (dc) superconducting quantum interference device (dc SQUID), 126–128  
operating principles, 518–519  
fabrication materials, 122–123  
high-temperature superconductors (HTSs):  
ac and dc Josephson effects, 250–251  
applications, 250  
circuit integration, 255–257  
configurations, 250  
damaged-induced junctions, 253  
deposition-interlayer junctions, 253–255  
grain boundary junctions, 251–253  
historical background, 248  
microbridges, 255  
principles of, 248–249  
properties of, 249–250  
proximity effect, 257–258  
reproducibility limits, 259  
SNS resistance, 258–259  
vs. LTS materials, 249  
YBCO materials, 249
- hysteretic Josephson junction, digital superconducting electronic circuits, 128–131
- low-probability punch through, voltage-state logic, 131
- radio frequency superconducting quantum interference device (RF SQUID), 518
- rapid single-flux quantum (RFSQ) logic, 131–134
- superconducting electronics, theoretical background, 117
- superconducting quantum interference device (SQUID):  
historical and theoretical background, 117–118  
principles of, 517–518
- tunneling devices:  
alternating current Josephson effect, 164–166  
digital applications, 169–170  
high-temperature superconductor systems, 170–172  
magnetic-field effects, 166  
metal-insulator-metal (MIM) tunneling, 163–164  
mixing and detection, 169  
related superconductive devices, 173  
resistively shunted junction (RSJ) model, 166–167  
superconducting quantum interference devices (SQUID), 167–169  
theoretical background, 163  
three-terminal devices, 172–173
- Josephson transmission line (JTL), rapid single-flux quantum (RFSQ) logic, 132–134
- Josephson vortex, critical current density, flux pinning, 99
- Joule heating:  
cable-in-conduit conductor stability:  
supercritical helium, 310  
superfluid helium (He-II) stability margin, 314–315  
capacitor energy storage theory, 332  
cryogenic stabilization, 207, 209  
Maddock-James-Norris equal-area theorem, 212–213  
Stekly criterion, 209–212  
flux jump stabilization, 319  
Peltier effect, 499–500  
quench protection, superconducting particle accelerators, 591
- Jumping ring experiments:  
Barkhausen effect, flux pinning and losses in superconductors, 51–52  
magnetic levitation, ac levitation, 24
- Jumpsum, Barkhausen noise, 48
- Kaolin processing, magnetic separators, 458–460
- Kapitza resistance, cable-in-conduit conductor stability, transient heat transfer, 315–316
- Kelvin relations:  
Peltier effect, 500

- Kelvin relations (*continued*)  
 temperature sensors, 602–603  
 thermoelectric refrigeration, 626–627
- Kerr rotation, magnetic recording materials, 40
- Kidney, diagnostic imaging, 359
- Kirchhoff's laws:  
 Josephson junctions, 124  
 radiation thermometry, 617–618
- Kramer's law, superconductor metallurgy, 201–203
- KTY linear sensors, semiconductor temperature sensors, 608
- La-(Ba,Sr,Ca)-Cu-O and Nd(Ce)-CuO compounds, high-temperature superconductors (HTSs):  
 structural properties, 282–283  
 substituents in, 288–290
- Lambda temperature ( $T_\lambda$ ):  
 superconductor magnetic cooling, superfluid helium (He-II), 590–591  
 superfluid helium (He-II) stability, CICC conductors, 311–312
- Landau-Ginzberg theory. *See* Ginzberg-Landau theory
- Landau-Lifshitz damping model, ferromagnetic resonance (FMR), 63–64
- Langevin equation:  
 Barkhausen effect, jump statistics, 49–50  
 paramagnetic materials, 31
- Lanthanide-aluminum-oxygen (LAO) substrates, high-temperature superconductors (HTSs), passive microwave devices, 555
- Laplace transform:  
 magnet design, 482–484  
 temperature sensor dynamic errors, 622
- Large coil task (LCT):  
 fusion reactors, superconducting magnetics, 572–574  
 sheathed flat cable manufacturing, EU-LCT conductor, 224–225  
 strand soldering techniques, 222–223  
 superconducting magnets, 115
- Large Hadron Collider (LHC):  
 cable insulation, 581  
 superconducting magnets, 114–115, 578
- Large helical device (LHD), fusion reactors, superconducting magnetics, stellarator structure, 574–575
- Large signal analysis, Schottky diode mixers, harmonic balance, 368
- Larmor frequency:  
 charged-particle motion, magnetic flux, 16  
 electron paramagnetic resonance (EPR), 62–63  
 magnetic resonance and, 59–61
- Larmor precession:  
 electron paramagnetic resonance (EPR), 62  
 magnetic resonance and, 59–61  
 nuclide constants, 479
- Larmor radius:  
 gyrotron technology, 395–396  
 magnetic moment, 383
- magnetized plasmas, 382–383
- Laser ablation, high-temperature superconductors, thin film development, 244–245
- Laser beam welding, cable-in-conduit coil (CICC) systems, 237
- Laser drivers, inertial confinement, fusion reactors, 393
- Latticebraids, cable-in-conduit coil (CICC) systems, 230–232
- Lattice structure,  $\beta$ -tungsten superconductors, 188–190
- LC circuit, magnetic levitation, 26
- Lead field, bioelectromagnetic source modeling, 470
- Lead resistance, temperature sensors, 610
- Lead titanate-based detectors, pyroelectric IR detectors, 431–432
- Leakage flux:  
 capacitor energy storage, dielectric leakage and currents, 334  
 properties of, 12  
 superconducting magnet quenching, insulating surface tracking, 86–88
- Legendre functions, magnet design, 482–484
- Lenz's law:  
 magnetic flux principles, 5  
 conservation, 6–7  
 transformers, 11–13  
 magnetic levitation, 23  
 plasma diamagnetism, 383
- Levenberg-Marquardt algorithm, bioelectromagnetic source modeling, 469
- Levitation:  
 magnetic levitation, pressure magnitude, 26–27  
 superconducting levitation:  
 applications, 567–568  
 bearing properties, 566–567  
 levitational phenomenology, 565–566  
 operating principles, 564–565  
 stable levitation, 563–564  
 theoretical models, 566
- Levitron magnetic levitation toy, 26
- Lifting magnetic separators, properties of, 457
- Limiting operating current ( $I_{lim}$ ):  
 cable-in-conduit conductor stability, supercritical helium, 309–310  
 stability margin calculations in He-I, CICC conductors, energy balance, 313
- Linear Hall effect transducers:  
 electrical current measurement, 405–406  
 magnetic field measurement, 405
- Linearity, temperature sensor resistance, 610
- Linear mixing, two-terminal nonlinear mixer devices, 365
- Linear variable-differential transformer (LVDT), magnetic sensing, 441–444
- Linear variable inductor, magnetic sensing, 439
- Linear variable-reluctance sensors, properties of, 436–437
- Lines of force (LOF), magnetic flux line and tube, 8
- Liquid helium:  
 cryogenic stabilization, safety issues, 216–217  
 superconductivity research, 1
- Liquid-phase epitaxy (LPE) garnet film, bubble memories, 40
- Lithium, nuclear fusion, tritium production, 378–379
- Liver, diagnostic imaging, 358
- Local oscillator (LO):  
 diode mixer topology and design, 370–371  
 frequency converters and mixers:  
 large signal analysis, harmonic balance, 368  
 theoretical background, 363–364  
 frequency multiplication, 365  
 harmonic mixers, 371  
 intermodulation, 365–366  
 parametric frequency conversion, 371–372
- London penetration depth:  
 critical current density, 92  
 high-temperature superconductors, 281  
 microwave-superconductor interaction, 154–155  
 microwave technology materials, 153  
 passive microwave circuits, 156
- London theory:  
 high-temperature superconductors (HTSs), 260  
 Meissner effect, 2  
 superconducting electronics, 117
- London two-fluid model, superconducting cavity resonators, radio frequency (RF) superconductivity, 529–530
- Long coil properties, superconducting magnet quenching propagation, cable-in-conduit coil (CICC) systems, 80–81
- Longitudinal periodicity, superconducting magnetics, 587
- Long-wavelength staring sensor (LWSS), bolometric detection, 427–428
- Lorentz force:  
 critical current density:  
 flux line lattice (FLL) structure, 96–97  
 summation theory, 99–100  
 critical current density ( $J_c$ ), flux pinning, 97–99  
 flux jumps, cryogenic stabilization, 207–208  
 Hall effect transducers, 402–403  
 high-temperature superconductors, magnetic properties, 261–262  
 magnetic flux:  
 plasma equilibrium, 16–18  
 type II superconductivity, 21  
 particle accelerators, superconducting magnetics:  
 dipole magnets, 577  
 quadrupole magnets, 577–578  
 superconducting magnetics:  
 components of, 587–588  
 radial support, 588–589  
 superconductivity research, 4
- Lorentz number, Peltier refrigeration, 502
- Lorentz relation, hard (permanent) magnetic materials, loudspeakers, 39–40
- Loss cone, magnetic mirroring, 384–385

- Loss curve:  
 capacitor energy storage, 335–336  
 steady state coupling loss, 145
- Loss maximum formula, hysteresis loss, 140
- Loss tangent, thin-film superconductor substrates, 153–154
- Lossy ferrites, soft magnetic shields, 39
- Loudspeakers, hard (permanent) magnetic materials, 39–40
- Lower limiting current, cable-in-conduit conductor stability, supercritical helium, 310
- Low-frequency noise, Schottky diode mixers, 367
- Low mass thin detector solenoids, operating principles, 412–413
- Low-noise amplifier (LNA), high-temperature superconductors (HTSs), passive microwave devices, 556–557
- Low-probability punch through, voltage-state logic circuits, 131
- Low-temperature superconductors (LTSs):  
 active microwave circuits, 160  
 critical current density:  
 resistive measurement, 93–94  
 scaling laws, flux pinning, 101–102  
 Josephson junctions, 125  
 vs. high-temperature superconductors (HTSs), 249  
 microwave technology materials, 153  
 recent developments in, 116  
 reciprocating magnetic systems, 458  
 solenoid electromagnetics, 455–456  
 superconducting digital circuits, 122  
 superconducting fault current limiters (SFCL), 544, 546  
 niobium-based materials, 549  
 superconductivity research, 4  
 transformer design, 596  
 cryostat requirements, 600  
 tunneling devices, Josephson junctions, theoretical background, 163  
 type II superconductors, alloys, 179–180  
 type I superconductors:  
 alloys, 179–180  
 properties of, 178
- Lung cancer, diagnostic imaging, 355
- Macroscopic response, magnetic materials, 30
- Maddock-James-Norris equal-area theorem, cryogenic stabilization, 212–213
- Maglev transport, magnetic levitation applications, 27–28
- Magnet cooling, particle accelerators:  
 cryostat structure, 591  
 superconductor critica temperature, 590  
 superfluid helium cooling, 590–591
- Magnetic anisotropy energy, ferromagnetic materials, 31
- Magnetic bearing, superconducting levitation, 566–567
- Magnetic confinement fusion, plasma physics, 381–387  
 closed magnetic field confinement devices, 385  
 cyclotron motion, 382  
 Debye length, 381  
 diamagnetism, 383  
 Larmor radius, 382–383  
 magnetic mirror confinement, 383–385  
 magnetic moments, 383  
 rotational transform, 385–387  
 tokamaks, 387
- Magnetic fields:  
 fusion reactor confinement, 380–381  
 gradient, magnetic separation and, 452  
 Josephson junction, tunneling devices, 166  
 strength, magnetic separation and, 452  
 superconducting magnetics, quality parameters, 584–585
- Magnetic flux:  
 charged-particle motion, 15–16  
 conversion, 10  
 critical current density, flux line lattice (FLL) structure, 94–97  
 defined, 5  
 demagnetization factor, 9–10  
 dynamo action and geomagnetism, 10–11  
 electrical machines, 11–15  
 ac generators, 13–15  
 dc generators, 13  
 transformers, 11–13  
 energy density, 6  
 ferromagnetic materials and shielding, 9  
 helicity, 10–11  
 inductance, 6  
 Lenz's law and flux conservation, 6–7  
 line and tube forces, 8  
 magnetomotive force and reluctance, 9  
 measurement of, 22  
 penetration and diffusion, 7–8  
 plasma equilibrium and flux surface, 16–18  
 Poynting flux, 7  
 principles of, 5–6  
 superconducting cyclotrons, 541–542  
 superconductors and, 18–22  
 ac performance, 21–22  
 intermediate state, 19  
 properties of, 18–19  
 type II superconductors, 19–21  
 thin detector solenoid design, 414–416  
 toroidal geometry, 8–9  
 vector potential, 6
- Magnetic forces, magnetic separation, 449–450
- Magnetic helicity, defined, 10–11
- Magnetic humps, separation properties, 457
- Magnetic induction, magnetic materials, 30
- Magnetic levitation:  
 ac levitation, 24  
 defined, 22  
 Earnshaw's theorem, 23  
 electrodynamic levitation, 24–26  
 electromagnetic casting, 27  
 electromagnetic levitation, 23–24  
 Maglev transport, 27–28  
 magnetic bearings, 28–29  
 magnetic forces in, 22–23  
 melting applications, 27  
 miscellaneous techniques, 26  
 permanent magnets, 23  
 pressure magnitude, 26–27  
 superconducting levitation:  
 applications, 567–568  
 bearing properties, 566–567  
 levitational phenomenology, 565–566  
 operating principles, 564–565  
 stable levitation, 563–564  
 theoretical models, 566  
 tuned resonators, 26
- Magnetic materials:  
 applications, 36–41  
 hard (permanent) magnets, 39–40  
 recording media, 40  
 soft materials, 37–38  
 ferrofluids, 41  
 historical background, 30  
 magnetic fields and response, 30  
 nonreciprocal materials, 40–41  
 taxonomy of, 30–36  
 antiferromagnetism, 34  
 diamagnetism, 30–31  
 double-exchange materials, 36  
 ferrimagnetism, 34–35  
 ferromagnetism, 31–32  
 metamagnetism, 35  
*M-H* loops, 32–33  
 negative exchange interaction, 33–34  
 paramagnetism, 31  
 spin glass state, 35–36  
 superparamagnetism, 36  
 thin films, 36
- Magnetic mirroring, plasma physics, 383–384
- Magnetic moment:  
 ferromagnetic resonance (FMR), 63–64  
 magnetic resonance and, 59–61  
 plasma physics, 383
- Magnetic noise, Barkhausen effect:  
 Barkhausen jump statistics, 49–50  
 domain wall-defect interaction, 44–47  
 equilibrium magnetization distribution, 44  
 experimental techniques, 50  
 flux pinning and losses in superconductors, 51–52  
 irreversible magnetization, 47–48  
 nondestructive testing, 50–51  
 principles of, 41–43  
 stress effects, 48  
 temperature dependence, 48–49
- Magnetic refrigeration:  
 components, 54–55  
 defined, 53–54  
 history and current status, 55  
 operating principles, 56–57  
 paramagnetic materials, 56  
 principle of, 54  
 temperature regulation, 57–58  
 theoretical background, 55–56
- Magnetic resonance:  
 electron paramagnetic resonance, 61–62  
 observation experiments, 62–63  
 ferromagnetic resonance, 63–64  
 nuclear magnetic resonance, 64  
 principles of, 58–61
- Magnetic resonance angiography (MRA):  
 brain imaging, 354  
 cardiac imaging, 357–358  
 vascular system, 362
- Magnetic resonance imaging (MRI):  
 bone tumors, 361

- Magnetic resonance imaging (MRI)**  
*(continued)*  
 brain imaging:  
   cerebrovascular disease, 353  
   neurodegenerative disease, 354–355  
   tumors, 354  
 breast cancer, 356–357  
 cardiac imaging, 357–358  
 chest imaging, lung cancer, 355  
 contrast agents, 350  
 female reproductive organ cancer imaging, 359  
 functional magnetic resonance imaging (fMRI), 350–351  
 joint disease, 360–361  
 liver imaging, 358  
 magnets for, 479–480  
   cryogenics, 490  
   design and construction techniques, 488–490  
   design principles, 480  
   field measurement, 486–487  
   pulsed gradient coils, 490  
   shielding parameters, 484  
   theoretical design, 480–481  
 nuclear magnetic resonance (NMR) and, 64  
 osteoporosis, 362  
 principles and applications, 348–351  
 superconducting magnets, 113–115  
 vascular system, 362  
 vertebral spine disease, 361  
*vs.* magnetic source imaging (MSI), 471
- Magnetic resonance spectroscopy (MRS):**  
 principles and applications, 349, 350  
 prostate imaging, 360
- Magnetic sensors:**  
 amorphous magnetic materials, 447–448  
 classification, 434–435  
 eddy current sensors, 440  
 Hall-effect sensors, 447  
 inductive sensors, 435–440  
   induction potentiometer, 439–440  
   linear and rotary variable-reluctance sensors, 436  
   linear variable inductor, 439  
   microsyn, 438  
   single-coil variable-reluctance sensor, 436–437  
   synchros, 438–439  
   variable-differential reluctance sensor, 437  
   variable-reluctance tachogenerators, 437–438  
 interference shielding and sensitivity, 448  
 magnetometers, 445–446  
   flux-gate magnetometers, 445  
   search coil magnetometers, 445–446  
   SQUID magnetometers, 446  
 magneto-optical sensors, 447  
 magnetoresistive sensors, 446–447  
 MOS magnetic field sensors, 445  
 thin films, 447  
 transformative sensors, 440–445  
   linear variable-differential transformer (LVDT), 441–444  
   rotary variable-differential transformer, 444–445
- Magnetic separation:**  
 applications, 458–459  
   chemical processing, 461  
   coal purification, 463  
   kaolin processing, 459–461  
   solid waste remediation, 462–463  
   titanium dioxide, 461  
   water treatment, 461–462  
 competing forces minimization, 453  
 electrodynamics, 449  
 equipment, 453–454  
   batch type, 453–454  
   carousel electromagnets, 458  
   continuous separators, 456–457  
   drum and pulley magnets, 457  
   grate-type magnets, 457  
   high-temperature superconducting reciprocating magnet systems, 458  
   high-temperature superconducting solenoid electromagnets, 456  
   induced-pole separators, 455  
   induced roll separators, 457  
   lifting magnets, 457  
   low-temperature superconducting reciprocating magnet systems, 458  
   low-temperature superconducting solenoid electromagnets, 455–456  
   plate magnets and magnetic humps, 457  
   resistive solenoid electromagnets, 455  
   wet drum magnetic separators, 457–458  
 field gradient, 452  
 field strength, 452  
 fluid dynamic effects, 450–451  
 high gradient separation characteristics, 452  
 magnetic forces, dynamics of, 449–450  
 magnetic phenomena, 449  
 matrix/filter systems, 451–452  
 minimizing competing forces, 451  
 open gradient magnet systems, 451  
 research background, 448–449  
 retention time, 452–453
- Magnetic solenoid, gyrotron systems, 397**
- Magnetic source imaging (MSI):**  
 bioelectromagnetic studies:  
   cellular sources, 467  
   electromagnetic calculations, 468  
   myocardium, 467–468  
   neurons, 467  
   numerical approaches, 468–469  
   primary current, 467  
   quasi-static approximation, 467  
 brain imaging:  
   auditory evoked fields, 471–472  
   cortical rhythms and reactivity, 472–473  
   elliptic foci, 473  
 cardiac studies:  
   distributed source reconstruction, 475  
   evoked fields, 475  
   ventricular preexcitation, 473–474  
   ventricular tachycardia, 474–475  
 future trends, 476  
 historical background, 464  
 instrumentation, 465–467  
 multichannel magnetometers, 466–467  
 neuromagnetic field detection, 465–466  
 MEG and MCG studies, 464–465  
 modeling issues, 475–476
- non-magnetic imaging modalities, 471  
 source modeling:  
   bioelectric and biomagnetic signals, 470–471  
   inverse problem, 469–470  
   regularization and constraints, 470
- Magnetocardiography (MCG):**  
 analytic solutions, 468  
 historical background, 464  
 multichannel magnetometers, 466–467  
 principles of, 464–465  
*vs.* magnetic resonance imaging (MRI), 471
- Magnetocrystalline anisotropy energy,**  
 Barkhausen effect, 42  
 equilibrium magnetization distribution, 44
- Magnetodiode sensors, properties of, 446**
- Magnetoelasticity, Barkhausen effect, stress field inhomogeneity, 45–46**
- Magnetoencephalography (MEG):**  
 analytic solutions, 468  
 epileptic foci, 483  
 historical background, 464  
 multichannel magnetometers, 466–467  
 principles of, 464–465  
*vs.* magnetic resonance imaging (MRI), 471
- Magneto-hydrodynamic (MHD) power generation:**  
 rotational transform, 386  
 superconducting magnets, 113–115
- Magnetometers:**  
 sensing properties, 445–446  
   flux-gate magnetometers, 445  
   search coil magnetometers, 445–446  
   SQUID magnetometers, 446  
   superconducting quantum interference device (SQUID) as, 525–526
- Magnetomotive force, principles of, 9**
- Magneto-optic sensors:**  
 intermediate (soft/hard) magnetic materials for, 40  
 properties of, 447
- Magnetoresistive sensors, properties of, 446**
- Magnetostriction:**  
 Barkhausen effect:  
   stress and noise statistics, 48–49  
   stress field inhomogeneity, 45–46  
 ferromagnetic materials, 32
- Magnetotransistor sensors, properties of, 446**
- Magnet stored energy, thin detector solenoid design, 415–416**
- Magnet technology:**  
 adiabatic demagnetization refrigerators (ADR), 54–55  
 superconductivity research, 4
- Main ring, particle accelerators, superconducting magnetics, 577**
- Mammography, breast cancer diagnostic imaging, 356–357**
- Manganese, magnetic resonance imaging (MRI), 350**
- Mass flow dependence, cable-in-conduit conductor stability, supercritical helium, 308–309**
- Matrix/filter systems, magnetic separation, 451–452**

- Maximum size of recoverable transition (MSRT) line, cryogenic stabilization, Wipf minimum propagation zone (MPZ), 214–215**
- Maximum-temperature requirement, quench protection, superconducting particle accelerators, 591**
- Maximum voltage to ground, superconducting particle accelerators, quench protection, 592**
- Maxwellian velocity:**  
 cyclotron motion, magnetized plasmas, 382  
 Larmor radius, magnetized plasmas, 382–383  
 nuclear fusion fuels, 376–377
- Maxwell's equations:**  
 magnetic flux:  
 Meissner effect, 2  
 penetration and diffusion, 7–8  
 plasma equilibrium, 17–18  
 Poynting flux principles, 7  
 magnetic levitation:  
 ac levitation, 24  
 electrodynamic levitation, 24–25  
 quasi-static approximation, 467
- Mean-time-to-failure (MTTF), high-temperature superconductors (HTSs), cryogenic packaging design, 561–562**
- Mechanical interrupters:**  
 cryogenic stabilization, 208  
 superconducting magnet quenching, coil protection circuits, 67
- Medical research, superconducting quantum interference device (SQUID) for, 526–527**
- Meissner effect (Meissner-Ochsenfeld effect):**  
 Barkhausen effect, flux pinning and losses in superconductors, 51–52  
 critical current density, 93–94  
 critical state magnetization model, 102–105  
 magnetic flux line lattice, 95–97  
 diamagnetic materials, 31  
 flux creep and FLL melting, 106  
 high-temperature superconductors (HTSs):  
 microstructures, 285–287  
 passive microwave devices, 554–555  
 theoretical background, 260  
 hysteresis loss, direct-current (dc) magnetization curve, 138–139  
 magnetic flux and:  
 intermediate state, 19  
 superconducting properties, 18–19  
 microwave technology, 152–153  
 superconducting electronics theory, 116  
 type I superconductors, 177, 182  
 type II superconductors, 2–3, 177
- Melt-powder-melt growth (MPMG) processing, bulk superconductors, 264**  
 texturing process, 265
- Melt-texture growth (MTG) process, YBaCuO superconductors, 265**
- Mercury superconductors:**  
 constituents in, 280–281  
 research background, 1  
 structural properties, 284–285  
 substituents in, 293–294
- Metal alloys:**  
 ductile superconductor metallurgy:  
 APC processing, 304–305  
 cabling procedures, 304  
 cold work processing, 300–302  
 composite assembly, 298–300  
 conventional processing, 297–298  
 extrusion, 300  
 final wire drawing, 303  
 niobium-titanium alloy system, 296–297, 304  
 precipitation heat treatment, 302  
 shaping procedures, 303  
 theoretical background, 295–296  
 twisting procedures, 303  
 Peltier effect in, 500–501  
 superconductor metallurgy, ternary elements, 201–203
- Metal-doping effects, YBCO superconductors, 270–272**
- Metal-in-gap (MIG) design, write recording heads, 38**
- Metal-insulator-metal tunneling, Josephson junctions, 163–164**
- Metallurgical research:**  
 ductile superconductor alloys:  
 APC processing, 304–305  
 cabling procedures, 304  
 cold work processing, 300–302  
 composite assembly, 298–300  
 conventional processing, 297–298  
 extrusion, 300  
 final wire drawing, 303  
 niobium-titanium alloy system, 296–297, 304  
 precipitation heat treatment, 302  
 shaping procedures, 303  
 theoretical background, 295–296  
 twisting procedures, 303  
 $\beta$ -tungsten superconductors, 188–190
- Metal-organic chemical vapor deposition (MOCVD):**  
 ferroelectric materials, microwave technology, 491–492  
 HTS thin films, 273–274  
 thin-film superconductor substrates, 154
- Metal-organic deposition (MOD), HTS thin films, 273–274**
- Metal oxide semiconductors (MOS), magnetic field sensors, 445**
- Metal oxide silicon field effect transistors (MOSFETs), quantum Hall effect, 403**
- Metamagnetism, properties of, 35**
- Metastability:**  
 cable-in-conduit conductors (CICC), 308  
 niobium-titanium composite superconductor metallurgy, 296–297
- M-H* curve, ferromagnetic materials, 32–33**
- Mica capacitors, properties of, 336**
- Microbridges, Josephson junction, 255**
- MicroSMES:**  
 hybrid UPS system, 510–511  
 motor drive system, 510  
 power quality, 511  
 series-connected systems, 509–510  
 superconducting fault current limiters (SFCL), 546
- Microstrip, ferroelectric materials, 492–493**
- Microsyn systems, variable-reluctance transducer, 438**
- Microwave technology:**  
 alternating current (ac) Josephson effect, 166  
 electron paramagnetic resonance (EPR), 62–63  
 ferroelectric materials:  
 applications, 494–498  
 design models, 492  
 guiding structures, 492–494  
 high-temperature superconductors, 498  
 phased array antenna, 495–497  
 theoretical background, 491–492  
 traveling wave antenna, 497–498  
 tunable filter, 495  
 varactor device, 494–495  
 voltage-controlled oscillator, 495  
 high-temperature superconductors (HTSs):  
 active microwave circuits, 159–162  
 cryogenic packaging, 540–562  
 delay lines, 560  
 film processing technology, 551–553  
 filter design, 557–560  
 interaction parameters, 154–156  
 materials, 153–154  
 nonlinearity, 155  
 passive microwave circuits, 156–159, 553–557  
 theoretical background, 152–153  
 magnetic materials for, 40–41
- MIT integral, quench protection, superconducting particle accelerators, hot spot temperature, 592**
- Millikelvin temperature range, Peltier refrigeration, 503–506**
- Mineral insulated metal sheathed (MIMS) sensors, thermocoupling, 614–615**
- MINIMAG experiment, low mass thin detector solenoids, 412–413**
- Minimum norm estimate:**  
 bioelectromagnetic source modeling, 470  
 magnetic source imaging (MSI), distributed sources, 475
- Minimum propagation zone (MPZ), cryogenic stabilization, Wipf minimum propagation zone (MPZ), 213–215**
- Minimum quench energy (MQE), superconducting magnetics, Lorentz force stability, 587–588**
- Mirror devices, fusion reactors, superconducting magnetics, 569–570**
- Mixer port impedances, Schottky diode mixers, 369**
- Mixing devices:**  
 bolometer mixers, 372  
 conversion loss, 366  
 harmonic mixers, 371  
 high-temperature superconductors, nonlinearity, microwave interaction, 155–156  
 Josephson junctions, tunneling devices, 169  
 mixer receiver noise, 366  
 negative resistance diode mixers, 372  
 parametric frequency conversion, 371–372  
 Schottky diode mixers, 366–370  
 applications, 367–368

- Mixing devices (*continued*)  
 conversion loss, 369  
 equivalent noise temperature, 369–370  
 harmonic balance, large signal analysis, 368  
 port impedance, 369  
 small signal analyses, 368–369  
 self-oscillating mixers, 372  
 superconducting devices, 372–375  
 hot electron bolometer (HEB) mixer, 374–375  
 SIS mixer, 373–374  
 theoretical background, 363–364  
 topology and design issues:  
 double balanced mixers, 371  
 double-double balanced mixers, 371  
 image rejection/enhancement, 371  
 single balanced mixers, 370–371  
 single-ended mixer, 370  
 two-terminal nonlinear devices:  
 frequency conversion, 364–365  
 frequency multiplication, 364  
 harmonic mixing, 365  
 intermodulation, 365–366  
 linear mixing, 365
- Modeling techniques, magnetic source imaging (MSI), 475–476
- Modified variable threshold logic (MVTL):  
 digital superconducting electronic circuits, 121  
 hysteretic Josephson junction, 129–131  
 flux quantization, 126
- Molecular beam epitaxy (MBE), high-temperature superconductors, thin film development, 246
- Molten materials, magnetic levitation, 27
- Monochromatic radiation thermometers, operating principles, 620
- Monotonic calibration curve, temperature measurement, 604–605
- Mott formula, Peltier coefficient, 500–501
- Multichannel magnetometers:  
 magnetic source imaging (MSI), 466–467  
 superconducting quantum interference device (SQUID), gradiometers, 521–522
- Multichip module (MCM) technology:  
 microwave applications, 161–162  
 superconducting electronics applications, 134–136
- Multilayer structures, Josephson junctions, circuit integration requirements, 255
- Multipacting, superconducting cavity resonators, radio frequency (RF) superconductivity, 530
- Multiple sclerosis, diagnostic imaging, 354–355
- Multiple stability parameters, cable-in-conduit conductor stability, supercritical helium, 310
- Multiplexer architecture, high-temperature superconductors (HTSs), filter design, 558–560
- Multipled frequency:  
 frequency converters and mixers, 363–364  
 parametric frequency conversion, 371–372
- Multiplier applications, Hall effect transducer, 407
- Multipole expansion, superconducting magnetics, 584
- “Mumetal” materials, soft magnetic shields, 39
- $\mu$ SMES (superconducting magnetic energy storage) device:  
 components, 513–514  
 configuration, 511  
 design optimization, 516–517  
 design principles, 515  
 excess storage (cost) vs. current, 515  
 excess storage (cost) vs. power (ac losses), 515–516  
 negative attributes, 513  
 positive attributes, 512–513  
 power quality, 511–517  
 vs. other storage devices, 511–512
- Mutual flux, properties of, 12
- Myocardium, magnetic source imaging (MSI), 467–468
- Néel temperature, antiferromagnetism, 34
- Negative exchange interaction, ferromagnetic materials, 33
- Negative resistance diode mixers, components of, 372
- Negative temperature coefficient (NTC) thermistors, semiconductor temperature sensors, 607–608
- Neurodegenerative disease, diagnostic imaging, 354–355
- Neuromagnetic fields, magnetic source imaging (MSI), 465–466
- Neurons, magnetic source imaging (MSI), 467
- Neutral beam heating, fusion reactors, 389–390
- Next European Torus (NET) project, sheathed flat cable manufacturing, niobium-titanium composite superconductor, 225–226
- Niobium-aluminum superconductors:  
 metallurgy, 189–190  
 production processes, 193
- Niobium compounds:  
 superconducting cavity resonators, 534  
 Nb/Cu cavities, 538  
 thermal breakdown, 537–538  
 superconducting fault current limiters (SFCL), 549
- Niobium nitride:  
 Josephson circuit fabrication, 122–123  
 superconductors made from, 187
- Niobium technology, Josephson junction, tunneling devices, 170
- Niobium-tin electromagnetic superconductor:  
 cable-in-conduit coil (CICC) systems, strand coating, 228–229  
 coupling current loss, interstrand resistance, 148  
 hysteresis loss:  
 anisotropy and variable angle orientation, 140–141  
 filament diameter, 141  
 zero-field crossing, 142  
 metallurgy, 190
- NMR magnets, 488  
 properties of, 109  
 sheathed flat cable manufacturing, 225–226  
 strand soldering techniques:  
 flat cables to copper pipe soldering, 224  
 hollow conductors, 221–222
- Niobium-titanium composite superconductor:  
 APC process, 304–305  
 cable-in-conduit coil (CICC) systems, strand coating, 228–229  
 coupling current loss, interstrand resistance, 148  
 hysteresis loss:  
 anisotropy and variable angle orientation, 140–141  
 filament diameter, 141  
 zero-field crossing, 141–142
- metallurgical research:  
 alloy fabrication, 297–298  
 APC processing, 304–305  
 cabling procedures, 304  
 cold work microstructures, 297, 300–301  
 cold work processing, 300–302  
 composite assembly, 298–300  
 conventional processing, 297–298  
 extrusion, 300  
 final wire drawing, 303  
 metastable phases, 296–297  
 niobium-titanium alloy system, 296–297, 304  
 precipitation heat treatment, 302  
 shaping procedures, 303  
 stable phases, 296  
 theoretical background, 295–296  
 twisting procedures, 303
- MRI magnets, 488–489  
 niobium-titanium-tantalum alloys, 304
- NMR magnets, 488  
 particle accelerators, superconductor magnetics, 579–580  
 properties of, 109, 187  
 sheathed flat cable manufacturing, EU-LCT conductor, 224–225  
 transformer design, 597, 598
- Noise:  
 high-temperature superconductors (HTSs), passive microwave devices, 555–556  
 receiver noise, frequency converters and mixers, 366  
 Schottky diode mixers, 367–368  
 superconducting quantum interference device (SQUID):  
 Josephson junctions, tunneling devices, 168  
 refrigeration, environmental noise, 522–523  
 sensitivity, 519  
 thermocouple temperature sensors, 615
- Noise equivalent power (NEP):  
 bolometric detection, 426  
 infrared radiation detector, 618–619
- Noise equivalent temperature difference (NETD):  
 bolometric detection, 426  
 pyroelectric IR detectors, amplifier noise, 431



- Nondestructive testing:**  
 Barkhausen effect, 50–51  
 superconducting quantum interference device (SQUID) for, 526
- Nonideal transformer:**  
 Barkhausen effect, type II superconductors, 52  
 properties of, 12
- Nonlinearity:**  
 high-temperature superconductors, microwave-superconductor interaction, 155–156  
 high-temperature superconductors (HTSs), passive microwave devices, 556–557  
 temperature sensor resistance, 610
- Nonlinear mixer devices, two-terminal nonlinear mixer devices:**  
 frequency conversion, 364–365  
 frequency multiplication, 364  
 harmonic mixing, 365  
 intermodulation, 365–366  
 linear mixing, 365
- Nonreciprocal materials, properties of, 40–41**
- Normal-insulator-normal (NIN) tunneling, Josephson junctions, 163–164**
- Normal-metal-insulator-superconductor (NIS) tunnel junction:**  
 Peltier refrigeration, 503–506  
 thermoelectric refrigeration, 630–631
- Normal resistance ( $R_n$ ), Josephson junctions, 249–250**  
 trilayer-geometry weak links, 254–255
- $n$ -type semiconductor, Peltier refrigeration, 501–506**
- Nuclear demagnetization refrigerators, properties of, 53–54**
- Nuclear fusion:**  
 fuel sources, 376–380  
 deuterium fuel, 377  
 fossil fuel energy density comparisons, 377–378  
 lunar helium, 379–380  
 reduced activation fuels, 379  
 tritium fuel, 378  
 tritium production, 378–379  
 fusion reactors, 375–376
- Nuclear magnetic resonance (NMR):**  
 high-temperature superconductor materials, applications, 278–279  
 magnets for, 479–480  
 construction techniques, 487–488  
 design principles, 480  
 field measurement applications, 486–487  
 Legendre functions, 482–484  
 manufacturing errors, 481  
 optimization methods, 483–484  
 shielding, 484  
 shimming, 484–486  
 theoretical design, 480–481  
 principles of, 64  
 superconducting magnets, 113–115  
 750 MHz proton frequency, 114  
 superconducting quantum interference device (SQUID) and, 526
- Nuclear power ratio, magnetic mirroring, 385**
- Null-flux geometry, electrodynamic levitation, 26**
- Numerical techniques, bioelectromagnetic field calculations, 468–469**
- Nyquist noise, superconducting quantum interference device (SQUID), detection coils, 520–521**
- Ohmic heating:**  
 fusion reactors, 389  
 thermoelectric refrigeration, 625–627
- Ohmic loss, supercapacitor technology, 341**
- Ohm's law:**  
 dynamo action and geomagnetism, 11  
 magnetic flux penetration and diffusion, 7–8  
 microwave-superconductor interaction, 154–155
- OMEGA conductor, strand soldering techniques, 220–221**
- $1/f$  noise, superconducting quantum interference device (SQUID), 519**
- One-dimensional model, stability margin calculations in He-I, CICC conductors, 313–314**
- Onsager relation, thermoelectric refrigeration, 626–627**
- Open gradient magnet systems (OGMS), magnetic separation, 451**
- Operating current, cable-in-conduit conductor stability, supercritical helium, 309–310**
- Optimization techniques, magnet design, 483–484**
- OR gates:**  
 complementary output switching logic (COSL) gate, voltage-state logic circuits, 130–131  
 digital superconducting electronic circuits, hysteretic Josephson junction, 128–131  
 rapid single-flux quantum (RFSQ) logic, 133–134
- Osteoarthritis, diagnostic imaging, 360–361**
- Osteoporosis, diagnostic imaging, 361–362**
- Output taper, gyrotron technology, 398**
- Output window, gyrotron technology, 399**
- Overhead lines, energy storage capacitors, 338**
- Oxidation pocket, high-temperature superconductors, thin film development, flash evaporation, 245–246**
- Oxide layer, electrolytic capacitors, 337**
- Oxide-powder-in-tube (OPIT) process:**  
 cryogenic stabilization, 206  
 high-temperature superconductors, wire and tape applications, 277
- Oxygen stability, high-temperature superconductors (HTSs):**  
 thin film development, 241–242  
 vacancies in, 288
- Pair-pair correlations, type I and II superconductors, Bardeen-Cooper-Schreiffer (BCS) theory, 181–182**
- Paper capacitors, properties of, 336–337**
- Parallel connections, capacitor energy storage, 334–335**
- Parallel plate, ferroelectric materials, 493–494**
- Parallel resonance frequency, quartz temperature sensors, 616**
- Paramagnetic materials:**  
 defined, 30  
 magnetic refrigeration, 54–55  
 properties of, 56  
 properties of, 31  
 YBCO superconductor substituents, 291
- Parameter regime, inertial confinement, fusion reactors, 392**
- Parametric frequency conversion, topology and components, 371–372**
- Parkinson's disease, diagnostic imaging, 355**
- Partial synchronization model, critical current density, scaling laws, flux pinning, 101–102**
- Particle accelerators:**  
 superconducting cavity resonators, 528–529  
 superconducting magnets, 114–115  
 bending and focusing magnets, 577–578  
 coil end design, 584  
 conductor and conductor insulation, 579–581  
 cooling mechanisms, 590–591  
 current-line distribution design, 581–582  
 field quality issues, 584–587  
 HERA, 578  
 LHC, 578  
 mechanical design issues, 587–590  
 properties of, 578–579  
 quench performance, 591  
 quench protection, 591–594  
 research and development issues, 579  
 RHIC, 578  
 sagitta, 584  
 SSC, 578  
 superconductivity principles, 578  
 system design, 577  
 Tevatron, 578  
 two-dimensional geometry, 582–584  
 UNK, 578
- Particle collisions, Tokamak fusion reactor, confinement design, 388**
- Paschen curve, superconducting magnet quenching, helium breakdown, 82–83, 88–89**
- Passing particles, Tokamak fusion reactor, bootstrap current, 391**
- Passive microwave circuits:**  
 antenna, 158–159  
 delay lines, 159  
 filters, 157–158  
 frequency agile devices, 158  
 high-temperature superconductors (HTSs), device fundamentals, 553–557  
 resonators, 156–157  
 superconductor materials, surface resistance, 155  
 theoretical background, 152–153  
 transmission line, 156
- Pauli principle, magnetic materials, 30**

- Peak surface fields, superconducting cavity resonators, 530–531
- Peltier effect:  
historical and theoretical background, 499–500  
metals and semiconductors, 500–501  
Millikelvin temperature range, 503–506  
refrigeration, 501–503  
thermocouple temperature sensors, 613  
thermoelectric effect and, 623–624
- Pendulum analog:  
Josephson junctions, 124–125  
super-conducting quantum interference device (SQUID), voltage-state logic circuits, 129–131
- Penetration depth:  
high-temperature superconductors (HTSs), 264  
passive microwave devices, 554–555
- Penetration field, hysteresis loss, filament diameter, 141
- Permalloy coils:  
soft magnetic inductors, 37–38  
soft magnetic shields, 39
- Permanent magnets, levitation of, 23
- Permeability, soft magnetic materials, 37
- Perovskites:  
high-temperature superconductors (HTS):  
Josephson junctions, 249  
structural properties, 281–282  
YBaCuO superconductors, 283–284  
rapid single-flux quantum (RF-SQ) logic circuits, 122–123
- Perturbation spectrum:  
flux jump stabilization, 324–326  
forced flow superconductor stability *vs.*, 307–308
- Phased array antenna, microwave ferroelectric materials, 495–496
- Phase locking, type I and II superconductors, 177  
Bardeen-Cooper-Schreiffer (BCS) theory, 182
- Phase-sensitive demodulation, linear variable-differential transformer (LVDT), 442–444
- Phase shift techniques:  
ferroelectric lens antenna, 496  
resonant-frequency (RF) super-conducting quantum interference device (RF SQUID), 118–119
- Phase stability, compact synchrotrons, 543–544
- Photon detectors, infrared radiation detector, 619
- Physical vapor deposition (PVD):  
thin film development, high-temperature superconductors (HTSs), 241  
thin-film superconductor substrates, 154
- Pick-up coils, superconducting magnet quenching detection, voltage sensors, 71
- Picture archiving and communication system (PACS), diagnostic imaging storage, 362
- Pinning force density ( $F_p$ ), superconductor metallurgy, 201–203
- Planar junctions:  
high-temperature superconductors (HTSs), 255  
filter design, 558
- Planar voids, superconducting magnet quenching, electrical field integrity, 85
- Planck's constant:  
magnetic resonance, 60–61  
radiation thermometry, 617
- Plant systems, inertial confinement, fusion reactors, 393
- Plasma confinement system, Tokamak fusion reactor, 387–388
- Plasma equilibrium, magnetic flux surfaces, 16–18
- Plasma exhaust system, Tokamak fusion reactor, 391
- Plasma frequency, Josephson junctions, 125  
ion-damaged weak links, 253
- Plasma heating, tokamak fusion reactor, 389–392  
bootstrap current, 390–391  
current drive, 390  
exhaust system, 391  
maintenance and materials, 392  
neutral beam injection, 390  
tritium breeding and thermal conversion blanket, 392
- Plasma oscillation, voltage-state logic punch through, 131
- Plasma physics, magnetic confinement fusion, 381–387  
closed magnetic field confinement devices, 385  
cyclotron motion, 382  
Debye length, 381  
diamagnetism, 383  
Larmor radius, 382–383  
magnetic mirror confinement, 383–385  
magnetic moments, 383  
rotational transform, 385–387  
tokamaks, 387
- Plastic capacitors, properties of, 337
- Plate magnets, separation properties, 457
- Platinum compounds, resistance temperature sensors, 605  
industrial metal resistance thermometers, 606–607
- Pneumonia, chest imaging, 355
- “Poisoned” weak links, Josephson junctions, 253
- Poisson's equation, bioelectromagnetic field calculations, 468
- Polar force, magnetic levitation, 22–23
- Polarized capacitors, properties of, 337
- Polo coil:  
cable-in-conduit coil (CICC) systems, strand bundling, 229–231  
fusion reactors, superconducting magnetism, 573–574
- Poloidal magnetic flux:  
demonstration poloidal coil (DPC-EX) coil, sheathed flat cable manufacturing, 227–228  
plasma equilibrium, 16–18
- Pool-boiling magnets, superconducting magnet quenching propagation, 78  
helium breakdown, 82–83
- Position sensors, Hall effect transducers, 407–408
- Positive temperature coefficient (PTC) thermistor, semiconductor temperature sensors, 608
- Positron emission tomography (PET). *See also* FDG positron emission tomography (FDG-PET)  
brain imaging:  
cerebrovascular disease, 353  
neurodegenerative disease, 354–355  
tumors, 354  
breast cancer, 357  
cardiac imaging, 357–358  
chest imaging, lung cancer, 355  
colorectal tumor imaging, 358–359  
diagnostic applications, 352  
*vs.* magnetic source imaging (MSI), 471
- Postanneal growth techniques, high-temperature superconductors, thin film development, 243–244
- Post-purification process, superconducting cavity resonators, thermal breakdown, 538
- Powder metallurgical systems,  $\beta$ -tungsten superconductors, 189–191
- Power capacitors, energy storage properties, 338–339
- Power devices, soft magnetic materials, 37
- Power dissipation, superconducting cavity resonators, 531
- Power Doppler ultrasound, principles and applications, 348
- Power handling:  
high-temperature superconductors (HTSs), passive microwave devices, 556–557  
thermoelectric conversion, 627
- Power quality:  
data, 507–509  
defined, 506  
standards, 507  
superconducting magnetic energy storage (SMES) configurations, 509–517  
hybrid UPS system, 510–511  
limits of, 513  
microSMES, 511  
motor drive system, 510  
 $\mu$ SMES, 511–517  
series-connected system, 509–510  
shunt-connected system, 509  
strengths of, 512–513  
theoretical background, 506–507
- Power supply systems:  
capacitor energy storage, theoretical background, 330–331  
detector magnetism, 416–418  
quench back, 417–418  
quench protection dump resistor, 416–417
- Poynting flux, principles of, 7
- Precession, magnetic resonance and, 58–61
- Precipitation heat treatment, ductile superconductor metallurgy, 302
- Precise resistance sensors, temperature measurement, 605
- Precision voltage, active microwave circuits, 161

- Precompression phenomenon,  
superconducting magnetics, azimuthal  
precompression, 588
- Pregnancy, diagnostic imaging, 359
- Pressure dependence:  
cable-in-conduit conductor stability,  
supercritical helium, 311  
superconducting magnet quenching  
propagation:  
cable-in-conduit coil (CICC) systems,  
80–81  
pool-boiling magnets, 78
- Primary current, magnetic source imaging  
(MSI), 467
- Propagation velocity, cryogenic stabilization,  
215
- Proportional-integral-differential (PID)  
control algorithms, magnetic  
levitation bearings, 29
- Proportionality coefficient, Peltier effect,  
499–500
- Prostate, diagnostic imaging, 359–360
- Proton, nuclear magnetic resonance (NMR)  
and, 64
- Proximity effect, Josephson junctions,  
257–258
- Pseudocapacitance, defined, 340
- p*-type semiconductor, Peltier refrigeration,  
501–506
- Pullthrough jacketing method, cable-in-  
conduit coil (CICC) systems, 235–  
237
- Pulmonary embolism, diagnostic imaging,  
355–356
- Pulsed gradient coils, MRI magnets, 490
- Pulsed laser deposition (PLD):  
high-temperature superconductors:  
thin film development, 244–245  
thin film processing, 273  
Josephson junctions, resistance control,  
258–259  
thin-film superconductor substrates, 154
- Pulse power units, capacitor energy storage,  
343–344
- Pulse-tube cryocoolers, superconducting  
electronic components, 123
- “Punch-through” problem, voltage-state logic  
circuits, 131
- Pyroelectric IR detectors, 430–433  
barium strontium titanate pyroelectric  
detectors, 432  
Johnson noise, 431  
lead titanate-based detectors, 431–432  
operating principles, 430  
temperature or radiation noise, 430–431  
YBaCuO detectors, 432–433
- Quadrupole field gradients, particle  
accelerators, superconducting  
magnetics, 577–578
- Quadrupole magnets:  
particle accelerators, superconducting  
magnetics, 577–578  
superconducting magnetics, current-line  
octuplet, 581–582
- Quality assurance, fusion reactors,  
superconducting magnetics, 576
- Quality factor (*Q*):  
high-temperature superconductors (HTSs):  
filter design, 557–560  
passive microwave devices, 553–557  
passive microwave circuit resonators,  
156–157
- Quantum dots, thermoelectric conversion,  
629–630
- Quantum flux parametron (QFP),  
superconducting electronic theory,  
122
- Quantum Hall effect, 403–404  
contact resistance, 404  
electrical metrology applications, 408–  
411  
sample dependence, 403  
semiconductor material dependence, 403
- Quantum mechanical exchange energy,  
Barkhausen effect, 42
- Quantum-scale thermoelectric refrigeration,  
principles of, 629–630
- Quantum wiring, Peltier refrigeration, 503
- Quartz temperature sensors, properties of,  
615–616
- Quasi-static approximation, magnetic source  
imaging (MSI), 467
- QUELL system, superconducting magnet  
quenching detection:  
advanced voltage sensors, 72–73  
fiberoptic temperature sensors, 75  
integrated volt-seconds, 74
- Quench and melt growth (QMG), YBaCuO  
superconductors, 265
- Quench back, detection magnetics, 417–  
418
- Quench detection, superconducting particle  
accelerators, 592
- Quenching, superconducting magnetics:  
detector magnetics, 416–418  
quench back, 417–418  
quench protection dump resistor,  
416–417  
electromagnetic superconductor  
protection, 111  
fusion reactors, superconducting magnetic  
protection, 575  
NMR magnets, 488  
particle accelerators:  
performance, 591  
protection, 591–594  
quench detection, 70–76  
advanced techniques, 72–75  
voltage sensors, 71–72  
quench propagation, 77–82  
adiabatic (potted) magnets, 77–78  
CICC magnets, 78–80  
long coil solutions, 80–82  
pool-boiling magnets, 78  
quench protection:  
adiabatic protection criterion, 76–77  
coil protection circuits, 65–70  
dump resistor, detection magnetics,  
416–417  
heaters, superconducting particle  
accelerators, 591–592  
principles of, 64–65  
superconducting fault current limiters  
(SFCL), 546–548
- Racetrack coil design, superconducting  
magnetics, 584
- Radial support, superconducting magnetics,  
588–589
- Radiation noise, pyroelectric IR detectors,  
431–432
- Radiation ratio thermometers, operating  
principles, 620–621
- Radiation thermometry:  
infrared radiation sensors, 618–619  
monochromatic radiation thermometers,  
620  
operating principles, 616–621  
radiation ratio thermometers, 620–621  
wide-band radiation thermometers,  
619–620
- Radiation thickness, thin detector solenoid  
design, 413–416
- Radio astronomy, superconducting  
electronics applications, 117
- Radio frequency (RF):  
electromagnetic fields, fusion reactor  
confinement, 380  
launcher system, gyrotron technology,  
398–399  
Schottky diode mixer applications, 367  
soft magnetic inductors, 38  
superconducting cavity resonators:  
superconductivity principles, 529–530  
surface resistance, 533
- Radio frequency superconducting quantum  
interference device (RF SQUID):  
applications, 118–119  
characteristics, 518  
historical and theoretical background,  
117–118  
input circuits, 520–522  
Josephson junctions, tunneling devices,  
168–169  
sensitivity, 519
- Radiology. *See also* Diagnostic imaging  
brain imaging, 353–355  
angiography, 354  
cerebrovascular disease, 353  
demyelinating disorders, 355  
neurodegenerative disease, 354–355  
tumors, 354  
breast cancer, 356–357  
chest imaging, 355–356  
lung cancer, 355  
pulmonary embolism, 355–356  
interventional, 353
- Radiology information systems (RIS),  
diagnostic imaging storage, 362
- Radionuclide emission imaging (nuclear  
medicine imaging):  
bone tumor imaging, 361  
diagnostic applications, 351–353  
positron emission tomography (PET), 352  
single-photon emission tomography  
(SPECT), 352–353
- Ramp junctions. *See also* Edge-geometry  
weak links  
Josephson junctions, high-temperature  
superconductor (HTS) materials,  
171
- Random access memory (RAM), intermediate  
(soft/hard) magnetic materials for, 40

- Rapid single-flux quantum (RSFQ) logic:  
 active microwave circuits, 161  
 alternating current (ac) Josephson effect, 166  
 components of, 131–134  
 digital superconducting electronic circuits, 120–122  
 high-temperature superconductors (HTSs), 122–123  
 Josephson junction, tunneling devices, 170
- Rare-earth metals, ferromagnetic materials, 32
- Rate-damp limit, cable-in-conduit conductor stability, 315
- React-and-wind technology:  
 sheathed flat cable manufacturing, ETL prototype, 226–227  
 strand soldering techniques, flat cables to copper pipe soldering, 224  
 superconductor metallurgy:  
 stabilization and barriers, 195–196  
 ternary elements, 201–203
- Reactive coevaporation, high-temperature superconductors, thin film development, 246
- Read heads, soft magnetic materials, 38–39
- Readout integrated circuits (ROICs), bolometric detection, 426–428
- Receiver noise, frequency converters and mixers, 366
- Reciprocating magnetic separators, properties of, 458
- Recording media:  
 Barkhausen noise, 51  
 intermediate (between soft and hard) materials, 40  
 soft magnetic materials:  
 read heads, 38–39  
 write heads, 38
- Rectangular waveguide, ferroelectric materials, 493–494
- Redox supercapacitors, defined, 340
- Reduced activation fuels, nuclear fusion, 379
- Reduced dimensionality systems, Peltier refrigeration, 503
- Reference junction correction, thermocouple temperature sensors, 615
- Reflection high-energy electron diffraction (RHEED), high-temperature superconductors, thin film development, in situ monitoring, 246–247
- Refrigeration:  
 adiabatic demagnetization refrigerators (ADR):  
 components of, 54–55  
 historical background, 55  
 properties of, 53–54  
 refrigerant construction (salt pill), 56  
 temperature regulation, 57–58
- Peltier effect, 501–506  
 millikelvin temperature range, 503–506  
 superconducting cavity resonators, 531–532  
 superconducting quantum interference device (SQUID), 522–523  
 closed cycle refrigeration, 522
- environmental noise, 522–523  
 thermoelectric conversion:  
 cryogenic cooling, 628–629, 631  
 principles of, 625–627  
 quantum-scale refrigeration, 629–630  
 superconducting tunnel junction, 630–631
- Relative permeability, soft magnetic materials, 37
- Relaxation phenomenon, magnetic resonance, 60–61
- Reluctance, magnetomotive force, 9
- Remanence:  
 Barkhausen effect, 42  
 ferromagnetic materials, 33
- Reproducibility, Josephson junctions, limits on, 259
- Residual magnetization, hysteresis loss, 138–139
- Residual resistance ratio (RRR):  
 coupling current loss, 148  
 low mass thin detector solenoids, 412–413  
 superconducting cavity resonators:  
 cavity fabrication and surface preparation, 534  
 radio frequency (RF) surface resistance, 533  
 thermal breakdown, 537–538
- Resistance:  
 Josephson junctions, control of, 258–259  
 temperature sensor resistance, 609–611
- Resistance temperature detector (RTD), industrial metal resistance thermometers, 606–607
- Resistance temperature sensors:  
 industrial metal resistance sensors, 605–607  
 precise resistance sensors, 605  
 properties of, 605–607
- Resistive electrical shims, magnet design, 486
- Resistively and capacitively shunted junction (RCSJ), Josephson effect, 251
- Resistively shunted junction (RSJ) model:  
 active microwave circuits, 159–160  
 Josephson junctions, 124–125  
 principles of, 249–250  
 tunneling devices, 166–169
- Resistive measurement, critical current density, 92–94
- Resistive solenoid electromagnets, 455
- Resistive temperature sensors, low-temperature ranges, 612
- Resistive transition analysis, flux flow state, 105
- Resonators, passive microwave circuits, 156–157
- Retention time, magnetic separation and, 452–453
- Reverse geometry imaging, digital x-ray systems and radiology, 347
- Reverse shear mode, Tokamak fusion reactor, 389
- Reynold's number, magnetic flux penetration and diffusion, 8
- Rheumatoid arthritis, diagnostic imaging, 360–361
- RHIC particle accelerator:  
 cable insulation, 581  
 characteristics of, 578  
 collar and yoke design, 590
- Risk assessment, fusion reactors, superconducting magnetics, 576
- RKKY interaction, ferromagnetic materials, 31
- Rolling assisted biaxially textured substrates (RABiTS), high-temperature superconductors, wire and tape applications, 277
- Root-mean-square (rms) fluctuation, Schottky diode mixers, 367
- Rotary variable-differential transformer, magnetic sensing, 444–445
- Rotary variable-reluctance sensors, properties of, 436–437
- Rotating magnetic field, ac generators, 14
- Rotational transform, nuclear reactors, 385–386
- Rutherford cables:  
 coupling loss, 147  
 particle accelerators, superconductor magnetics, 580
- Safety issues, cryogenic stabilization, 216–217
- Sagitta bending, superconducting magnetics, 584
- Salt pills, adiabatic demagnetization refrigerators (ADR), 56
- Saturation magnetization:  
 coupling current loss, 146  
 ferromagnetic materials, 31–32
- Scaling laws, critical current density, flux pinning models, 100–102
- Schmitt-trigger circuit, Hall effect position sensors, 408
- Schottky diode mixers, frequency mixing, 366–370  
 applications, 367–368  
 conversion loss, 369  
 equivalent noise temperature, 369–370  
 harmonic balance, large signal analysis, 368  
 port impedance, 369  
 small signal analyses, 368–369
- Schrödinger equation, Josephson junctions, 123–124
- Scintimammography, breast cancer imaging, 357
- Search coil magnetometers, properties of, 445–446
- Second Law of Thermodynamics,  $\mu$ SMES (superconducting magnetic energy storage) device, 513
- Seebeck effect:  
 Peltier refrigeration, 502  
 theoretical background, 499–500  
 thermocouple temperature sensors, 612–613  
 thermoelectric effect and, 623–624
- Self-field limitation:  
 depairing critical current density, 94  
 hysteresis loss, 143–144
- Self-oscillating mixers, characteristics of, 372

- Semiconductors:  
 Peltier effect in, 500–501  
 doping techniques, 502–503  
 superconducting fault current limiters (SFCL), 548–549  
 temperature sensors, 607–609  
 active sensors, 609  
 bulk silicon sensors, 608–609  
 thermistors, 607–608  
 thermoelectric refrigeration, 626–627
- Sensitivity devices:  
 electromagnetic interference, 448  
 superconducting quantum interference device (SQUID), 519
- Sensor technology, magnetic levitation bearings, 29
- Serial resonance frequency, quartz temperature sensors, 616
- Series connections:  
 capacitor energy storage, 335  
 superconducting magnetic energy storage (SMES) device, 509–510  
 superconducting particle accelerators, quench protection, magnet string, 592–593
- Series interleaf protection circuits, superconducting magnet quenching:  
 coil protection circuits, 68  
 cold dump circuits, 68–69
- Shajii quench theory, superconducting magnet quenching propagation, cable-in-conduit coil (CICC) systems, 79–81
- Shape anisotropy:  
 ductile superconductor metallurgy, 303  
 ferromagnetic materials, 32
- Shapiro steps:  
 alternating current (ac) Josephson effect, tunneling devices, 165–166  
 Josephson effect, 251  
 Josephson junctions, grain boundaries, 251
- Shear model, flux line lattice, summation theory, 100
- Sheathed flat cables, forced flow manufacturing, 224–228  
 DPC-EX JAERI conductor, 227–228  
 ETL wind-and-react flat conductor, 226–227  
 EU-LCT conductor, 224–225  
 niobium-tin sheathed flat cables, 225–226  
 TMC-FF JAERI conductor, 227
- “Shell effect,” flux jump stabilization:  
 dynamic stability criterion, 321  
 heat transfer coefficients, 322–324
- Shielding currents:  
 critical current density, 93–94  
 magnetic flux line lattice, 94–97  
 magnet design, MRI systems, 484  
 MRI magnets, 489–490  
 superconducting fault current limiters (SFCL), quenching applications, 547–548  
 superconducting quantum interference device (SQUID), refrigeration shielding, 523
- Shielding loss, electromagnetic superconductor, time-varying field operation, 112
- Shielding sensors, electromagnetic interference, 448
- Shields, soft magnetic materials, 39
- Shimming, magnet design, 484–486  
 ferromagnetic shims, 485–486  
 resistive electrical shims, 486  
 superconducting axial shims, 484–485  
 superconducting radial shims, 485
- Shockley theory, active semiconductor sensors, 609
- Short coil properties, superconducting magnet quenching propagation, cable-in-conduit coil (CICC) systems, 80
- Shot noise, Schottky diode mixers, 367
- Shunt-connected SMES, configuration, 509–510
- Shunt impedance, superconducting cavity resonators, 532–533
- Sideband noise:  
 frequency converters and mixers, mixer receiver noise, 366  
 Schottky diode mixers, 370
- Signal analysis, Schottky diode mixers, 368–369
- Signal attenuation, superconducting quantum interference device (SQUID), detection coils, 520–521
- Signal-to-noise ratio, bolometric detection, 426
- Signal transformers, soft magnetic inductors, 37–38
- Silicon-iron alloys, power transformers, 37
- Silsbee’s hypothesis:  
 cryogenic stabilization, 205  
 depairing critical current density, 94
- Single balanced mixer, topology and design, 370–371
- Single-coil linear variable-reluctance sensor, properties of, 436–437
- Single-electron logic (SEL), Josephson junction tunneling devices, 170
- Single ended mixer, topology and design, 370
- Single-photon emission tomography (SPECT):  
 brain imaging:  
 cerebrovascular disease, 353  
 tumors, 354  
 breast cancer imaging, 357  
 chest imaging, lung cancer, 355  
 diagnostic applications, 352–353
- Single sideband (SSB) noise, frequency converters and mixers, 366
- SINIS Peltier refrigerator, configuration, 505–506
- Sinusoidally wound stator, ac generators, 14–15
- Skeletal system, diagnostic imaging, 360–362
- Skin effect, microwave-superconductor interaction, 154–155
- Slab models, flux jump stabilization, 318
- Slew rate limitation, active magnetic bearings (AMBs), 29
- Slotline, ferroelectric materials, 493
- Small perturbation, flux jump stabilization, critical current density ( $J_c$ ), 324–326
- Small signal analysis, Schottky diode mixers, 368–369
- Smooth pipe conductors, forced flow manufacturing, 218–224  
 electroplating assembly, 219–220  
 strand soldering, central copper pipe, 220–224  
 superconductor pipes, 218–219
- Snoek’s law:  
 ferrimagnetism, 35  
 soft magnetic inductors, 38
- Soft magnetic materials:  
 inductors, 37–38  
 permeability, 37  
 power devices, 37  
 properties and applications, 37–39  
 read heads, 38–39  
 shields, 39  
 write heads, 38
- Solenoids:  
 detector solenoid design criteria:  
 coil electrical connections and leads, 419  
 cold mass support system, 418–419  
 cryogenic heat sink support structure, 419  
 gyrotron systems, 397  
 high-energy physics particle detector magnetics, 412–413  
 coil electrical connections and leads, 419–420  
 cold mass support system, 418–419  
 cryogenic heat sink solenoid support, 419  
 detector solenoid end design criteria, 418–420  
 historical background, 412  
 low mass thin detector magnets, 412–413  
 thin solenoid design criteria, 413–416  
 thin solenoid parameters, 413  
 high-temperature superconducting (HTS) electromagnets, 456  
 low-temperature superconducting (LTS) electromagnets, 455–456  
 magnet design theory, 480–481  
 manufacturing errors, 481  
 resistive solenoid electromagnets, 455
- Solid-state reaction processing, YBCO bulk superconductors, 265
- Solid-state refrigeration, Peltier effect, 501–506
- Solid waste remediation, magnetic separator equipment, 461–462
- Source modeling, magnetic source imaging (MSI):  
 bioelectric and biomagnetic signals, 470–471  
 inverse problem, 469–470  
 regularization and constraints, 470
- Spatially changing magnetic field, coupling current loss, 147
- Spherical inclusions, Barkhausen effect, domain wall pinning, 47
- Spherical void, superconducting magnet quenching, electrical field integrity, 85
- Spinel ferrites:  
 ferrimagnetism, 35  
 nonreciprocal materials, 41
- Spin glass state, properties of, 35–36

- Spin-lattice relaxation time, magnetic resonance, 61
- Spin-orbit coupling, ferromagnetic materials, 32
- Spin-spin relaxation, magnetic resonance, 61
- Sputtering techniques, high-temperature superconductors:  
 metal-organic chemical vapor deposition (MOCVD), 273–274  
 thin film development, 244, 273
- Stability criteria:  
 cryogenic stabilization, 209  
 Maddock-James-Norris equal-area theorem, 212–213  
 Stekly criterion, 209–212  
*vs.* flux-jump stabilization, 208  
 Wipf minimum propagation zone (MPZ), 213–215
- flux jump stabilization, 319  
 adiabatic criterion, 319–320  
 cryogenic criterion, 321  
 dynamic criterion, 320–321  
 equations and calculations, 317–319  
 general stability criterion, 319  
 homogeneous composites, 321–322  
 nomenclature, 327  
 outer shell thickness and heat transfer, 322–324  
 sample determinations, 324  
 small perturbation, critical state model, 324–326  
 theoretical background, 317
- forced flow superconductors:  
 cable-in-conduit conductors, 308–314  
 energy balance, 312–313  
 He-I stability calculations, 312–313  
 He-II stability calculations, 314–315  
 one-dimensional model, 313–314  
*vs.* perturbation spectrum, 307–308  
 principles of, 306–308  
 Stekly criterion, 306–307  
 transient heat transfer, 315–316  
 zero-dimensional model, 313
- Stability margin ( $\Delta E$ ), cable-in-conduit conductors (CICC), 308  
 He-I calculations, 312–314  
 supercritical helium, 308–311  
 superfluid helium (He-II), 311–312, 314–315
- Stabilization methods:  
 cryogenic stabilization:  
 advanced theory and thermal management problems, 215  
 early magnet applications, 206–207  
 historical background, 204–206  
 instability disturbances, 207–208  
 distributed heat source, 208  
 flux jumps, 207–208  
 mechanical disturbances, 208  
 safety issues, 216–217  
 techniques for, 208–215  
 Maddock-James-Norris equal-area theorem, 212–213  
 Stekly criterion, 209–212  
*vs.* flux-jump stabilization, 208  
 Wipf minimum propagation zone (MPZ), 213–215  
 thermophysical and heat data, 215–216
- superconductor metallurgy, 195–200  
 cabling degradation, 197  
 stress and strain, 197–200
- Stabilizer materials, ductile superconductor metallurgy, composite assembly, 298–299
- Stable phases, niobium-titanium composite superconductor metallurgy, 296
- Standard platinum resistance thermometer (SPRT):  
 circuit measurement, 610–611  
 low-temperature resistive sensors, 612  
 precise resistance sensors, 605
- Standing wave ratio (SWR), ferroelectric lens antenna, 497
- Steady state coupling loss, formulas for, 145
- Steady-state errors, temperature sensors, 621–622
- Stefan-Boltzmann constant:  
 bolometric detection, 425  
 wide-band radiation thermometers, 619–620
- Stekly criterion:  
 cryogenic stabilization, 209–212  
 forced flow superconductors, 306–307
- Stellarators:  
 fusion reactors, superconducting magnetics, 574–575  
 rotational transform production, 386–387
- STEP-1 prototype, sheathed flat cable manufacturing, TMC-FF JAERI conductor, 227
- Step-edge grain boundary junctions, high-temperature superconductors (HTSs), 252
- Step-edge SNS junctions, high-temperature superconductors (HTSs), 253–254
- Stewart-McCumber parameter:  
 active microwave circuits, 159–160  
 Josephson junctions, resistively shunted junction (RSJ) model, 167
- Stirling cryocooler, superconducting electronic components, 123
- Stokes' theorem, magnetic flux and vector potential, 6
- Strand coating, cable-in-conduit coil (CICC) systems, 228–229
- Strand geometry:  
 cable-in-conduit coil (CICC) systems, 229–230  
 ductile superconductor metallurgy, 299–300
- Strand soldering techniques, smooth pipe conductors, 220–224  
 flat cable solder to copper pipes, 223–224  
 LCT conductor, 222–223  
 niobium-tin hollow conductor, 221–222  
 OMEGA conductor, 220–221
- Stray coil capacitance, superconducting magnet quenching, 85–86
- Stress-accelerated grain boundary oxidation (SAGBO), cable-in-conduit coil (CICC) systems, incoloy 908 conduit, 233
- Stress cone/stress grading materials, superconducting magnet quenching, insulating surface tracking, 87–88
- Stress effects, Barkhausen noise, 48
- Stress fields, Barkhausen effect, inhomogeneity, 45–46
- Stress/strain characteristics, superconductor metallurgy, 197–200
- Stroke, brain imaging techniques, 353
- Strong focusing technique:  
 alternating gradient machines, compact synchrotrons, 544  
 magnetic levitation, 26
- Strontium-titanate bicrystal:  
 direct-current (dc) super-conducting quantum interference device (dc SQUID), 127–128  
 high-temperature superconductors, thin film development, 243
- Substituents, high-temperature superconductors (HTSs), 287–294  
 Bi-, Tl-, and Hg-based compounds, 293–294  
 La-(Ba,Sr,Ca)-Cu-O and Nd(Ce)-CuO compounds, 288–290  
 oxygen and oxygen vacancies, 288  
 YBCO compounds, 290–293
- Substrate composition, high-temperature superconductors, thin film development, 243
- Summation theory, critical current density, flux pinning, 99–100
- Supercapacitors:  
 applications, 342–344  
 intermediate storage, 343  
 pulse power units, 343–344  
 transport applications, 342–343  
 energy storage properties, 339–342  
 design technology, 340–341  
 equivalent circuits, 341  
 integration techniques, 341–342
- Superconducting axial shims, magnet design, 484–485
- Superconducting cavity resonators:  
 fabrication and surface preparation, 534  
 figures of merit, 530  
 future research issues, 538–539  
 niobium/copper cavities, 538  
 overlapping field emission, 534–536  
 passive microwave circuits, 156–157  
 peak surface fields, 530–531  
 power dissipation, 531  
 radio frequency (RF) superconductivity principles, 529–530  
 radio frequency (RF) surface resistance, 533  
 refrigerator requirements, 531–532  
 shunt impedance, 532–533  
 theoretical background, 528–529  
 thermal breakdown, 536–538  
 thermometry based diagnostics, 531
- Superconducting critical current. *See also* Critical current density  
 defined, 91–92
- Superconducting cyclotrons, theory and principles of, 540–542
- Superconducting electromagnets:  
 forces and stress, 112–115  
 high-temperature superconductors, 109–110  
 properties of, 107–108  
 protection issues, 111–112

- stability, 110–111
- technical superconductors, 108–109
- time-varying operation, 112
- Superconducting electronics:
  - digital electronic principles, 128–134
    - applications and future research, 134–136
    - rapid-single-flux-quantum (RSFQ) logic, 131–134
    - voltage-state logic, 128–131
  - historical and theoretical background, 116–123
  - principles, 123–128
    - dc superconducting quantum interference device (SQUID) magnetometer, 126–128
    - flux quantization, 125–126
    - Josephson junctions and Josephson effect, 123–125
- Superconducting fault current limiters (SFCL):
  - cryogenics, 549–550
  - electric utility grid application, 545–546
  - ferromagnetic types, 548
  - quenching types, 546–548
  - semiconductor types, 548–549
  - superconductor requirements, 549
  - theoretical background, 544–545
- Superconducting levitation:
  - applications, 567–568
  - bearing properties, 566–567
  - levitational phenomenology, 565–566
  - operating principles, 564–565
  - stable levitation, 563–564
  - theoretical models, 566
- Superconducting magnetic energy storage (SMES) device:
  - capacitor energy storage, 331
  - high-temperature superconductor materials, 279
  - power quality, 509–517
    - hybrid UPS system, 510–511
    - limits of, 513
    - microSMES, 511
    - motor drive system, 510
    - $\mu$ SMES, 511–517
    - series-connected system, 509–510
    - shunt-connected system, 509
    - strengths of, 512–513
- Superconducting magnetics:
  - case histories, 89
  - cyclotron and synchrotron theory, 540
  - electrical integrity, 82–89
    - design allowables, 88
    - field concentrations, 83–84
    - helium breakdown, 82–83, 88–89
    - insulating surface tracking, 86–88
    - insulation voids, 84–85
    - stray coil capacitance, 85–86
  - examples of, 113–115
  - fusion reactors:
    - coil structures, 569–575
    - cost optimization, 575–576
    - design issues, 575–576
    - electrical insulation design, 575
    - mirror devices, 569–570
    - quench protection, 575
    - risk and quality assurance, 576
  - stellarators, 574–575
  - theoretical background, 568–569
  - tokamaks, 570–574
  - particle accelerators, 114–115
    - bending and focusing magnets, 577–578
    - coil end design, 584
    - conductor and conductor insulation, 579–581
    - cooling mechanisms, 590–591
    - current-line distribution design, 581–582
    - field quality issues, 584–587
    - HERA, 578
    - LHC, 578
    - mechanical design issues, 587–590
    - properties of, 578–579
    - quench performance, 591
    - quench protection, 591–594
    - research and development issues, 579
    - RHIC, 578
    - sagitta, 584
    - SSC, 578
    - superconductivity principles, 578
    - system design, 577
    - Tevatron, 578
    - two-dimensional geometry, 582–584
    - UNK, 578
  - quench detection, 70–76
    - advanced techniques, 72–75
    - voltage sensors, 71–72
  - quench propagation, 77–82
    - adiabatic (potted) magnets, 77–78
    - CICC magnets, 78–80
    - long coil solutions, 80–82
    - pool-boiling magnets, 78
  - quench protection:
    - adiabatic protection criterion, 76–77
    - coil protection circuits, 65–70
    - principles of, 64–65
- Superconducting pipes, forced flow manufacturing, 218–219
- Superconducting quantum interference device (SQUID):
  - active microwave circuits, 159–162
  - analog-to-digital converter (ADC) applications, 135–136
  - applications, 523–527
    - geophysical applications, 526
    - laboratory applications, 523–526
    - medical applications, 526–527
    - nondestructive testing and evaluation, 526
  - bulk superconductor structural characterization, 267
  - dc SQUID magnetometer, 126–128
  - flux quantization, 126
  - high-temperature superconductor materials, applications, 278–279
  - historical and theoretical background, 117–118
  - input circuits, 520–522
    - detection coils, 520–521
    - gradiometers, 521–522
  - Josephson effect, 517–518
  - Josephson junctions, 125
    - ramp junctions, 171–172
    - step-edge grain boundary junctions, 252–253
  - tunneling devices, 167–169
  - limitations on, 519
  - magnetic source imaging (MSI), neuromagnetic fields, 465–466
  - magnetometer/susceptometer applications, 525–526
  - magmeter sensing, 446
  - operating principles:
    - dc SQUID, 518–519
    - RF SQUID, 518
    - sensitivity parameters, 519
  - principles of, 517
  - quantum Hall effect and, 411
  - refrigeration, 522–523
    - closed cycle refrigeration, 522
    - environmental noise, 522
    - noise reduction, 523
    - shielded rooms, 523
  - tunneling devices, Josephson junctions, theoretical background, 163
  - voltage-state logic circuits, two- and one-junction models, 129–131
- Superconducting radial shims, magnet design, 485
- Superconducting Supercollider (SSC):
  - cable insulation, 581
  - characteristics of, 578
- Superconducting tunnel junction, thermoelectric refrigeration, 630–631
- Superconducting upconverting parametric amplifier, development of, 119
- Superconductivity:
  - critical magnetic fields and critical temperatures, 1–2
  - engineering applications, 3–4
  - historical background, 1
  - Meissner effect, 2–3
  - theoretical development, 4
- Superconductor-insulator-normal (SIN) tunnel device, Josephson junctions, 164
- Superconductor-insulator-superconductor (SIS) tunnel device:
  - active microwave circuits, 161
  - frequency conversion and mixing, 372–374
  - Josephson junctions, 125
  - magnetic-field effects, 166
  - mixing and detection, 169
  - properties of, 164–165
  - ramp junctions, 171–172
  - resistively shunted junction (RSJ) model, 166–167
- Superconductor-normal-superconductor (SNS) tunnel device:
  - alternating current (ac) Josephson effect, 165–166
  - edge-geometry weak links, 254
  - Josephson junctions:
    - circuit integration requirements, 255–256
    - high-temperature superconductor (HTS) materials, 171
    - proximity effect, 258
    - ramp junctions, 171–172
    - resistance control, 258–259
  - step-edge SNS junctions, 253–254

- Superconductors:**  
 Barkhausen effect, flux pinning and losses in, 51–52  
 beta-tungsten, metallurgy, 188–190  
 bronze conductors, 194–195  
 heat-treatment principles and conditions, 200–201  
 magnetic flux and, 18–22  
 ac performance, 21–22  
 intermediate state, 19  
 properties of, 18–19  
 type II superconductors, 19–21  
 production methods, 190–194  
 bronze process, 191  
 composite diffusion process, 190  
 external diffusion process, 192  
 internal tin process, 191–192  
 jelly roll process, 192  
 niobium-aluminum production, 193–194  
 powder metallurgy process, 190–191  
 in situ process, 192  
 surface diffusion process, 190  
 tube process, 193  
 properties of, 187  
 stabilization and barriers, 195–200  
 cabling degradation, 197  
 stress and strain, 197–200  
 ternary elements, 201–203
- Supercritical helium:**  
 cable-in-conduit conductor stability:  
 heating time scale, 310  
 mass flow dependence, 308–309  
 multiple stability, 310  
 operating current dependence, 309–310  
 operating field dependence, 310  
 temperature and pressure parameters, 311  
 superconductor magnetic cooling, 590
- Superdirective antennas, passive microwave circuits, 158**
- Super exchange. See Indirect exchange**
- Superfluid helium (He-II):**  
 cable-in-conduit conductor (CICC)  
 stability, stability margin, 311–312, 314–315  
 superconductor magnetic cooling, 590–591
- Superparamagnetism, properties of, 36**
- Surface acoustic wave (SAW) filters, passive microwave circuits, 157–158**
- Surface diffusion process,  $\beta$ -tungsten superconductor production, 190**
- Surface impedance, high-temperature superconductors (HTSs), passive microwave devices, 554–555**
- Surface resistance/reactance:**  
 microwave-superconductor interaction, 154–155  
 passive microwave circuits, 156  
 superconducting cavity resonators, 533
- Surface superconductivity, type I and II superconductors, 184**
- Susceptometers, superconducting quantum interference device (SQUID) as, 525–526**
- Symmetry properties, diode mixers, 370**
- Synchros, magnetic sensing, 438–439**
- Synchrotron oscillations, compact synchrotrons, 543–544**
- Synchrotron radiation:**  
 applications and characteristics, 542–543  
 compact synchrotron, operating principles, 543–544
- Tachogenerators, variable-reluctance tachogenerators, 437–438**
- Tape materials:**  
 high-temperature superconductors (HTSs):  
 first generation processes, 276  
 OPIT method, 276  
 second generation, YBCO conductors, 276–277  
 $\beta$ -tungsten superconductor metallurgy, 189–190
- TBCCO films, high-temperature superconductors, processing technology, 551–553**
- Technical superconductors:**  
 composite superconductors, 108–109  
 high-current superconductors, 109  
 niobium-tin, 109  
 niobium-titanium, 109  
 properties of, 107–109
- Temperature coefficient of resistance (TCR), bolometric detection, 424–425**  
 YBaCuO semiconductors, 428–429
- Temperature dependence:**  
 Barkhausen noise, 48–49  
 cable-in-conduit conductor stability, supercritical helium, 311  
 critical current density, scaling laws, flux pinning, 102  
 flux creep and FLL melting, 105–106  
 magnetic refrigeration, 57–58  
 passive microwave circuits, transmission lines, 156  
 pyroelectric IR detector noise, 430–431
- Temperature sensors:**  
 errors and uncertainties, 621–622  
 dynamic errors, 622  
 steady-state errors, 621–622  
 international temperature scale (ITS), 603–604  
 measurement principles, 604–605  
 physical properties, 602–603  
 quartz temperature sensors, 615–616  
 radiation thermometry, 616–621  
 infrared radiation sensors, 618–619  
 monochromatic radiation thermometers, 620  
 radiation ratio thermometers, 620–621  
 wide-band radiation thermometers, 619–620  
 resistance measurement, 609–611  
 resistance sensors, 605–607  
 industrial metal resistance sensors, 605–607  
 precise resistance sensors, 605  
 resistive sensors, low temperature ranges, 612  
 semiconductor sensors, 607–609  
 active semiconductor sensors, 609  
 bulk silicon sensors, 608  
 thermistors, 607–608  
 thermocouple sensors, 612–615
- Ternary elements, superconductor metallurgy, 201–203**
- Tevatron superconducting particle accelerator:**  
 cable insulation, 581  
 characteristics of, 578
- Thallate superconductors:**  
 bulk processing, 265–266  
 constituents in, 280–281  
 structural, electrical, and magnetic properties, 268–270  
 structural properties, 284–285  
 substituents in, 293–294
- Thallination process, HTS thin film production, 274**
- Thermal breakdown, superconducting cavity resonators, 536–538**
- Thermal conductivity, flux jump stabilization, 321**
- Thermal conversion blanket, tokamak fusion reactor, 392**
- Thermal diffusion, cable-in-conduit conductor stability, supercritical helium, 309**
- Thermal-hydraulic quenchback (THQB), superconducting magnet quenching propagation, cable-in-conduit coil (CICC) systems, 81–82**
- Thermal noise:**  
 infrared radiation detector, 618–619  
 Schottky diode mixers, 367
- Thermistors:**  
 low-temperature resistive sensors, 612  
 semiconductor temperature sensors, 607–608
- Thermocouple temperature sensors:**  
 circuit measurement, 615  
 physical principles, 612–613  
 sensor design, 614–615  
 thermocouple properties, 613–614
- Thermodynamic stability, thin film development, 241–242**
- Thermoelectric conversion:**  
 efficiency technology, 627–628  
 future research issues, 631–632  
 power generation, 627  
 refrigeration:  
 cryogenic cooling, 628–629, 631  
 principles of, 625–627  
 quantum-scale refrigeration, 629–630  
 superconducting tunnel junction, 630–631  
 theoretical background, 623  
 thermoelectric effects, 623–624
- Thermoelectric effects:**  
 theoretical background, 499  
 thermoelectric conversion, 623–624
- Thermometry-based diagnostics, superconducting cavity resonators, 531**
- Thermophysical data, cryogenic stabilization, 215–216**
- Thermopiles, infrared radiation detector, 618–619**
- Thevenin equivalent resistance, temperature sensors, 610**
- Thin detector solenoids:**  
 cryogenic cooling, 420–422



- design criteria, 413–416
- low mass technology, 412–413
- parameter definition, 413
- Thin film magnetic materials:
  - ferroelectric materials, microwave technology, 491–492
  - high-temperature superconductors (HTSs):
    - applications, 278–279
    - device processing, 552
    - epitaxy, 242–243
    - flash evaporation, 245–246
    - historical background, 240–241
    - laser ablation, 244–245
    - materials, 241
    - metal-organic chemical vapor deposition, 273–274
    - microwave packaging, 552–553
    - molecular beam epitaxy, 246
    - physical/chemical deposition techniques, 271, 273
    - postanneal techniques, 243–244
    - processed film characterization, 274
    - processing technology, 551–553
    - pulsed laser deposition, 273
    - reactive coevaporation, 246
    - in situ techniques, 244, 246–247
    - in situ vs. ex-situ* processing, 271
    - sputtering, 244, 273
    - structural, electrical, and mechanical properties, 274–276
    - thermodynamics, 241–242
  - properties of, 36
  - sensing properties, 447
  - write heads, 38
- Thin-film superconductors:
  - microwave technology materials, substrates for, 153–154
  - thermoelectric conversion, 628
- Third-order intercept (TOI), high-temperature superconductors, nonlinearity, microwave interaction, 155–156
- Thomson effect:
  - theoretical background, 499–500
  - thermocouple temperature sensors, 613
  - thermoelectric effect and, 623–624
- Three-conductor belted cables, energy storage capacitors, 338
- Three-terminal tunneling devices, Josephson junctions, 172–173
- Time decay, superconducting magnetics, 586
- Time-dependent equation, superconducting magnet quenching propagation, cable-in-conduit coil (CICC) systems, 79
- Time-varying field:
  - charged-particle motion, magnetic flux, 16
  - electromagnetic superconductor operation, 112
- Titanium dioxide, magnetic separation, 460
- TMC-FF conductor, sheathed flat cable manufacturing, 227
- Tokamak fusion reactor, 387–392
  - anomalous diffusion, 388
  - bootstrap current, 390–391
  - confined particle collision, 388
  - current drive, 390
  - enhanced confinement, 388–389
  - enhanced reversed shear mode, 389
  - inductive plasma current, 387
  - maintenance and materials, 392
  - neutral beam injection, 390
  - plasma confinement system, 387–388
  - plasma exhaust system, 391–392
  - plasma heating methods, 389–390
  - plasma limiters, 387
  - rotational transform production, 387
  - superconducting magnetics:
    - coil structure, 570–574
    - quenching, multiple magnets, 68
  - tritium breeding and thermal conversion blanket, 392
- Toroidal coils:
  - fusion reactors, superconducting magnetics, 570–574
  - sheathed flat cable manufacturing, TMC-FF JAERI conductor, 227
  - superconducting magnets, 113
- Toroidal geometry:
  - magnetic flux and fields, 8–9
  - plasma equilibrium, 16–18
  - nuclear reactors, rotational transform, 385–386
- Torque, magnetic resonance and, 58–61
- Training phenomenon, cryogenic stabilization, 207
  - forced flow superconductors, 307–308
- Transformative sensors, properties of, 440–445
  - linear variable-differential transformer (LVDT), 441–444
  - rotary variable-differential transformer, 444–445
- Transformers:
  - magnetic flux, 11–13
  - soft magnetic materials, 37
  - superconductivity principles:
    - configuration, 597
    - design and analysis, 597–600
    - state-of-the-art technology, 600–601
    - theoretical background, 596–597
- Transient coupling loss, formulas for, 145–146
- Transient heat transfer, cable-in-conduit conductor stability, 315–316
- Transistor sensors, active semiconductor sensors, 609
- Transition metals, superconductivity research, 1–2
- Transit time factor, superconducting cavity resonators, 530
- Transmission line:
  - microwave technology, ferroelectric materials, 492
  - passive microwave circuits, 156
- Transport applications, capacitor storage in, 342–343
- Transport currents:
  - coupling loss in, 146–147
  - critical current density, 93–94
  - flux jump stabilization, equations and calculations, 317–318
  - hysteresis loss, 142–143
  - processed superconducting thin films, 274
- Transrectal ultrasound (TRUS), prostate imaging, 359–360
- Transverse resistivity, coupling current loss, 144–145
- Trapped particles, Tokamak fusion reactor, bootstrap current, 391
- Traveling normal zones (TNZs), cryogenic stabilization, 215
- Traveling wave antenna, ferroelectric materials, 497–498
- Trilayer-geometry weak links, Josephson junctions, 254–255
- Tritium fuel:
  - nuclear fusion:
    - decay, 378
    - production, 378–379
  - Tokamak fusion reactor, breeding and thermal conversion blanket, 392
- Tube mill jacketing process, cable-in-conduit coil (CICC) systems, 234–235
- Tube process,  $\beta$ -tungsten superconductors, 192
- Tuberculosis, chest imaging, 355
- Tunable filters, microwave ferroelectric materials, 495
- Tune resonators, magnetic levitation, 26
- $\beta$ -Tungsten superconductors, metallurgy, 188–190
- Tunneling devices, Josephson junction:
  - alternating current Josephson effect, 164–166
  - digital applications, 169–170
  - high-temperature superconductor systems, 170–172
  - magnetic-field effects, 166
  - metal-insulator-metal (MIM) tunneling, 163–164
  - mixing and detection, 169
  - related superconductive devices, 173
  - resistively shunted junction (RSJ) model, 166–167
  - superconducting quantum interference devices (SQUID), 167–169
  - theoretical background, 163
  - three-terminal devices, 172–173
- Twisting process, ductile superconductor metallurgy, 303
- Two-dimensional geometry, superconducting magnetics, 582–584
  - coil geometry, 582
  - current shell approximations, 582–583
  - iron yoke contribution, 583
  - limits of coil geometry, 584
  - operating margin, 583–584
  - symmetry parameters, 582
- Two-dimensional scanning, ferroelectric lens antenna, 496
- Two-fluid superconductor model, microwave-superconductor interaction, 154–155
- Two-terminal nonlinear mixer devices:
  - frequency conversion, 364–365
  - frequency multiplication, 364
  - harmonic mixing, 365
  - intermodulation, 365–366
  - linear mixing, 365
- Type I superconductors:
  - BCS theory and, 180–181
  - pair-pair correlations, 181–182
  - Cooper pairing, 181

- Type I superconductors (*continued*)  
 critical magnetic field parameters, 184–185  
 electromagnetic properties, 107  
 examples of, 177–180  
   high-temperature materials, 180  
   intermediate materials, 178  
   low-temperature materials, 178–180  
 Ginzburg-Landau equations, 182–183  
 hysteresis loss, 138  
 Meissner flux exclusion, 182  
 phase locking and wave function rigidity, 182  
 physical characteristics, 180–184  
 properties of, 177  
 surface superconductivity, 184  
 vortex attraction, 184  
 vortex repulsion, 183–184  
 vortex state, 183  
 zero electrical resistance, 182
- Type II superconductors:  
 Barkhausen effect, nonideal structures, 52  
 BCS theory and, 180–181  
   pair-pair correlations, 181–182  
 Cooper pairing, 181  
 critical current density ( $J_c$ ), flux pinning, 97–102  
 critical magnetic field parameters, 184–185  
 cryogenic stabilization, 205–206  
   early magnet applications, 206–207  
 depairing critical current density, 94  
 electromagnetic properties, 107–108  
 examples of, 177–180  
   high-temperature materials, 180  
   intermediate materials, 178  
   low-temperature materials, 178–180  
 Ginzburg-Landau equations, 182–183  
 hysteresis loss, 138  
   dc magnetization curve, 138–139  
 magnetic flux, 19–21  
   flux flow, 20  
   flux penetration and jump, 20–21  
 magnetic properties, 261–262  
 Meissner effect, 2–3  
 Meissner flux exclusion, 182  
 niobium-titanium composite  
   superconductor, metallurgical  
   research, 295–296  
 phase locking and wave function rigidity, 182  
 physical characteristics, 180–184  
 properties of, 177  
 superconducting electronics theory, flux  
   quantization, 125–126  
 superconducting levitation, 564  
 surface superconductivity, 184  
 vortex attraction, 184  
 vortex repulsion, 183–184  
 vortex state, 183  
 zero electrical resistance, 182
- Ultracapacitor. *See also* Supercapacitors  
 defined, 339
- Ultrasound:  
 female reproductive organ cancer imaging, 359
- fetal and pregnancy imaging, 359  
 principles and applications, 347–348
- Uncertainty measurement, temperature  
 sensors, 621–622
- Uncooled infrared detector arrays:  
 bolometric detectors, 424–430  
   operating principles, 424–426  
 vanadium oxide microbolometer arrays,  
 424–426  
 YBaCuO microbolometer arrays,  
 428–430
- pyroelectric IR detectors, 430–433  
 barium strontium titanate pyroelectric  
 detectors, 432  
 Johnson noise, 431  
 lead titanate-based detectors, 431–432  
 operating principles, 430  
 temperature or radiation noise, 430–431  
 YBaCuO detectors, 432–433  
 theoretical background, 424
- Uninterruptible Power Supply (UPS) SMES  
 system, 510–511
- Universal scaling regimes, superconducting  
 magnet quenching propagation, cable-  
 in-conduit coil (CICC) systems, 81
- UNK superconducting particle accelerator:  
 cable insulation, 581  
 characteristics of, 578
- Upconversion:  
 frequency converters and mixers, 363–364  
 parametric frequency conversion, 371–372  
 two-terminal nonlinear mixer devices, 365
- Vanadium oxide microbolometer arrays,  
 bolometric detection, 426–427
- Varactor, microwave ferroelectric materials,  
 494–495
- VAR balance equation, energy storage  
 capacity, 339
- Variable angle orientation, hysteresis loss,  
 140–141
- Variable-reactance circuits, microwave  
 ferroelectric materials, varactors,  
 494–495
- Variable-reluctance sensors:  
 linear and rotary configurations, 436–438  
 microsyn, 438  
 variable-differential reluctance sensor,  
 437  
 variable-differential tachogenerator,  
 437–438
- Vascular system, diagnostic imaging, 362
- Vector potential, magnetic flux, 6
- Ventricular preexcitation, magnetic source  
 imaging (MSI), 473–474
- Ventricular tachycardia, magnetic source  
 imaging (MSI), 474–475
- Vertebral spine disease, diagnostic imaging,  
 361
- Vibrating sample magnetometer (VSM), bulk  
 superconductor structural  
 characterization, 267
- Viscous drag force, magnetic separation and,  
 453
- Void fraction:  
 cable-in-conduit coil (CICC) systems,  
 strand bundling, 229–230
- coupling current loss, interstrand  
 resistance, 148
- Voids:  
 Barkhausen effect, domain wall pinning,  
 46–47  
 critical current density ( $J_c$ ), flux pinning,  
 98–99  
 superconducting magnet quenching, solid  
 insulation voids, 84–85
- Voltage-controlled oscillator (VCO):  
 microwave ferroelectric materials, 495  
 temperature sensor resistance, 610
- Voltage induction, ac generator, 13–14
- Voltage measurement, superconducting  
 quantum interference device (SQUID)  
 for, 525
- Voltage sensors, superconducting magnet  
 quenching detection:  
 advanced techniques, 72–73  
 conventional taps and bridges, 71  
 detection circuits, 71–72  
 fiberoptic temperature sensors, 74–75  
 integrated volt-second detection, 74  
 internal termination, 73–74  
 joint voltage extraction, 73  
 pick-up coils, 71
- Voltage-state logic:  
 4 x 4 packet switch, 135  
 digital superconducting electronic circuits:  
 components, 128–131  
 theoretical background, 120–121
- Voltage waveform smoothing, capacitor  
 energy storage, 331
- Volumetric heat capacity, stability margin  
 calculations in He-I, CICC  
 conductors, energy balance, 312–313
- Von Klitzing constant, quantum Hall effect,  
 409–411
- Vortex attraction, type I and II  
 superconductors, 184
- Vortex repulsion, type I and II  
 superconductors, 183–184
- Vortex structure:  
 Barkhausen effect:  
   flux pinning and losses in  
   superconductors, 51–52  
   type II superconductors, 52  
 critical current density, flux pinning,  
 98–99  
 high-temperature superconductors,  
 magnetic properties, 262  
 type II superconductors:  
   low-temperature superconductivity,  
   178–179  
   physical properties, 183
- Water treatment systems, magnetic  
 separator equipment, 460–461
- Wave heating, fusion reactors, 389
- Weak Interacting Massive Particles  
 (WIMPs), superconducting quantum  
 interference device (SQUID) search  
 for, 526
- Weak links:  
 grain boundary Josephson junctions, 250  
 Josephson junctions, 253  
 edge-geometry weak links, 254

- triple-geometry weak links, 254–255
- Well-cooled regimes, stability margin
  - calculations in He-I, CICC
  - conductors, energy balance, 312–313
- Wet drum magnetic separators, properties of, 457–458
- Whole coil normality, superconducting magnet quenching propagation, cable-in-conduit coil (CICC) systems, 78–79
- Wide-band radiation thermometers, operating principles, 619–620
- Wiedemann-Franz law:
  - Peltier refrigeration, 502
  - superconducting cavity resonators, cavity fabrication and surface preparation, 534
- Wien's law, radiation thermometry, 617
- Winding topology:
  - electromagnetic superconductors, 112
  - fusion reactors, superconducting magnetics, 568–569
- Window-frame coil design, superconducting magnetics, 584
- Wipf minimum propagation zone (MPZ), cryogenic stabilization, 213–215
- Wire products:
  - ductile superconductor metallurgy, 303
  - high-temperature superconductors (HTSs):
    - first generation processes, 276
    - OPIT method, 276
    - second generation, YBCO conductors, 276–277
  - NMR magnets, diametric properties, 487–488
  - resistance temperature detectors (RTDs), 606–607
  - thermocouple temperature sensors, 613–614
- Wolff-Parkinson-White (WPW) syndrome, magnetic source imaging (MSI), 473–474
- Write heads, soft magnetic materials, 38
  
- X-ray diffraction (XRD):
  - bulk superconductor structural characterization, 266–267
  - high-temperature superconductors (HTS),
    - YBaCuO structural properties, 283–284
    - processed superconducting thin films, 274
    - thallate superconductors, 268–269
    - thin film structural, electrical, and magnetic properties, 274–276
- X-ray imaging:
  - applications, 345–346
  - breast cancer, 356–357
  - cardiac imaging, 357–358
  - synchrotron radiation, 542–543
- X-ray lithography, synchrotron radiation, 543
- X-ray transmission computed tomography (X-ray CT):
  - bone tumor imaging, 361
  - brain imaging techniques:
    - cerebrovascular diseases, 353–354
    - neurodegenerative disease, 354–355
  - breast cancer, 356–357
  - chest imaging:
    - lung cancer, 355
    - pulmonary embolism, 356
  - colorectal tumor imaging, 358–359
  - female reproductive organ cancer imaging, 359
  - joint disease imaging, 360–361
  - kidney imaging, 359
  - liver imaging, 358
  - principles of, 346–347
  - prostate imaging, 359–360
- YBaCuO superconductors:
  - bolometric detection, microbolometer arrays, 428–429
  - pyroelectric IR detectors, 432–433
  - structural properties, 283–284
  - substituents in, 290–293
  - texturing processes, 265
- YBCO superconductors:
  - applications, 277–279
  - bulk superconductor processing, 264–266
  - metal-doping effects, 270–271
  - structural, electrical, and magnetic properties, 267–268
  - texturing, 265
  - YBCO 123 solid-state reaction technique, 265
  - constituents in, 280–281
  - cryogenic stabilization, 206
  - direct-current (dc) superconducting quantum interference device (dc SQUID), 127–128
  - high-temperature superconductors (HTS):
    - electronics, 116
    - thin film processing technology, 551–553
  - Josephson junctions:
    - bicrystal grain boundaries, 252
    - circuit integration requirements, 256
    - edge-geometry weak links, 254
    - groundplains, 256–257
    - planar/in-line junctions, 255
    - properties, 249
    - reproducibility limits, 259
    - step-edge grain boundary junctions, 252–253
    - step-edge SNS junctions, 253–254
  - rapid single-flux quantum (RFSQ) logic circuits, 122–123
  - substituents in, 290–293
  - superconducting levitation, 564
  - superconducting transformer design, 598–599
  - thin film development:
    - high-temperature superconductors (HTSs), 241
    - oxygen stability, 241–242
  - wire applications, 276–277
- “Ying-Yang” coils, superconducting magnets, 113
- Young's modulus, superconductor metallurgy, stress/strain characteristics, 198–200
- Yttrium high-temperature superconductors (HTSs), properties of, 110
  
- Zeeman energy, magnetic resonance, 60–61
- Zero-dimensional model, stability margin calculations in He-I, CICC conductors, 313
- Zero electrical resistance, type I and II superconductors, 182
- Zero-field crossing, hysteresis loss, 141–142