

Practical Antenna Handbook

Sixth Edition

Joseph J. Carr
George W. Hippisley
Karl F. Warnick

Mc Graw Hill

About the Authors

Karl F. Warnick is a professor of Electrical and Computer Engineering at Brigham Young University and a Fellow of the Institute of Electrical and Electronics Engineers. His research in the fields of antenna theory, wireless communications, computational modeling, and imaging sensors for radio astronomy has been published in over 300 papers, books, and presentations.

Joseph J. Carr was a military electronics technician and the author of several popular electronics books, including McGraw Hill's *Secrets of RF Design*, and *Old Time Radios! Restoration and Repair*. He wrote a monthly column for *Nuts & Volts* magazine.

George W. (Bud) Hippisley, W2RU, earned his BSEE degree from MIT and was formerly chief operating officer for a major supplier of electronics to the cable TV industry. A longtime active amateur radio operator who has won or ranked nationally in many competitive on-the-air operating events, he has given talks on the basics of antennas and ionospheric propagation to radio clubs and other groups for more than 35 years.

Front cover: Two towers support the author's monoband Yagi antennas for 40, 20, 15, 6, and 2 meters. The taller tower is shunt-fed on 160 meters and supports one end of an 80-meter dipole; the shorter tower supports a 40-meter dipole. *In the foreground*: a section of open-wire transmission line feeding the 80-meter dipole. *Out of view*: Beverage wires for low-band receiving and a triband Yagi for 10, 15, and 20 meters.

Practical Antenna Handbook

Joseph J. Carr
George W. Hippisley
Karl F. Warnick

Sixth Edition



McGraw Hill books are available at special quantity discounts to use as premiums and sales promotions, or for use in corporate training programs. To contact a representative please e-mail us at bulksales@mcgraw-hill.com.

Practical Antenna Handbook, Sixth Edition

Copyright © 2025, 2012, 2001, 1998, 1994, 1989 by McGraw Hill LLC. All rights reserved. Printed in the United States of America. Except as permitted under the United States Copyright Act of 1976, no part of this publication may be reproduced or distributed in any form or by any means, or stored in a data base or retrieval system, without the prior written permission of the publisher.

1 2 3 4 5 6 7 8 9 LCR 29 28 27 26 25 24

Library of Congress Control Number: 2024943448

ISBN 978-1-260-13225-0
MHID 1-260-13225-0

The pages within this book were printed on acid-free paper.

Sponsoring Editor
Lara Zoble

Project Manager
Tasneem Kauser, KnowledgeWorks Global Ltd.

Copy Editor
KnowledgeWorks Global Ltd.

Proofreader
KnowledgeWorks Global Ltd.

Production Supervisor
Richard Ruzycka

Composition
KnowledgeWorks Global Ltd.

Art Director, Cover
Anthony Landi

Information contained in this work has been obtained by McGraw Hill LLC. ("McGraw Hill") from sources believed to be reliable. However, neither McGraw Hill nor its authors guarantee the accuracy or completeness of any information published herein, and neither McGraw Hill nor its authors shall be responsible for any errors, omissions, or damages arising out of use of this information. This work is published with the understanding that McGraw Hill and its authors are supplying information but are not attempting to render engineering or other professional services. If such services are required, the assistance of an appropriate professional should be sought.

Contents at a Glance

Part I Fundamentals

1 Introduction to Radio Communications
2 Radio Waves
3 Radio Wave Propagation
4 Ionospheric Propagation
5 Antenna Basics
6 Transmission Lines and Impedance Matching
7 Antenna Arrays

Part II Building-Block Antennas

8 Dipoles
9 Large Wire Loop Antennas
10 Multiband and Tunable Wire Antennas
11 Vertically Polarized Antennas

Part III Directional Antenna Arrays

12 Wire Arrays
13 Vertical Arrays
14 The Yagi-Uda Beam Antenna
15 Cubical Quads and Delta Loops

Part IV Antennas for Specialized Applications

16 Receiving Antennas
17 Hidden and Limited-Space Antennas
18 Mobile and Marine Antennas
19 Emergency and Portable Antennas

20 Antennas for 160 Meters
21 VHF and UHF Antennas

Part V Microwave Antennas
22 Microwave Waveguides
23 Microwave Antennas
24 Antenna Noise
25 Radio Astronomy Antennas
26 Radio Direction-Finding (RDF) Antennas
27 Antennas for Wireless Devices

Part VI Tuning, Troubleshooting, and Design Aids
28 Antenna Tuners (ATUs)
29 Antenna Modeling Software
30 Testing and Troubleshooting

Part VII Mechanical Construction and Installation
31 Supports for Wires and Verticals
32 Towers
33 Grounding for Safety and Performance
34 Zoning, Restrictive Covenants, and Neighbors

Index

Contents

Preface to the Sixth Edition
Preface to the Fifth Edition
Acknowledgment

Part I Fundamentals

1 Introduction to Radio Communications

2 Radio Waves
- 2.1 Water-wave Analogy
- 2.2 Frequency and Wavelength
- 2.3 Properties of Electromagnetic Waves
  - 2.3.1 Radio Waves Are the Result of Accelerated Electrical Charges
  - 2.3.2 Electromagnetic Waves Have a Finite Speed
  - 2.3.3 Radio Waves Are Different from Static or Quasi-Static Charge Fields
- 2.4 Radio Wave Intensity
- 2.5 Phase Velocity
- 2.6 Isotropic Sources
- 2.7 Receiving Antennas
- 2.8 Free Space Path Loss
- 2.9 Polarization
- 2.10 Band Designations
- 2.11 Radio Spectrum Access, Regulation, and Coordination

3 Radio Wave Propagation

3.1 The Earth's Atmosphere
3.2 Reflection, Refraction, Diffraction, and Dispersion
3.2.1 Diffraction Phenomena
3.3 Propagation Paths
3.4 Ground-Wave Propagation
3.4.1 Ray Path Curvature
3.4.2 Horizon Distance
3.4.3 Surface Waves
3.4.4 Ground-Wave Propagation Frequency Dependence
3.5 Tropospheric Propagation
3.6 Fading Mechanisms

4 Ionospheric Propagation
4.1 The Sun's Effect on the Ionosphere
4.1.1 Sunspots
4.2 Ionospheric Layers
4.2.1 D Layer
4.2.2 E Layer
4.2.3 F Layer
4.3 Measures of Ionospheric Propagation
4.3.1 Critical Frequency
4.3.2 Virtual Height
4.3.3 Maximum Useable Frequency (MUF)
4.3.4 Lowest Useable Frequency (LUF)
4.4 Regular Ionospheric Variations
4.4.1 Diurnal (Daily) Variation
4.4.2 27-Day Cycle
4.4.3 Seasonal Variation
4.4.4 11-Year and 22-Year Cycles
4.5 Ionospheric Disturbances
4.5.1 Sporadic E Layer
4.5.2 Sudden Ionospheric Disturbances (SIDs)

4.5.3 Ionospheric Storms
4.6 Ionospheric Sky-Wave Propagation
4.6.1 Angle of Incidence and Radiation Angle
4.6.2 Geography and the MUF
4.6.3 Great Circle Paths
4.6.4 Ionospheric Fading and Diversity
4.7 Using the Ionosphere
4.7.1 Skew Paths
4.7.2 Gray Line Propagation
4.7.3 Long Path
4.7.4 Ionospheric Scatter Propagation
4.7.5 Auroral Propagation
4.7.6 Meteor Scatter Propagation
4.8 Propagation Predictions
4.9 The Ionosphere and the Magic of Radio

5 Antenna Basics
5.1 Circuit Fundamentals
5.1.1 Superposition Principle
5.1.2 Electric Charge
5.1.3 Current
5.1.4 Resistors
5.1.5 Capacitors
5.1.6 Inductors
5.1.7 Phasors
5.1.8 Complex Impedance
5.1.9 RLC Circuit
5.1.10 Resonance
5.1.11 Quality Factor
5.1.12 Sources
5.1.13 Units
5.2 Fields
5.2.1 Voltage and the Electric Field

5.2.2 Current and the Magnetic Field
5.2.3 Displacement Current and Maxwell's Equations
5.3 The Hertzian Dipole
5.4 The Half-Wave Dipole
5.4.1 Standing Waves
5.4.2 Velocity of Propagation and Antenna Length
5.4.3 Half-Wave Dipole Feedpoint Impedance
5.5 Antenna Losses and Radiation Efficiency
5.6 Near Field and Far Field
5.7 Antenna Pattern
5.8 Gain and Directivity
5.9 Effective Isotropic Radiated Power
5.10 Reciprocity
5.11 Receiving Antennas and Effective Area
5.12 Friis Transmission Equation and Link Budgets
5.13 Signal to Noise Ratio
5.13.1 SNR and Radiation Efficiency
5.14 Electrically Small Antennas and Bandwidth Limitations

6 Transmission Lines and Impedance Matching
6.1 Types of Transmission Lines
6.1.1 Parallel Conductor Lines
6.1.2 Coaxial Lines
6.1.3 Microstrip, Stripline, and Coplanar Waveguide
6.2 Lumped Element Model
6.3 Characteristic Impedance
6.4 Phase Velocity
6.5 Loss in Transmission Lines
6.6 Transmission Line Responses

6.6.1 Step-Function Response of a Transmission Line
6.7 Reflection Coefficient
6.8 Sinusoidal Steady State
6.8.1 Standing Wave Ratio
6.8.2 Mismatch Losses
6.9 Impedance Matching in Antenna Systems
6.9.1 Distributed Matching Networks
6.9.2 Coaxial Cable Baluns
6.9.3 Stub Matching
6.9.4 Quarter-Wave Matching Section
6.9.5 Series Matching Section
6.10 The Smith Chart
6.10.1 Smith Chart Components
6.10.2 Normalized Impedance
6.10.3 Pure Resistance Line
6.10.4 Constant Resistance Circles
6.10.5 Constant Reactance Circles
6.10.6 Plotting Impedance and Admittance
6.10.7 Transmission Line Distance Scale
6.10.8 Radially Scaled Parameters
6.10.9 Transmission Line Calculations
6.10.10 Stub Matching with the Smith Chart
6.10.11 Incorporating Loss
6.10.12 Frequency Response Plots
6.11 S-parameters and Microwave Networks

7 Antenna Arrays
7.1 Array Basics
7.2 Array Factor
7.3 Feeding Arrays
7.4 Parasitic Arrays
7.5 Ground Effects

7.6 Lossy Ground

Part II Building-Block Antennas

8 Dipoles
8.1 The Half-Wave Dipole
8.1.1 Dipole Construction
8.1.2 The Dipole Feedpoint
8.1.3 The Dipole Radiation Pattern
8.1.4 Tuning the Dipole Antenna
8.2 Modified Dipoles
8.2.1 Inverted-V Dipole
8.2.2 Sloping Dipole ("Sloper")
8.2.3 Broadband Dipoles
8.2.4 Shortened Coil-Loaded Dipoles
8.2.5 Bent Dipole
8.2.6 Extended Double Zepp Antenna
8.2.7 Multiband Fan Dipole
8.3 Summary

9 Large Wire Loop Antennas
9.1 Half-Wavelength Large Loops
9.2 1-Wavelength Large Loops
9.3 Delta Loop
9.4 Half-Delta Sloper (HDS)
9.5 2λ Bisquare Loop Antenna
9.6 Near-Vertical Incidence Skywave (NVIS) Antenna

10 Multiband and Tunable Wire Antennas
10.1 Multiband Antennas
10.1.1 Trap Dipoles
10.1.2 Multiple Dipoles
10.1.3 Log-Periodic Antenna

10.2 Tuned Feeder Antennas
10.2.1 G5RV Multiband Dipole
10.2.2 Longwire Antenna
10.2.3 Off-Center-Fed Dipole
10.2.4 Windom

11 Vertically Polarized Antennas
11.1 Vertical Dipole
11.2 Grounded Vertical Monopole
11.3 Non-Quarter-Wavelength Vertical Monopoles
11.4 Ground Systems for the Grounded Monopole
11.5 Feeding the Ground-Mounted Vertical
11.6 Elevated Ground-Plane Antenna
11.7 Vertical versus Horizontal Polarization
11.8 Supporting the Vertical Antenna

Part III Directional Antenna Arrays

12 Wire Arrays
12.1 Longwire Antennas
12.1.1 Longwire Pattern and Gain
12.1.2 Nonresonant Longwire Antennas
12.2 Arrays of Longwires
12.2.1 Vee Beams
12.2.2 Nonresonant Vee Beams
12.2.3 Rhombics
12.3 Bobtail Curtain
12.4 Half-Square
12.5 20-m ZL-Special Beam

13 Vertical Arrays
13.1 Two-Element Array
13.2 360-Degree Directional Array

13.3 Four-Square Array
13.4 Receive-Only Vertical Arrays
13.5 Grounding

14 The Yagi-Uda Beam Antenna
14.1 Theory of Operation
14.2 Three-Element Yagi Antennas
14.3 Yagi Antennas with Four or More Elements
14.4 Yagi Construction
14.4.1 Rotatable Dipole
14.4.2 Multielement Yagi Construction
14.4.3 Rotatable Three-Element 20-Meter Yagi
14.4.4 Wire Yagis
14.5 Multiband Yagis
14.6 Yagis with Loaded Elements
14.7 Feeding the Yagi
14.7.1 Balun Feed
14.7.2 Gamma Match
14.7.3 T-Match
14.7.4 Hairpin Match
14.8 Moxon Beam
14.9 Stacking Yagi Antennas
14.9.1 Stacking for Gain
14.9.2 Stacking for Limited Space
14.10 Custom Yagi Designs and Further Reading

15 Cubical Quads and Delta Loops
15.1 Cubical Quad
15.1.1 Quad Versus Yagi
15.1.2 Feeding the Quad
15.1.3 Construction Ideas
15.2 Delta Loop
15.3 Diamond Loop

15.4 Concluding Thoughts

Part IV Antennas for Specialized Applications

16 Receiving Antennas

16.1 Beverage or "Wave" Antenna
16.1.1 The Bidirectional Beverage
16.1.2 Phased Beverages
16.1.3 Snake or BOG Antenna
16.1.4 Limitations of the Beverage Antenna
16.2 Receiving Loops
16.2.1 Small Loop Receiving Antennas
16.2.2 Air Core Frame ("Box") Loops
16.2.3 Transformer Loops
16.2.4 Tuning Schemes for Loop Antennas
16.2.5 The Sports Fan's Loop
16.2.6 Shielded Loop Antennas
16.2.7 Sharpening the Loop
16.2.8 Loop Amplifier
16.2.9 Self-Resonance for VLF/LF Loops
16.2.10 Coaxial-Cable Loop Antennas
16.3 The EWE Antenna
16.4 Pennants, Flags, and the K9AY Loop
16.5 Longwire Antennas

17 Hidden and Limited-Space Antennas

17.1 Hidden Antennas
17.2 Limited-Space Antennas

18 Mobile and Marine Antennas

18.1 Mobile Antennas for VHF and UHF
18.2 Mobile Antennas for HF and MF
18.3 Types of Short Mobile Antennas

18.4 Tuning HF Mobile Antennas
18.5 Marine Radio Antennas

19 Emergency and Portable Antennas
19.1 Improving Radiation Efficiency
19.2 Reducing Ground Losses
19.3 Selecting the Right Polarization
19.4 Temporary Supports
19.5 Summary

20 Antennas for 160 Meters
20.1 Vertical Monopoles
20.2 λ/4 Vertical Monopole
20.3 Shortened Verticals
20.4 Loading the Tower
20.5 Insulated-Base Loaded Verticals
20.6 Phased Verticals
20.7 Horizontal Antennas
20.7.1 Shortened Horizontal Antennas
20.8 Inverted-L Antennas
20.8.1 Dual-Band Inverted-L
20.8.2 Phased Inverted-L Antennas
20.9 Flat-Top or T Antenna
20.10 80-Meter Dipole on 160 m
20.11 Receiving on the Transmit Antenna

21 VHF and UHF Antennas
21.1 Ground-Plane Vertical
21.2 Coaxial Vertical
21.3 5/8-Wavelength 2-Meter Antenna
21.4 J-Pole Antennas
21.5 Collinear Vertical
21.6 Grounds for Ground-Plane Verticals
21.7 Yagi Antennas

21.7.1 Six-Meter Yagi
21.7.2 Two-Meter Yagi
21.8 Halo Antennas
21.9 Quad Beam Antennas
21.10 VHF/UHF Scanner Band Antennas
21.10.1 "Scanner-Vision" Antennas
21.10.2 Scanner Skyhooks
21.11 VHF/UHF Antenna Impedance Matching

Part V Microwave Antennas

22 Microwave Waveguides
22.1 Waveguide Basics
22.2 Waveguide Modes
22.3 The TEM Mode
22.4 Rectangular Waveguide
22.5 Boundary Conditions
22.6 Rectangular Waveguide Modes
22.7 Waveguide Excitation
22.8 Bouncing Wave Interpretation
22.9 Velocity and Wavelength in Waveguides
22.10 Cutoff Frequency (f_c)
22.11 Waveguide Impedance
22.12 Waveguide Terminations
22.13 Waveguide Joints and Bends
22.14 Waveguide Coupling Methods

23 Microwave Antennas
23.1 Dipole Antenna
23.2 Horn Antenna
23.3 Reflector Antennas

23.4 Parabolic Dish Antennas
23.4.1 Parabolic Dish Feed Geometries
23.5 Array Antennas
23.5.1 Slot Array Antennas
23.5.2 Electronically Steered Arrays

24 Antenna Noise
24.1 Thermal Noise
24.2 Noise Figure and Equivalent Noise Temperature
24.2.1 Noise Figure (*F*)
24.2.2 Noise Temperature (T_e)

25 Radio Astronomy Antennas
25.1 Signals from the Sun
25.2 Signals from Jupiter
25.2.1 Ring Antenna
25.2.2 Directional Discontinuity Ring Radiator
25.2.3 Helical Antennas
25.2.4 Multiple Helical Antennas
25.3 Interferometeric Arrays
25.4 Dish Antennas

26 Radio Direction-Finding (RDF) Antennas
26.1 Sense Antenna Circuit
26.2 RDF at VHF and UHF
26.3 Fox Hunting
26.4 Tracking Down RFI
26.5 Shortwave and AM BCB "Skip" RDF
26.6 Adcock Antennas
26.7 Watson-Watt Adcock Array
26.8 Doppler RDF Antennas
26.9 Wullenweber Array

26.10 Time Difference of Arrival (TDOA) Array
26.11 Switched-Pattern RDF Antennas
26.12 Double-Ducky Direction Finder (DDDF)
26.13 Summary

27 Antennas for Wireless Devices
27.1 Design Considerations for Small, Low-Gain Antennas
27.2 Planar Antennas
27.3 Patch Antennas
27.4 Planar Inverted F Antenna (PIFA) and Derivatives
27.5 Chip Antennas
27.6 Dielectric Resonator Antenna
27.7 Modules and Integrated Antennas

Part VI Tuning, Troubleshooting, and Design Aids

28 Antenna Tuners (ATUs)
28.1 ATU Circuit Configurations
28.1.1 L-Section Network
28.1.2 Π Networks
28.1.3 Split-Capacitor Network
28.1.4 Transmatch Circuit
28.2 Baluns
28.2.1 Toroid Impedance-Matching Transformers
28.3 Ferrite Core Inductors
28.3.1 Toroidal Transformers
28.3.2 Mounting Toroid Cores
28.3.3 How Many Turns?
28.3.4 Ferrite Rods

29 Antenna Modeling Software
29.1 Antenna Analysis and Synthesis

29.2 Types of Numerical Methods
29.3 Finite Difference and Finite Element Methods
29.4 Method of Moments
29.5 Numerical Electromagnetics Computation (NEC)
29.5.1 A Typical NEC-2 GUI
29.5.2 Advanced Capabilities
29.5.3 Accuracy and Computation Time
29.6 Terrain and Propagation Modeling Software
29.7 Summary

30 Testing and Troubleshooting
30.1 A Review of SWR
30.2 Antenna Analyzers
30.3 Spectrum Analyzers
30.4 Vector Network Analyzers
30.5 Field Strength Meters
30.6 RF Wattmeters
30.7 Testing Coaxial Cable
30.8 Dummy Loads

Part VII Mechanical Construction and Installation

31 Supports for Wires and Verticals
31.1 Antenna Safety
31.2 Wire Antenna Construction
31.2.1 Center Insulators and Feedline Connection
31.2.2 End Insulators
31.2.3 Folded Dipole Construction
31.2.4 Supporting the Wire Antenna
31.3 Trees, Masts, and Other Supports
31.4 Halyards and Support Lines
31.5 Strain Relief
31.6 Installing Vertical Antennas

32 Towers
- 32.1 Types of Towers
- 32.2 Tower Fundamentals
- 32.3 Guyed Towers
- 32.4 Bracketed Towers
- 32.5 Foldover Towers
- 32.6 Crank-up Towers
- 32.7 Self-Supporting Towers
- 32.8 Tapered Cross-Section Towers
- 32.9 Rotating Poles
- 32.10 Tower Bases
- 32.11 Guy Wires and Anchors
- 32.12 Erecting the Tower
- 32.13 Rotators, Masts, and Thrust Bearings
- 32.14 Preowned Towers and Accessories
- 32.15 Additional Information

33 Grounding for Safety and Performance
- 33.1 Grounding for Lightning Protection
- 33.2 Grounding for Power Distribution Safety
- 33.3 Station Grounding
  - 33.3.1 Ground Wire Tuner
- 33.4 Grounds for Antenna Efficiency
- 33.5 Soil Conductivity
- 33.6 Grounding with Radials
- 33.7 Elevated Grounds

34 Zoning, Restrictive Covenants, and Neighbors
- 34.1 Federal Regulations
- 34.2 State Regulations
- 34.3 Local Zoning
- 34.4 Private Sector Restraints

Index

Preface to the Sixth Edition

My introduction to antennas came with the old black and white television set that my siblings and I used to watch *Happy Days*, *The Muppet Show*, and Saturday morning cartoons. The antenna was a few feet of twin-lead antenna wire. I had a bit of a knack for moving the wire into just the right configuration to clear up the snow on the screen and get a watchable image. That, along with taking apart old radios and electronic equipment, got me off to an early start in the field of wireless communication.

My task with this new edition was to preserve the hands-on feel of the book and its endless old-school charm while adding a modern touch. Fortunately, my skills are almost entirely complementary to those of the original author—which is a polite way of saying that I'm an academic researcher with limited experience when it comes to many aspects of everyday radio communications. The goal was to retain the practical focus of the book while allowing the reader some measure of confidence that the book incorporates recent advances in antenna theory, design, and practice.

Changes include moving foundational chapters like impedance matching and the Smith chart earlier in the book and updating the notation and treatment of transmission line theory and antenna analysis. New topics have been added or given increased emphasis. References to older sources and appendices with out of date or easily searchable information have been removed.

Karl F. Warnick
Provo, Utah

Preface to the Fifth Edition

My paternal grandfather was born and raised in Newfoundland, on the outskirts of St. John's, its capital city. As a young man, he came to the United States in 1903 to find work. To his dying day he understood very little about radio or television beyond the simple mechanics of adjusting the volume control or changing the station. Yet he became directly responsible for my nearly lifelong love affair with "wireless" communications when he gave me an AM broadcast band clock radio for my 13th birthday. It was my very first radio receiver, and, although my grandfather thought I would prize the alarm on it, the "magical" feature I came to love instead was the Sleep knob. Now I could climb into bed and fall asleep to the sounds of faraway stations, my parents and I secure in the knowledge that the radio would turn itself off.

As I gained familiarity with my newfound window on the world, I marveled at the distant stations I could hear on certain nights, and soon I was keeping a paper diary, or log, of broadcast station call letters I was able to pluck from the cacophony of late-night radio waves. Without knowing the name of my obsession, I had become a broadcast band DXer! ("DX" is radio shorthand for "long distance.")

At about the same time, my next-door neighbor obtained his amateur radio license, and soon I was spending occasional evenings in his basement, sitting silently by his side while he made two-way contacts with like-minded hobbyists around the world. The seed was sown, and with the help of the local radio club I obtained my own "ham radio" license before I turned 14. I had plunged headlong into a passionate love affair with a technology that has shaped my life ever since, from my choice of college major to my choice of career and even my choice of where to live on this planet.

But it wasn't until I was in the midst of preparing new material for this fifth edition of *the Practical Antenna Handbook* that I realized

my grandfather had another, albeit indirect, connection with the world of radio I had so quickly taken to.

Picture the world in the fall of 1901: The only way to travel between North America and Europe is by giant steamship; the Wright brothers haven't yet made their first flight on the sands of the North Carolina Outer Banks; there is no such thing as radio or TV broadcasting; it will be more than a half century before the first artificial satellite beams its telemetry data down to Earth; and it will take three-quarters of a century for the Internet to be born. In 1901, the only way to communicate across the sea is by sending mail on the oceangoing steamships or by wire-line telegraphy, using the relatively new undersea cables that had been laid in the late 1800s.

In Europe, meanwhile, young inventor Guglielmo Marconi is having scant success convincing governments and corporations of the value of his recently discovered ability to signal across towns and villages without wires. Discouraged with his native Italy's lackluster response to his experiments, he has come to England, where he has received encouragement from that government—especially after his success in communicating across the English Channel.

But something more is needed to capture the public's attention. So Marconi boards a ship and, with some of his assistants, heads to Newfoundland after setting up a transmitting station overlooking the Atlantic near the westernmost point of England, at Poldhu, Cornwall. On December 11, 1901, Marconi and his North American team hoist a wire with a large kite launched from Signal Hill at the entrance to the bay at St. John's, and on the following day Marconi reports hearing the letter "S"—the prearranged code letter being sent by other assistants back at Poldhu—in the headphones of his receiving apparatus! As they say, the rest is history.

Did my grandfather ever meet Marconi? Was he aware of the transatlantic experiments being conducted within a few miles of his home? Did he visit Signal Hill to see what all the fuss was about? I'll probably never know. But it's fun to speculate.

In distinct contrast to water waves and sound waves, or to the Internet, which moves information from place to place largely via optical fibers and other physical media, radio communication takes

place with no supporting medium at all! Electromagnetic waves launched at the transmitting site travel essentially forever, until intercepted at a receiving site. The task of efficiently converting transmitter output power into radio waves capable of traveling into the infinite reaches of space or of picking up those waves for amplification and conversion to audible sound waves by my clock radio falls to a class of devices known as antennas.

It is no exaggeration to say that without antennas, there would be no radio. Thus, to design radio or television equipment, to be an amateur radio operator or shortwave listener, or to successfully install a wireless Internet service for a group of users, a practical understanding of antenna basics and basic antennas is a necessity—hence the title of this book.

In the decade since the fourth edition of the *Practical Antenna Handbook* first appeared, entirely new classes of specialized antennas—including medium-frequency (MF) directional receiving antennas for limited spaces and budgets—have come into existence and are in daily use by thousands of radio amateurs and shortwave listeners around the world. For many older antenna designs, theories of antenna operation originating in crude experiments from as long as a century ago have been either confirmed or supplanted by new theories and practices based on modern analytic techniques, coupled with tests conducted with new generations of test equipment, and backed up by extensive computer modeling using inexpensive and readily available software that until just a few years ago was found only in the laboratories of universities and large corporations.

Scientific and technical information transfer has also been upended by the Internet. Today, numerous online discussion groups allow experts to report and critique recent antenna findings in near real time while hundreds or thousands of other professionals and rapt hobbyists "listen in."

In putting together the fifth edition, every chapter of the *Practical Antenna Handbook* has been reviewed and revisions made wherever appropriate. Occasionally, antenna designs or topics that are too esoteric or arcane for the overview of basic antenna types this book represents have been winnowed out to make room for new material

of broader or more topical interest. In many chapters, new explanations of fundamental concepts are the result of developing and refining effective (i.e., "lightbulb") answers to the most-asked questions during antenna talks I have long given at radio club meetings and conventions. Similarly, the use of the half-wave dipole as a conceptual building block for "quick and dirty" analysis of many other antenna types found throughout this book has been well-received over the years.

With few exceptions, the use of math in this book is kept at or below high school level. Perhaps the only "advanced" math concept in the book is the use of the imaginary number operator *j* to simplify and clarify the manipulation and display of complex numbers.

Two structural changes have been made to the *Practical Antenna Handbook*:

- The order of presentation of the material has been revised to collect related topics in groups of chapters and to help the book as a whole flow logically from the most basic concepts and antennas to more complicated configurations and techniques.
- The fourth edition of the Practical Antenna Handbook included a CD with supplemental information and spreadsheets on it. Today it is far more practical to commit this material—and more—to a Web site, where it can be updated and augmented, as appropriate, over the lifetime of the book. Go to www.mhprofessional.com/carr5 for:
    - Tables of worldwide geographic coordinates and antenna dimensions vs frequency
    - Supplier updates
    - Author's blog
    - Additional photographs and schematics
    - Links to tutorials and specialized calculators

Before I became captivated by radio, my goal had been to be a journalist. Throughout my adult life I have never lost that love of writing, and during my career I have occasionally had the good fortune to be able to marry the two interests in a few brief work assignments. Ordinarily, writing a book about antennas would represent nirvana for me, except that I cannot lose sight of the fact that I have this opportunity only because Joe Carr passed away shortly after completing the fourth edition. So I hope you will find this fifth edition a worthy expansion and updating of Joe's fine work over the years, and I ask your indulgence as I dedicate it to my grandfather, Archie Munn Hippisley, who started me on this path many years ago.

George W. (Bud) Hippisley, W2RU
Old Forge, New York

Acknowledgment

As with other fields, in engineering and the sciences virtually all advances in knowledge build upon accomplishments and results that are the work of others. Thus, an overview volume such as this must necessarily draw upon and reflect the collective wisdom of the larger community of contributors to the state of the art. To identify all such individuals would require a second book, even larger than this one.

However, certain people who provided specific guidance, inspiration, and resources should not go unrecognized. In particular:

Guy Olinger, K2AV; Jim Lux, W6RMK; Tom Rauch, W8JI; Jim Brown, K9AY; and Chuck Counselman, W1HIS—exemplary members of a cadre that is advancing the collective state of the art in radio communications technology by freely disseminating and sharing their considerable knowledge, analytic skills, and experimental results with others via the Internet;

W. L. Myers, K1GQ, for a near-lifetime of stimulating dialog and many kindnesses, and for introducing me to *cocoaNEC*, antenna modeling software for the Macintosh by Kok Chen, W7AY;

Roy Lewallen, W7EL, the author of EZNEC, and Dan Maguire, AC6LA, for NEC-2 modeling guidance;

Professors Amir Fariborz (SUNY Institute of Technology) and Orlando Baiocchi (University of Washington/Tacoma Institute of Technology) for making electromagnetics interesting and a lot more fun during my second "tour" around the EM classroom; and

James Rautio, AJ3K, author of the first antenna modeling software (*Annie*) I ever used, who has leveraged his unparalleled appreciation of James Clerk Maxwell's accomplishments into

concrete projects to help preserve Maxwell's legacy for generations to come.

In amateur radio, those who provide assistance to newcomers to the hobby are known as *Elmers*. Some of us can point to a single such Elmer; others, to many. Long ago, I was fortunate to have many, virtually all of whom have since passed on. In the area of antennas, one deserves posthumous recognition:

1. Carl Herrling, ex-K2TVT, who helped a 13-year-old neophyte radio amateur and budding author construct and erect his very first antennas—including a three-wire dipole for 80 meters!

On the publishing side of this project, I must mention:

1. Tom Nutt-Powell, an author in an entirely unrelated field, whose one-sentence motivational "mantra" (that I can't repeat here) kept me on track when it seemed as though I would never find the right words—or, for that matter, *any* words;
2. Ginny Carroll, whose efficiency, competence, and shared love of the English language made the final stages of the review and editing process a joy; and
3. Judy Bass, who enthusiastically carried the ball for this project within McGraw Hill, and who worked with me to develop a schedule we all could live with.

Finally, my entire family has been enthusiastic beyond belief, but two members in particular deserve specific praise:

1. Nathan Dougall, my stepson, whose gift to me of the fourth edition of this book started this whole project, even as he embarked on his own book; and
2. Linda, my wife, who was instantly supportive at the onset of this project, and who has since administered the loving

but firm prodding I periodically needed.

Bud Hippisley

PART I

Fundamentals

CHAPTER 1
Introduction to Radio Communications
CHAPTER 2
Radio Waves
CHAPTER 3
Radio Wave Propagation
CHAPTER 4
Ionospheric Propagation
CHAPTER 5
Antenna Basics
CHAPTER 6
Transmission Lines and Impedance Matching
CHAPTER 7
Antenna Arrays

CHAPTER 1
Introduction to Radio Communications

Useful communication by radio has been with us for the entire twentieth century. Originally determined to be mathematically possible by the Scottish inventor James Clerk Maxwell in the mid-1860s, radio waves remained science fiction until 1887, when Heinrich Hertz was first to convert Maxwell's theoretical equations into laboratory experiments that demonstrated and confirmed their existence. Guglielmo Marconi picked up the ball and through the 1890s took the first productive steps toward commercializing radio communication. Nikola Tesla from Serbia, Alexander Fessenden from Canada, Alexander Popov from Russia, and Lee DeForest from the United States were early pioneers in advancing the state of the art. Thus began the international story of radio.

Early wireless was widely seen as a neighborhood or cross-town endeavor of limited usefulness, but that changed on a fateful December day in Newfoundland when the Morse letter "S" tickled Guglielmo Marconi's ears. When Marconi successfully demonstrated wireless telegraphy as a commercially viable technology, radio communications began in earnest. By the turn of the century, "wireless telegraphy" (as radio was called then) was sparking (pun intended) the imaginations of countless people across the world. The "wireless" aspect so radically changed communications that the word is still used to denote all radio communications.

Wireless telegraphy was immediately pressed into service by shipping companies because it provided an element of safety that had been missing in the pre-wireless days. Early shipping companies advertised that their ships were safer because of the new wireless installations aboard. Wireless telegraphy proved its usefulness on the high seas in 1909 when two ships collided in the foggy Atlantic Ocean and were sinking. Passengers and crew members of both

ships were in imminent danger of death in the icy waters, but radio operator Jack Binns became the first man in history to send out a maritime distress call. His call for help in Morse code was received and relayed by nearby ships, allowing another vessel to come to the aid of the stricken pair of ships.

Radio communication prior to about 1916 was carried on via telegraphy (i.e., the on-off keying of a radio signal in the Morse code). But one night in 1916 there was more magic. Radio operators and monitors up and down the Atlantic seaboard—from the Midwest to the coast and out to sea for hundreds of miles—heard something that must have startled them out of their wits: crackling out of the "ether," amid the whining of Alexanderson alternators and the *zzzchht* of spark-gap transmitters, came a new sound—a human voice. Engineers and scientists at the Naval Research Laboratory in Arlington, Virginia, had transmitted the first practical amplitude-modulated (AM) radio signals. Earlier attempts had been successful as scientific experiments but did not employ commercially available equipment.

Radio activity in the early years was unregulated and chaotic. Today it is quite structured and heavily regulated. At the international level, radio is coordinated by the International Telecommunications Union (ITU) in Geneva, Switzerland, based on treaties signed at World Administrative Radio Conferences (WARCs) held every few years. Each sovereign country retains control of its own internal radio regulations. In the United States, matters pertaining to radio communications are handled by the Federal Communications Commission (FCC), headquartered in Washington, DC.

Along with commercialization of radio communications came a group of dedicated amateur radio operators. Amateur radio has grown from a few thousand "hams" prior to World War I to more than 900,000 today, about one-third of them in the United States. Amateurs were ordered off the air during World War I, and the hobby almost did not make a comeback after the war. There were, by that time, many powerful commercial interests greedily coveting the frequencies used by amateurs, and they almost succeeded in

keeping postwar amateurs off the air. But the amateurs won the dispute—at least partially.

In those days, low frequencies with wavelengths longer than 200 m (i.e., 20–1500 kHz) were considered most valuable for communications. The cynical attitude attributed to the commercial interests regarding amateurs was "Put 'em on two hundred meters and below . . . they'll never get out of their backyards there!" But there was a surprise in store for those commercial operators. Wavelengths shorter than 200 m are in the frequency region we now call shortwaves. In 1919, it was not suspected by anyone that shortwaves could be used for anything other than local communication. Today, shortwaves are well known for their ability to support communications over transcontinental distances.

One of the authors of this book once heard an anecdote from an amateur radio operator "who was there." In the summer of 1921, this man owned a beautiful, large, wire "flattop" antenna array for frequencies close to 200 m on his family's farm in southwestern Virginia. Using those frequencies, he was accustomed to regularly communicating several hundred miles into eastern Ohio and down to the Carolinas. Returning home at Thanksgiving of 1921 during his first semester at college, he noticed that his younger brother had replaced the long flattop array with a short dipole antenna. Confronting his brother over the incredible sacrilege, he was told that they no longer used 150 to 200 m but were using 40 m instead. Everyone "knew" that 40 m was useless for communications over more than a few blocks, so (undoubtedly fuming) the older sibling took a turn at the key. He sent out a "CQ" (general call) and was answered by a station with a call sign similar to "8XX." Thinking that the other station was in the eighth US call district (at the time covering West Virginia, Ohio, Michigan, and western New York), he asked him to relay a message to a college buddy in Cincinnati, Ohio. The other station replied in the affirmative, but suggested that ". . . you are in a better position to reach Cincinnati than I, I am *French* 8XX." Call signs in 1921 did not yet include any system of national prefixes, such as we have today. The age of international amateur communications had arrived! And with it came a new problem:

National identifiers in call signs became necessary, which is why all US call signs begin with K, W, N, or portions of the A block.

During the 1930s, radio communications and broadcasting spread rapidly as technology and techniques improved. World War II became the first war to be fought with extensive electronics. Immediately prior to the war, the British developed a new weapon called radio detection and ranging, or radar as it is now called. This tool allowed them to be forewarned of German aircraft streaming across the English Channel, planning to strike targets in the United Kingdom. The German planes were guided by (sophisticated for their times) wireless highways in the sky, while British fighters defended the home islands by radio vectoring from ground controllers. With night fighters equipped with the first centimeter wavelength (i.e., microwave) radar, the Royal Air Force was able to strike the invaders accurately—even at night. The first kill occurred one dark, foggy, moonless night when a *Beaufighter* closed on a spot in the sky where the radar in the belly of the plane said an enemy plane was flying. Briefly thinking he saw a form in the fog, the pilot cut loose a burst from his quad mount of 20-mm guns slung in the former bomb bay. Nothing. Thinking that the new toy had failed, the pilot returned to base— only to be told ground observers had reported that a German *Heinkle* bomber fell from the overcast sky at the exact spot where the pilot had his ghostly encounter.

In the intervening years, wireless technologies have exploded. Over short ranges, radio frequency signals link a wide variety of accessories over distances from a few inches to several feet or more to computers and mobile devices. Medium-range systems including wireless networking and cellular communications are as commonly used and as familiar as breathing. Long-range radio communication systems are used daily by people all around the world. Global positioning systems use radio frequency signals to provide location information to drivers, hikers, pilots, and mobile device users across the globe. Voice, data, and video services have been available for many years from dish antennas aimed at satellites in geosynchronous orbit 23,000 mi from the earth’s surface. More recently, medium earth orbit (MEO) and low earth orbit (LEO) have

opened up a wide range of new sensing technologies and radio communication services, with smallsats and cubesats dramatically reducing the cost of spaceborne radio transceivers. Amateur operators are able to communicate worldwide on low power and have even launched their own satellites.

Whether visible or integrated inside a wireless device, communications systems require antennas at both ends of the channel. The antenna is one of the most important parts of receiving and transmitting systems (Fig. 1.1). That is what this book is all about.



FIGURE 1.1 A variety of radio antennas on Sandia Peak, near Albuquerque, New Mexico. Photograph by Daniel Mayer, distributed under a CC BY-SA 3.0 license.

CHAPTER 2
Radio Waves

Today it is well recognized that radio signals travel in a wavelike manner, but that fact was not always so clear. It was well known in the first half of the nineteenth century that wires carrying electrical currents produced a magnetic field surrounding the wire that was capable of exerting a force over short distances. This knowledge formed the basis for the invention of electrical motors and generators. But it was not until 1887 that German physicist Heinrich Hertz succeeded in demonstrating the ability to transmit and receive radio signals between two separate electrical circuits in his laboratory. In so doing, he experimentally confirmed Maxwell's theoretical work from a quarter century earlier, predicting the existence of previously unimagined radio signals that, like light, were electromagnetic waves.

Like the induction field, the electromagnetic wave is created by an electrical current. Unlike the induction field, the electromagnetic wave requires a changing electric current. And once launched, the radiated field further differs from the induction field in that it is self-sustaining and no longer depends on the existence of a conductor or a current for its path or its amplitude. Once it leaves the conductor, the radiated electromagnetic wave propagates through space according to the universal equations governing wave motion of all kinds.

2.1 Water-wave Analogy

Radio wave propagation can be visualized with a water-wave analogy. Although not a perfect analogy, it serves to illustrate the point that a wave represents an energetic disturbance moving away from a source. Figure 2.1 shows a body of water into which a ball is

dropped (Fig. 2.1*a*). When the ball hits the water (Fig. 2.1*b*), it displaces water at its point of impact and pushes a leading wall of water away from itself. The ball continues to sink (Fig. 2.1*c*) and the wave propagates away from it until the energy is dissipated. Although Fig. 2.1 shows the action in only one dimension (a side view), the actual waves propagate outward in all directions, forming concentric circles when viewed from above.

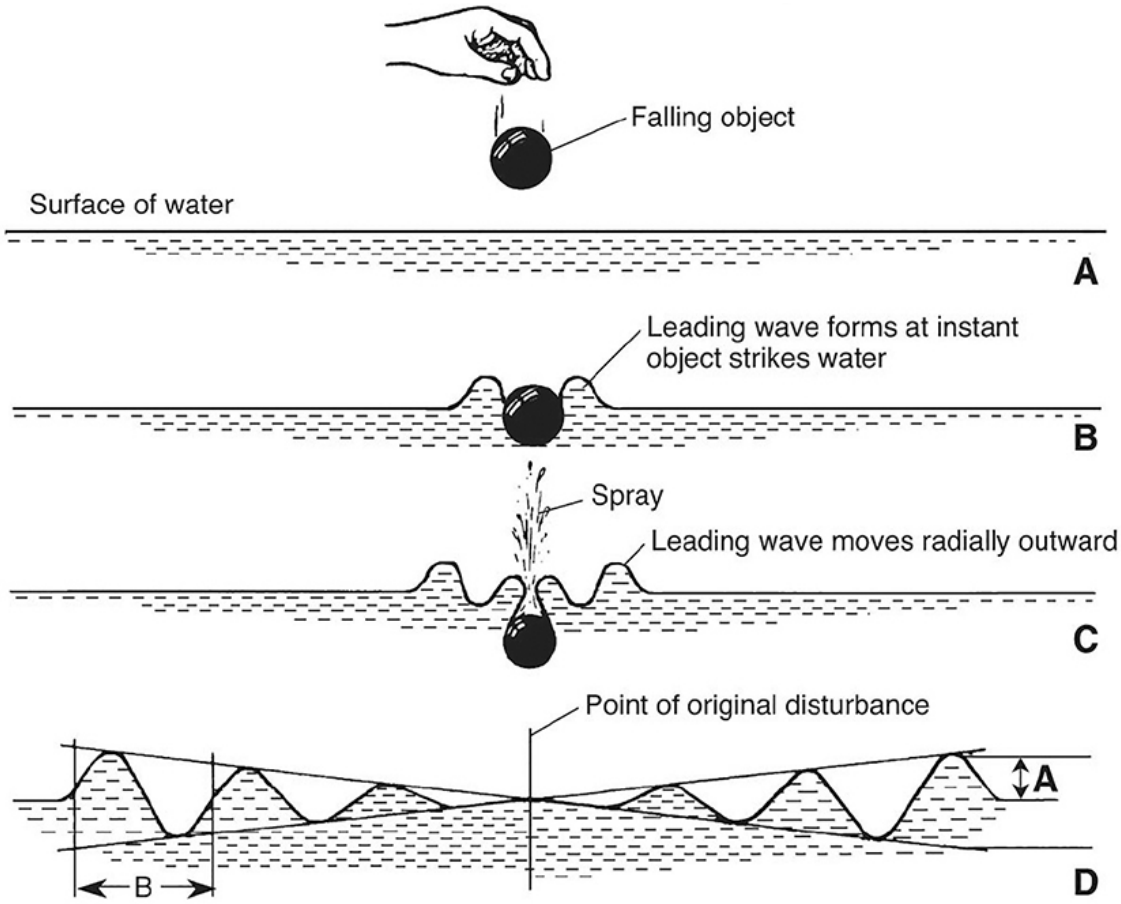
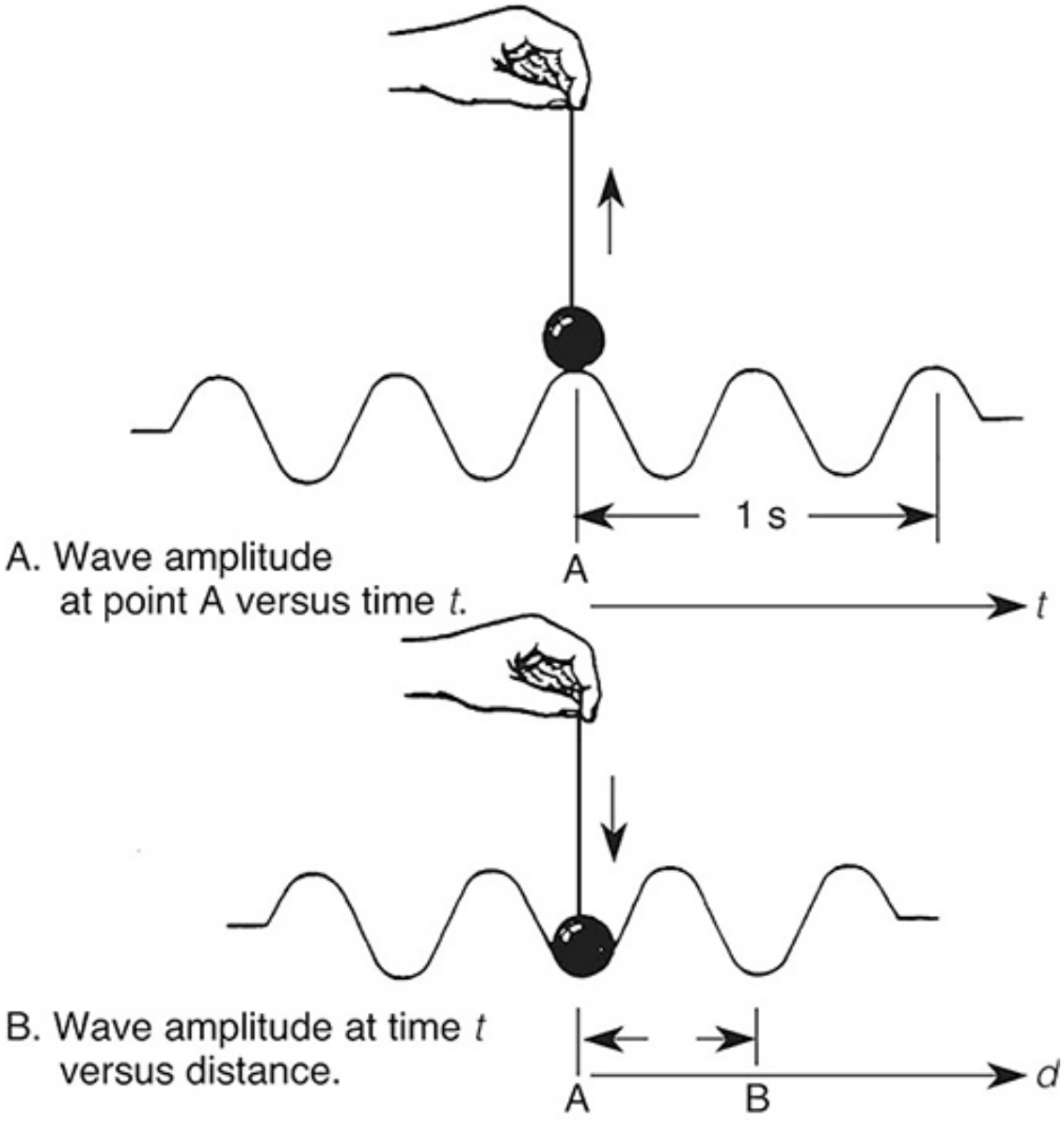


FIGURE 2.1 A ball dropped into water generates a wavefront that spreads out from the point of original disturbances.

As the wave produced by the dropped ball moves away from the point of origin, it is damped—that is, it will decline in amplitude on

successive crests until the energy is dissipated by friction between adjacent water molecules and the wave ceases to exist. To make the analogy to radio waves more realistic, the wave must be generated in a continuous fashion, as shown in Fig. 2.2*a*: A ball is dipped into a body of water at a cyclic rate, successively reinforcing new wave crests on each dip. New waves continue to be generated at point A as long as the ball continues to oscillate up and down. The result is a continuous wave train spreading out through the water.



F IGURE 2.2 A bobbing ball on a string demonstrates continuous wave generation. (A) Wave amplitude at point A versus time *t*. (B) Wave amplitude at time *t* versus distance *d*.

An interesting aspect of the water-wave analogy is that individual water molecules show little horizontal movement across the surface of the water. Each water molecule primarily moves up and down with only a small amount of horizontal displacement. Nonetheless, a clearly visible wavefront travels outward from the point where the ball was dropped. The disturbance associated with the wave and

the energy carried by the disturbance moves without carrying water molecules along with it. Similarly, the motion of a radio wave does not require that electrons travel with it. We shall see in our discussion of basic antennas that maximum radiation occurs at right angles to the direction of travel of electrons along a wire antenna.

2.2 Frequency and Wavelength

The quantitative properties of a wave are measured by its frequency (*f*) and wavelength (*λ*). The frequency is the number of oscillations (or cycles) in the wave's amplitude per unit of time. Frequency is often expressed as the number of cycles per second (cps). If the time between successive positive peaks at any single point in space is 0.5 s, as suggested by Fig. 2.2*a*, exactly two complete amplitude cycles of the wave occur in 1 s, so the frequency of the wave created by the oscillating ball is 2 cps. We can make the measurement from consecutive observable points elsewhere on the waveform, such as negative peaks, zero crossings, etc. In honor of Heinrich Hertz, the unit of frequency was renamed the Hertz (Hz) many years ago. Since the units are equivalent
$(1\ \mathrm{Hz} = 1\ \mathrm{cps})$
, the wave in Fig. 2.2 has a frequency of 2 Hz.

Because the frequencies best suited for efficient radio propagation are so high, they are usually expressed in kilohertz ($\mathrm{kHz} = 1000\mathrm{s}$ of Hz), megahertz ($\mathrm{MHz} = 1{,}000{,}000\mathrm{s}$ of Hz), or gigahertz ($\mathrm{GHz} = 1{,}000{,}000{,}000\mathrm{s}$ of Hz). Thus, the frequency of the Boston radio station WBZ, operating near the middle of the amplitude modulation (AM) broadcasting band, can be properly expressed as 1,030,000 Hz, 1030 kHz, or 1.03 MHz, all of which are equivalent and equally correct.

The wavelength of any wave is the physical distance between adjacent corresponding points of a spatial waveform, measured at the exact same time. In Fig. 2.2*b*, the wavelength (*λ*) is drawn as the distance between successive negative peaks in any direction from the source, but we could also measure the wavelength from the distance between successive positive peaks, or between any two corresponding features on successive waves as we move outward from the source in a straight line.

Wavelength is proportional to the reciprocal of the frequency. The relationship is given by

$$f\lambda = v$$

, where *f* is the frequency in Hz, *λ* is the wavelength in meters, and *v* is the velocity of propagation in meters per second (m/s). In each case, *v* depends on certain characteristics of the medium (water, air, etc.) in which the wave propagates. In water at room temperature, the velocity of a sound wave is typically 1482 m/s, or about 4861 ft/s—just a bit less than a mile per second. In air, sound waves travel at about 343 m/s, or 1125 ft/s. In contrast, radio waves in free space propagate at the speed of light—approximately 300,000,000 m/s (or 186,000 mi/s) in free space and almost exactly the same in the earth's atmosphere. For historical reasons, the lowercase letter *c* is usually used to represent the speed of light. Thus:

$$f = \frac{c}{\lambda} \simeq \frac{300{,}000{,}000}{\lambda} \quad (2.1)$$

where *f* is in Hertz (Hz), *λ* is in meters (m), and *c* is in meters per second (m/s).

For convenience, these equations are often recast with frequency in kilohertz or megahertz:

$$f(\mathrm{kHz}) = \frac{300{,}000}{\lambda} \quad (2.2)$$

$$f(\mathrm{MHz}) = \frac{300}{\lambda} \quad (2.3)$$

Thus, a wavelength of 300 m corresponds to a frequency of 1 MHz (the middle of the AM broadcast band), while a wavelength of 10 m corresponds to 30 MHz (just above what we call the amateur 10-m band, which lies between 28.0 and 29.7 MHz). While radio receiver tuning dials in North America are usually calibrated in kilohertz or megahertz, in Europe and the rest of the world it is not uncommon to find radio dials calibrated in meters, the universal unit of wavelength, as well as in frequency.

When a wave of any type—be it a mechanical wave, sound wave, or radio wave— passes from one linear medium to another, only its wavelength changes, not its frequency. A 1-kHz tone in the air will still be a 1-kHz tone in water, whereas a radio signal having a wavelength of 40 m in free space may have a wavelength of 30 m in one of the solid-dielectric transmission lines discussed in Chap. 6.

The place where the water-wave analogy breaks down most profoundly is in the need for a medium of propagation. Water waves propagate by actually moving water molecules; that is, water is the medium through which the wave propagates. Similarly, sound travels to your ears by vibrating the molecules of the air or water between you and the guitar string or other source of the mechanical vibration. Maxwell's equations make it clear that electromagnetic waves can exist in laboratory vacuum jars and in free space. Neither Maxwell nor other scientists of the late 1800s could conceive of the "action at a distance" in a vacuum predicted by his equations, so they invented a hypothetical medium called the ether (or *æther*) for transporting electromagnetic waves, including radio waves and light, from one place to another. A profoundly important experiment conducted in the late nineteenth century by American physicists Michelson and Morley conclusively disproved the existence of an ether, yet scientists, engineers, the press, and the lay public continued to refer to a medium for electromagnetic wave propagation until at least the 1920s. Even today, radio enthusiasts refer to the "stuff" out of which radio waves arrive as the "ether," but the term is merely an archaic, linguistic echo of the past. The medium in which electromagnetic waves propagate can

include air, earth, water, and other materials, but it may also just be empty space.

2.3 Properties of Electromagnetic Waves

The study of electromagnetic wave radiation and transmission is typically a semester-long, calculus-intensive upper-class college course. Here we shall limit ourselves to summarizing some of the key issues of electromagnetism central to our understanding of how antennas work. Our goal here is to help you develop an intuitive feel for the generation and propagation of electromagnetic fields in the radio portion of the electromagnetic spectrum without subjecting you to advanced math.

2.3.1 Radio Waves Are the Result of Accelerated Electrical Charges

Acceleration is the rate of change of a body's velocity and velocity is the rate of change of a body's position. Electrons traveling at a constant velocity along a wire create a constant current in that wire. Even though stationary electric charges or constant (or direct) currents may create static or slowly changing electric or magnetic fields, they do not generate radio waves.

An alternating current in a wire or other conductor accelerates the charges on the wire and has the potential to radiate, but a few hurdles have to be overcome. For one thing, the acceleration undergone by electrons in response to the application of an alternating voltage is proportional not only to source amplitude but also to the frequency of the source. High-frequency circuits radiate waves more easily than low-frequency ones. Further, accelerated charges must also be hosted by a physical structure that promotes EM radiation by forcing, if you will, the electric and magnetic fields to leave the vicinity of the radiating structure. That's what distinguishes an efficient antenna from a lumped circuit component or a transmission line, and why antenna geometry is so important.

2.3.2 Electromagnetic Waves Have a Finite Speed

Even though the speed of light (c) is extremely high, it's still a finite number. It takes, for instance, about 8 minutes for light emitted from the surface of our sun to reach us $(93{,}000{,}000\ \mathrm{mi}/186{,}000\ \mathrm{mi/s} = 500\ \mathrm{s})$. Because there is a finite time delay before the radiated fields from an accelerated electron reach a distant point, the radiated wave from an accelerated electron exists as a real entity, carrying energy from source to receiver,

2.3.3 Radio Waves Are Different from Static or Quasi-Static Charge Fields

Many of us have experienced a variety of high school experiments helping us to visualize the strength and spatial distributions of electric and magnetic fields. Iron filings on a sheet of paper with a bar magnet underneath, for instance. While these demonstrations help us understand that invisible entities called fields exist, static or quasi-static (slowly varying) fields are different from radio waves in many respects:

- The static fields around electric charges are radially directed outward from the charge sources, like needles from a porcupine's body. Static magnetic fields wrap around a straight wire carrying a constant current. A static electric field can exist without a corresponding magnetic field, and vice versa. In contrast, a radio wave requires the simultaneous existence of linked time-varying electric and magnetic fields at right angles to each other.
- In the absence of radiation, electric field lines must start and end on charged bodies. But the field lines for radio waves become detached from their source and travel through free space closed upon themselves.
- An antenna is a physical conducting structure deliberately designed to facilitate the departure into space of linked time-varying electric and magnetic fields originating in the

accelerated charges (resulting from time-varying currents) on the conductor. Sometimes poorly laid-out circuits become unintentional radiators.

- A radio wave propagates outward from the radiator with its electric and magnetic fields at right angles to the direction of wavefront motion and to each other. Radio waves propagating through free space are transverse electromagnetic (TEM) waves, consisting of mutually perpendicular electric and magnetic fields (see Fig. 2.3) varying and traveling together in synchronism.

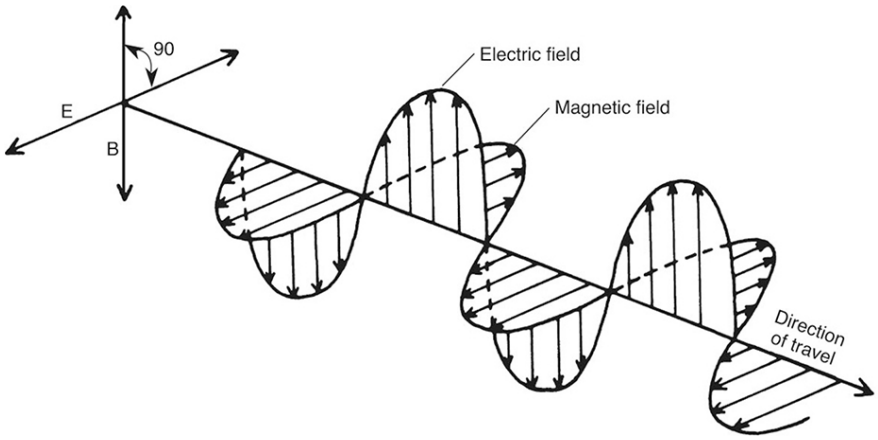


FIGURE 2.3 A transverse electromagnetic (TEM) wave consists of electric and magnetic fields at right angles to each other and to the direction of wave propagation.

- The strength of static and quasi-static electric and magnetic fields falls off as $1/r^2$, where *r* is the distance from the source to a receiver at a distant point. In sharp contrast, the strength of a TEM radio wave falls off according to a slower decay rate with distance, $1/r$. We can detect radio waves from accelerating charges much, much farther away than we can detect the effects of the static charge distributions (such as by observing iron filings around a magnet).

- Radio waves are exactly like light—whether visible, infrared (IR), or ultraviolet (UV)—and other free-space radiation except for their frequency and wavelength. Radio waves occupy the lowest frequencies of the entire electromagnetic (EM) spectrum; hence, they have much longer wavelengths than the other forms of radiation we encounter.
- Radio waves are real, not just artificial or abstract concepts to help our visualization. Radio waves carry energy with them, expanding outward in all directions from the source. When a radio wave encounters a receiving antenna, it delivers energy to that antenna and to any associated load. If the load impedance includes a real, or resistive part, real power is dissipated in it—power that can only have come from the arriving wave.

2.4 Radio Wave Intensity

In free space, a radio wave expanding outward over an ever-growing spherical surface centered on the source (accelerated electrons on a conductor) neither gains nor loses energy. Thus, with the passage of time and distance, the original amount of energy is spread over a larger and larger surface (the wavefront). A helpful analogy is to visualize the process of blowing up a balloon. The total energy in the wavefront is analogous to the total rubber in the skin of the balloon. As you blow into the balloon, the balloon expands and the skin is stretched and becomes thinner. Similarly, as the wave moves away from the source, the radio wavefront occupies a larger and larger spherical surface. The total power spreads out over a larger sphere and the power per unit area of the surface at a point on the sphere becomes smaller.

Thus, even though there are no obstacles in free space to dissipate the energy in the wavefront, if we could measure the amplitude of the passing wavefront at points progressively farther from the source we would see that it decreases with distance from

the source. The radio wave decays or is reduced in apparent power as it propagates from the transmitter to the receiver. Because a fixed amount of power in the wavefront is spread over a larger and larger surface area, the amplitude of the wave is inversely proportional to the total area of the sphere's surface. That area is proportional to the square of the radius of the surface from the source, so radio waves at all frequencies suffer losses according to the inverse square law.

Let's take a look radio wave decay more closely. Electric fields (E-fields) are measured in units of volts per meter (V/m), or in the subunits millivolts per meter (mV/m) or microvolts per meter (μV/m). If a vertically oriented E-field of 10 mV/m propagates across your body and you are about 2 m tall, a head-to-toe voltage difference of

$$(2\ \mathrm{m}) \times (10\ \mathrm{mV/m}) = 20\ \mathrm{mV}$$

, is generated. As a radio wave travels outward from its source, the electric field vector falls off in direct proportion to the distance traveled. The reduction of the E-field is linearly related to distance (i.e., if the distance doubles, the E-field voltage vector halves). Thus, a 10-mV/m E-field that is measured 1 mi from the transmitter will be 5 mV/m at 2 mi. This means less voltage induced by the wave at the terminals of a receiving antenna a weaker signal as the receiving antenna moves away from the source.

The power in an electrical system with a time-harmonic or sinusoidal voltage waveform over time is

$$P = \frac{V^2}{2R} \quad (2.4)$$

where *R* is the resistance in ohms (Ω) and *V* is the peak electrical potential in volts (V).

In the case of a radio wave, the *R* term is replaced with the intrinsic impedance (Z or η) of free space, which is approximately 376.7 Ω, and the voltage is replaced by the electric field intensity. If the peak E-field intensity is 10 V/m, then the power density of the signal is

$$S = \frac{(10\ \mathrm{V/m})^2}{2 \cdot 377\ \Omega} \simeq 0.133\ \mathrm{W/m^2} = 13.3\ \mathrm{mW/m^2} \quad (2.5)$$

As the radio wave propagates away from the source, the electric field intensity decays inversely with distance. Since the power density involves the square of the electric field intensity, the power density, measured in watts per square meter ($\mathrm{W/m^2}$), falls off according to the square of the distance.

This phenomenon is shown graphically in Fig. 2.4. Here, you can see a lamp emitting a light beam that falls on a surface (*A*), at distance *L*, with a given intensity. At another surface (*B*) that is 2*L* from the source, the same amount of energy is distributed over an area (*B*) that is four times as large as area *A*. Thus, the power density falls off according to $1/r^2$, where *r* is the distance from the source. This is called the inverse square law.

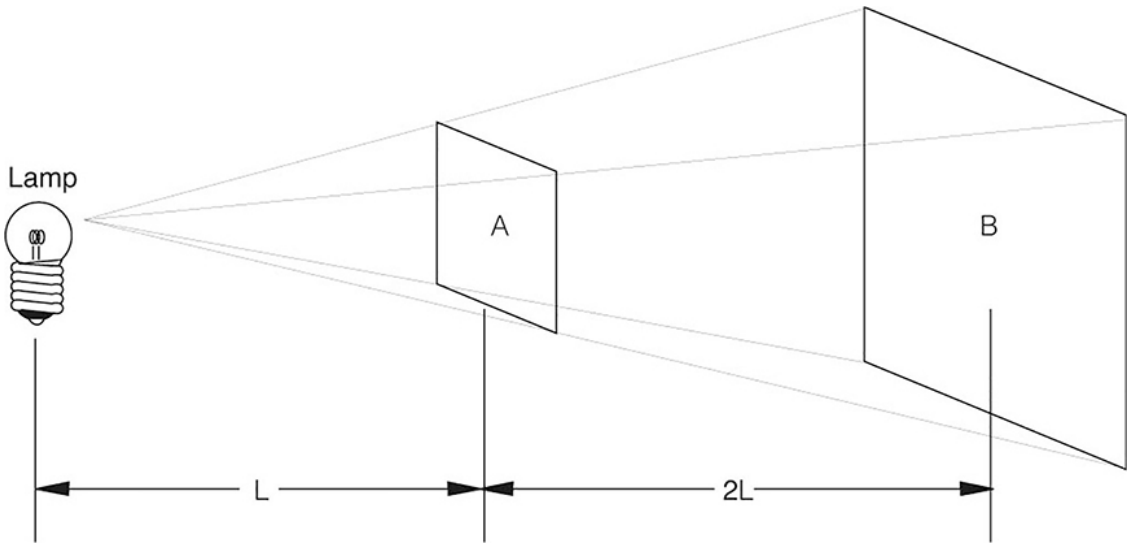


FIGURE 2.4 Wave propagation obeys the inverse square law.

An advancing wave can experience additional reductions in amplitude when it passes through matter. In addition to the decay in power density due to spreading of the wave, dissipative attenuation occurs. At microwave frequencies attenuation is caused by the oxygen and water vapor content of the air surrounding our globe. At other frequencies, losses originating in other materials can be found. These effects are dependent on the relationship of molecular distances to the wavelength of the incident wave. For many frequencies of interest, the effect is extremely small, and radio waves in the atmosphere behave almost as though they were in the vacuum of free space. Some materials, such as wood, that are opaque throughout the spectrum of visible light are fairly transparent at most radio frequencies.

2.5 Phase Velocity

As mentioned earlier in this chapter, an EM wave travels at the speed of light, designated by *c*, which is about 300,000,000 m/s (or 186,000 mi/s). The velocity of the wave slows in dense media, but in air the speed is so close to the "free-space" value of *c* that the same velocity is used for both air and the near-perfect vacuum of outer space when solving most practical problems. In pure water, which is much denser than air, the speed of radio signals is about one-ninth that of the free-space speed. (Contrast this with sound waves, which travel almost four times faster in water than in air.) This same phenomenon shows up in the velocity factor (v_F) of transmission lines. In coaxial cable, the value of v_F is about 0.80, which means a signal propagates along the line at a speed of 0.80*c*, or 80 percent of the speed of light.

2.6 Isotropic Sources

In dealing with both antenna theory and radio-wave propagation, a fictitious device called an isotropic source (or isotropic radiator) is sometimes used for simpler mathematics and for the sake of comparison to real antennas. You will see the isotropic model mentioned at several places in this book. An isotropic source is a tiny source that radiates energy equally well in all directions. The radiation pattern is a perfect sphere with the isotropic antenna at the center. Because such an idealized source generates uniform output in all directions, and its geometry is easily determined mathematically, the signal intensity can be calculated easily from basic geometric principles.

While an isotropic source is simple to describe compared to more complicated radiating antennas, no real electromagnetic source is fully isotropic. The isotropic source is an idealized mathematical radiator that helps us understand other radiators and is used as a reference to compare the strength of different radiating sources.

For the isotropic radiator, the power per unit area on the surface of a spherical wavefront or power density is

$$S_{\text{av}} = \frac{P_t}{4\pi r^2} \quad (2.6)$$

where
P_t
is the total power radiated by the source and *r* is the radius of the sphere in meters (i.e., distance from radiator to point in question). This is the total power radiated by the source divided by the area of a sphere of radius *r*.

As a transmitted wave propagates away from an isotropic source, the radiated sphere of energy gets ever larger. At a great distance from the center, a small slice of the advancing wavefront is essentially a flat plane, as shown in Fig. 2.5. This situation is analogous to the apparent flatness of the prairie in the midwestern United States, even though the surface of the earth is really spherical. If you could see the electric and magnetic field vectors,

they would appear to be at right angles to each other in the flat-plane wavefront.

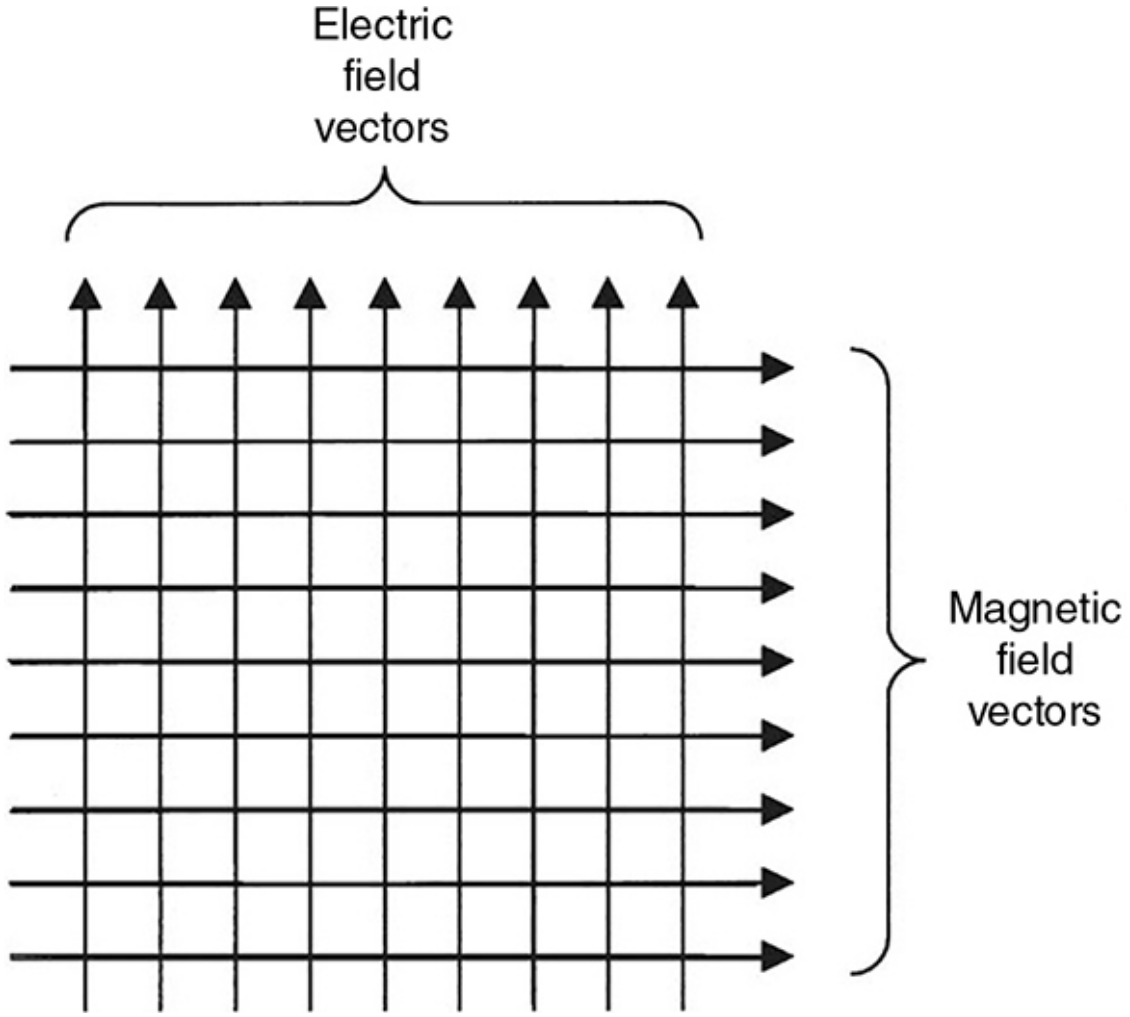


FIGURE 2.5 Far from the source, a spherical wavefront can be treated as a plane wave with E- and H-fields at right angles to each other.

Example 2.1 Calculate the power density at a distance of 1 km (1000 m) from a 1000 W isotropic source.

Solution

$$\begin{aligned} S &= \frac{P}{4\pi r^2} \\ &= \frac{(1000\ \mathrm{W})}{4\pi(1000\ \mathrm{m})^2} \\ &= 7.95 \times 10^{-5}\ \mathrm{W/m^2} \end{aligned}$$

2.7 Receiving Antennas

For a receiving antenna, the ability to collect power from the EM wave and deliver it to the load is measured by the effective area. When the isotropic radiator is used as a receiver, its effective area is $A_e = \lambda^2/4\pi$. For other antennas, this value for the effective area is increased (or decreased) by the gain of the antenna (see Sec. 5.11). The power delivered by a receiving antenna to matched load is the incident power density times the effective area of the receiving antenna, or

$$P_L = S_{\text{inc}} A_e \quad (2.7)$$

where
S_{inc}
is the power density incident on the receiving antenna in
$\mathrm{W/m^2}$
.

By combining Eqs. (2.6) and (2.7), with one isotropic radiator transmitting total power
P_t
to an isotropic receiver a distance *r* away, the power delivered to the receiving antenna's load is

$$P_L = \frac{P_t \lambda^2}{(4\pi r)^2} \quad (2.8)$$

where
$\lambda = c/f$
is the wavelength of the signal.

2.8 Free Space Path Loss

Armed with formulas for the transmitting and receiving properties of an isotropic antenna, we can derive an expression for ordinary free-space path losses (FPL) between an isotropic transmitter antenna and an isotropic receiver antenna, expressed in dB as

$$L(\text{dB}) = -20\log_{10}\left(\frac{\lambda}{4\pi r}\right) \quad (2.9)$$

Alternatively, the free space path loss can be computed using

$$L(\text{dB}) = 20\log_{10}\ r + 20\log_{10}\ f + k \quad (2.10)$$

where *r* is the path length, *f* is the frequency in megahertz (MHz), and *k* is a constant that depends on the units of *r* as follows:

$$\begin{aligned} k &= 32.4 \text{ if } r \text{ in kilometers} \\ &= 36.58 \text{ if } r \text{ is in statute miles} \\ &= 37.80 \text{ if } r \text{ is in nautical miles} \\ &= -37.87 \text{ if } r \text{ is in feet} \\ &= -27.55 \text{ if } r \text{ is in meters} \end{aligned}$$

In Sec. 5.12, we reconsider the free space path loss and signal strength at the receiver for a link between two antennas that may not be isotropic radiators.

2.9 Polarization

The polarization of an EM wave is, by definition, the direction or orientation of the electric field as the wave propagates. Often the electric field oscillates along a single vector direction and the

polarization of the radio wave is linear. To specify the polarization of a radio wave, we give the orientation or angle of the electric field vector with respect to some agreed-upon reference.

In some cases, the reference is the earth's surface, although that may well be meaningless for antennas in outer space. In the case of small mobile antennas, the reference may be the closest large body of conducting material, such as the roof of a car. Figure 2.6*a* is an example of vertical polarization because the E-field is vertical with respect to the earth, the assumed reference plane. If the fields are interchanged (as in Fig. 2.6*b*), the EM wave is said to be horizontally polarized.

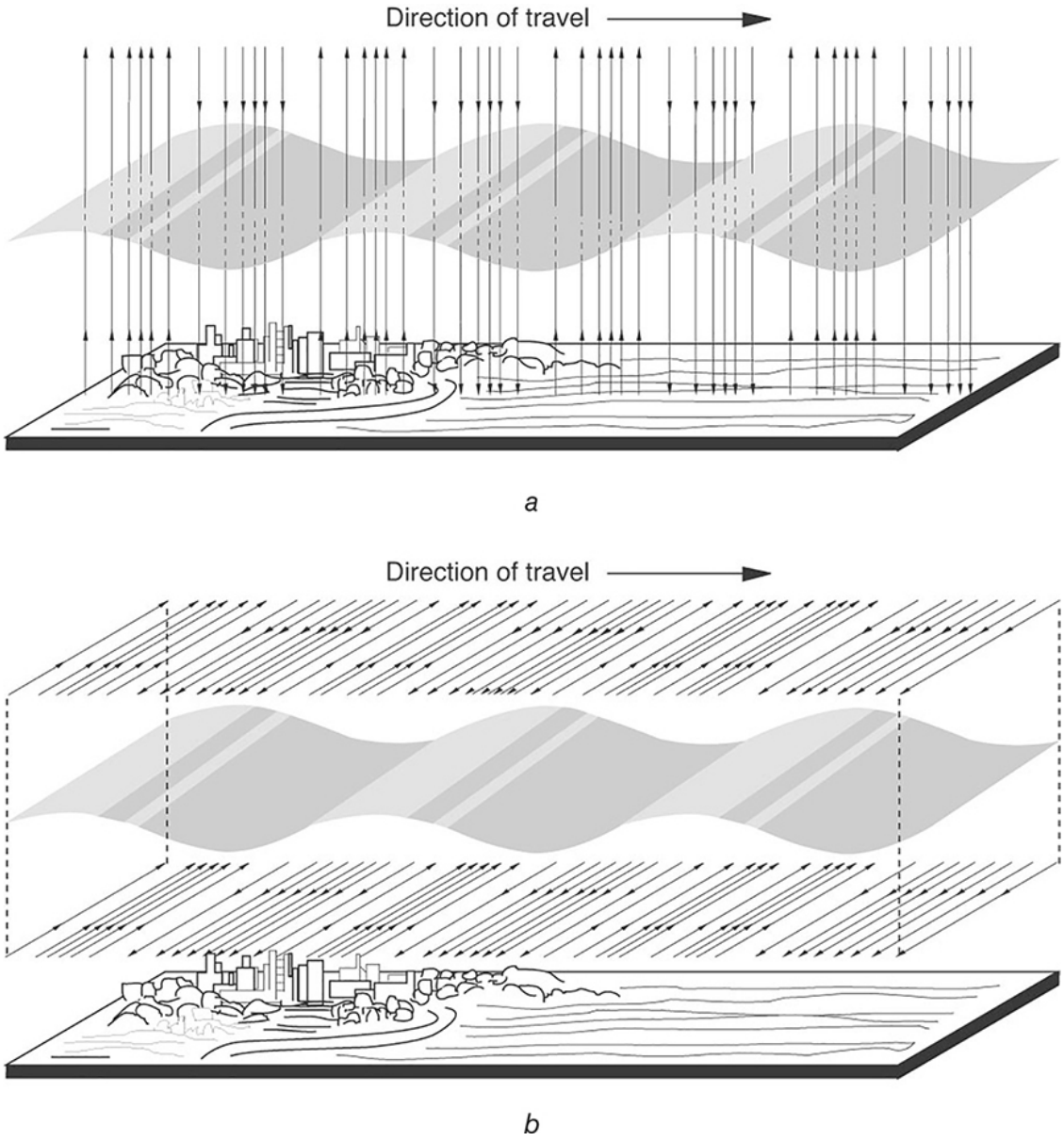


FIGURE 2.6 Wave polarization is determined by the direction of the electric field lines of force. (*a*) Vertically polarized electromagnetic wave. (*b*) Horizontally polarized wave.

If a transmitting antenna radiates a wave with a given polarization, for best performance, the receiving antenna should be aligned with the polarization of the EM wave it receives. In many cases, the polarization at the receiver is the same as it was when radiated by the transmitter, but structures and atmospheric layers

can in some cases change the polarization state of the wave, and adjustments can be made to maximize alignment between the receive antenna and the polarization of the incident wave.

For more complex antennas, the polarization may be circular or elliptical. For these polarization states, the direction of movement of the electric field vector around a circle or ellipse with respect to the direction of propagation of the wave has a sense, which is usually specified as right-handed or left-handed, with the thumb in the direction of wave propagation and the fingers curled in the direction of the electric field vector's rotation.

2.10 Band Designations

The properties of radio waves depend strongly on frequency. Low-frequency radio waves with long wavelengths penetrate buildings and objects and diffract readily around obstacles. High-frequency waves experience more blockage and propagation tends to be more line of sight. Historically, lower frequency radio waves were more readily generated and received, so the general progression over time has been from lower frequency to higher frequency applications.

For convenience in referring to radio communication services at different frequencies, band designations are used. Extremely low frequency (ELF), super low frequency (SLF), ultra low frequency (ULF), and very low frequency (VLF) are used for specialized applications that require deep penetration into earth or ocean and represent frequencies below 30 kHz.

Above a few tens of kHz, the number of radio services explodes into the hundreds, if not thousands. The low frequency (LF) band (30–300 kHz) and medium wave (MW) or medium frequency (MF) band (300 kHz–3 MHz) are used for a variety of services including amplitude modulation (AM) broadcast radio and amateur radio.

The high frequency (HF, 3–30 MHz) band is unique because of ionospheric skip propagation which permits long-distance communication, sometimes world-wide, without the aid of

repeaters. Services include international short-wave radio broadcast, over the horizon radar systems, and popular amateur radio bands for distant (DX) contacts.

Very high frequency (VHF, 30–300 MHz) and ultra high frequency (UHF, 300 MHz–1 GHz) bands are used for hundreds of shorter-range line-of-sight services including digital TV broadcast, citizens band radio, two-way mobile business service, and emergency communication. The former VHF analog TV broadcast channels are now reallocated to other communication services. In these bands, repeater transmission is often used to extend coverage range.

The microwave bands cover 1 GHz to 100 GHz and are uses for a variety of services including satellite up and down links (e.g., global positioning satellites, satellite TV, and broadcast radio), long-haul trunkline communications (though largely replaced now by fiber optics), commercial mobile wireless services, wireless internet (WiFi), and short-range wireless device connections like Bluetooth. The average person directly uses dozens of services operating in these large swaths of radio spectrum every day and indirectly hundreds of others.

In the early 1900s, the "high frequency" and "ultra high frequency" bands were at the state of the art of what could be handled with the radio equipment of the day. As technology progressed, communication in the GHz frequency range and above became possible. Since there are only so many comparative prefixes beyond "ultra," for frequencies above 1 GHz a system of letter band designations was created. This has led to the historical quirk that the "very high" and "ultra high" frequency bands are actually relatively low in frequency compared to the bands used by modern communication services like cellular networks and satellite voice, data, and video links.

Letter band designations begin with L band (1–2 GHz), S band (2–4 GHz), C band (4–8 GHz), and X band (8–12 GHz). These are followed by Ku, K, and Ka bands (12–18 GHz, 18–27 GHz, and 27–40 GHz, respectively). At the high end of the microwave range are V band (40–75 GHz) and W band (75–100 GHz).

For broad references to larger frequency ranges, more general terms are used. Radio waves with frequencies from around 1 GHz to 100 GHz are referred to as microwaves. Antenna sizes scale very roughly with wavelength. In contrast to the tall antenna towers needed for HF bands, the antennas needed for radio communications at frequencies in the GHz range seemed rather tiny, hence the name microwaves. Frequencies in the range of 30–300 GHz with wavelengths on the order of a few millimeters or less are millimeter band or mm-wave frequencies.

Above the mm-wave range is the THz band. Waves with frequencies from around 1 to 10 THz are used for security scanning, chemical sensing, and other applications. Due to the sophisticated and expensive electronics needed to generate and process signals at these frequencies, the THz band represents territory that is only beginning to be exploited for practical applications. Waves at THz frequencies and higher are similar in their behavior to visible light. The THz band overlaps with the infrared range (up to around 400 THz) that sits just below visible light in frequency.

2.11 Radio Spectrum Access, Regulation, and Coordination

In one sense, radio spectrum is nearly unlimited. The widths of named frequency bands tend to increase with frequency and the band designations have a sort of logarithmic nature. The HF band offers a few MHz of bandwidth, whereas X band offers many GHz.

The feasible data rate for a given communication signal is roughly proportional to the bandwidth. A hypothetical radio link occupying frequencies from ELF through UHF band (this would be impossible in practice due to the complex patchwork of national and international regulations governing radio services) could transport roughly three gigabits of data per second. Just one of the letter-designated bands, Ka band alone, offers this same amount of potential bandwidth. A tiny slice of V or W band offers more

potential data bandwidth than all the frequencies in the ELF, SLF, ULF, VLF, HF, VHF, and UHF bands combined. Thus, there is a huge demand for new technologies and services in the high microwave and mm-wave ranges as they offer so much potential bandwidth.

The catch is that radio waves at frequencies in the GHz range and above are easily blocked. Services operating at high frequencies tend to be relatively short range or point to point. This is exploited by cellular networks, as frequencies can be reused from one cell to another without causing interference due to the rapid decay with distance of high frequency radio waves over terrain and structures. New communication technologies operating in the microwave and mm-wave bands will offer enormous data handling capability, but they operate over short distances (e.g., in picocells or femtocells) with mainly line of sight propagation.

The long-range transmission capability of radio waves with frequency below 1 GHz means that the lower frequency bands will almost certainly remain extremely valuable. It's worth mentioning here that antennas operating at frequencies from low MHz to a few GHz are the focus of this book.

With these ideas in mind, we can think of radio spectrum as a consumable resource similar to oil and minerals in the ground. Only so many radio services can operate in a given geographic area and bandwidth. If nearby transmitters operate at or near the same frequency without any coordination, interference occurs and users experience degradation in quality of services.

To avoid interference, radio services need to be regulated. Once a frequency band has been approved by a regulatory body for an application, hardware is developed, built, and purchased and people begin to use the band. This represents significant economic value and the approved band usage cannot be easily changed. Once a frequency band is allocated, it's allocated and that is very hard to undo.

There is only so much bandwidth available in desirable frequency ranges that combine reasonable bandwidth with good geographic coverage. Given that radio spectrum is a limited resource, regulators seek to ensure that various radio services can effectively

and safely coexist. Companies and government institutions support research and development aimed at creating new technologies for avoiding interference, sharing bandwidth, and coordinating radio services. Legal frameworks seek to enable reasonable access to radio spectrum for all people, regardless of economic or social status. Spectrum access, regulation, coordination are exceedingly important to society and will continue to be so as new radio services are developed.

CHAPTER 3
Radio Wave Propagation

To choose the right antenna for our purposes, we first need an understanding of how radio waves get from one place to another. Intuitively, radio signal propagation seems similar to light propagation; after all, light and radio signals are both electromagnetic waves. But the propagation of radio signals is not the simple matter that it seems at first glance. The intensity and other properties of an electromagnetic wave are influenced as the wave travels through the atmosphere and around mountains, hills, and other terrain features. Simple predictions based on the spreading of energy as a wave propagates can fail drastically at radio frequencies due to diffraction, interaction with the atmosphere, and many other practical effects.

In the microwave region of the spectrum, electromagnetic waves are influenced by atmospheric pressure and water vapor content of the air through which the wave propagates. Near-earth microwave propagation differs from that of the very high frequency (VHF) and high frequency (HF) bands. Perhaps the most amazing difference is found in the medium-frequency (MF) and high-frequency (HF) portions of the electromagnetic spectrum, where solar ionization of the upper reaches of the atmosphere provides a refracting layer that acts much like a mirror at these frequencies and supports long-distance "skip" communications essential to such uses of radio as international broadcasting and amateur DXing ("DX" is radio shorthand for "long distance").

In order to help develop an understanding of what an antenna is expected to accomplish and which design parameters are important for effective radio communications, this chapter examines electromagnetic wave propagation phenomena at radio frequencies.

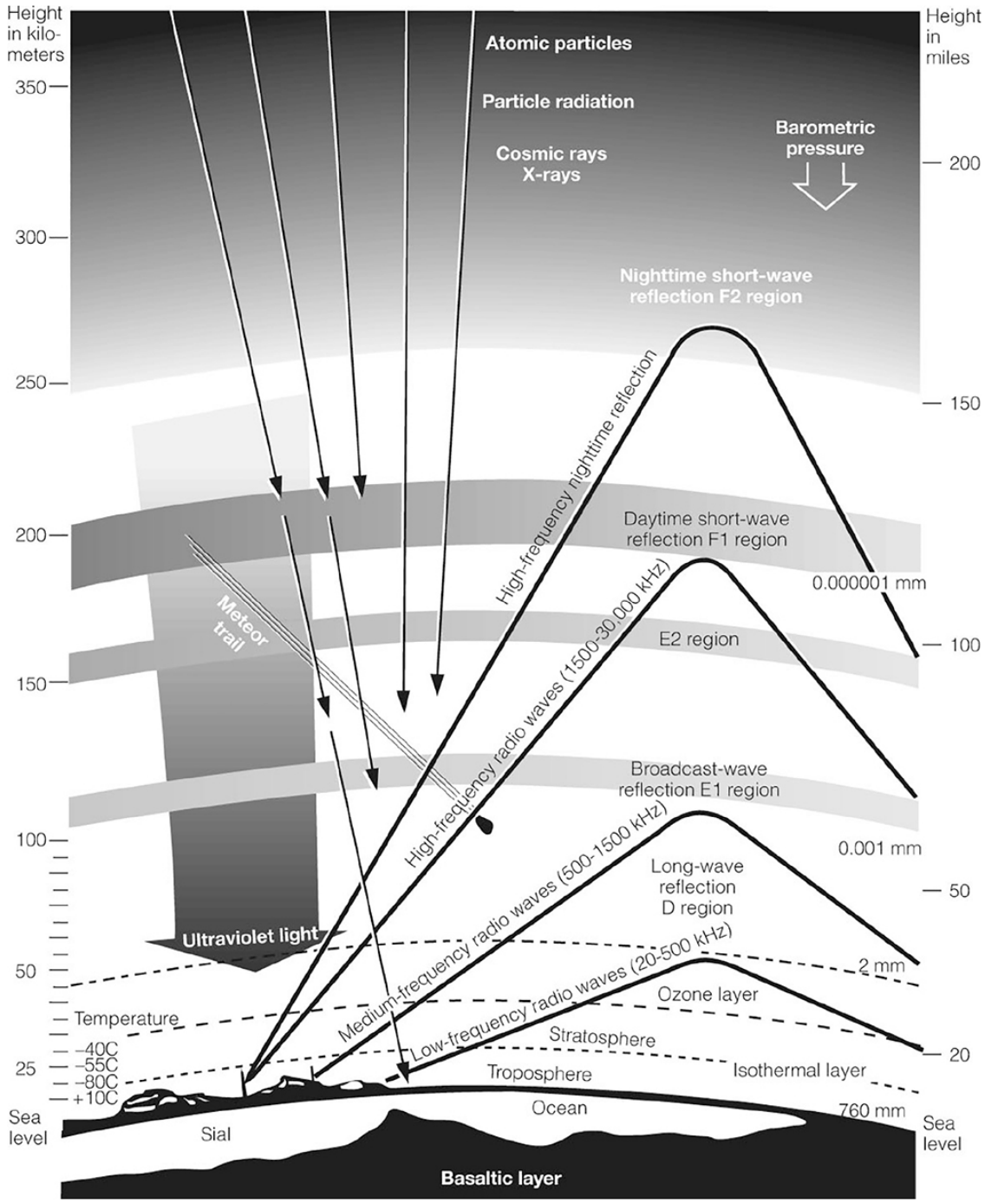
3.1 The Earth's Atmosphere

As Marconi and his contemporaries would have immediately realized if they could have witnessed today's data communications between NASA's ground stations and their interplanetary probes in the near vacuum of outer space, electromagnetic waves do not need an atmosphere in order to propagate. But when a radio wave does propagate in the earth's atmosphere, it interacts with it. As a result, the radio wave will be subject to three potential effects:

- It will suffer increased attenuation (relative to normal free-space path loss over a comparable distance).
- Its path may be bent or redirected.
- Its polarization may be altered.

All of these effects vary with frequency and from one region of the atmosphere to another.

The atmosphere (Fig. 3.1) consists of three major regions: troposphere, stratosphere, and ionosphere. Distinctions between these regions are a matter of definition based on gas densities and behaviors by groups of scientists. The boundaries between these regions are not very well defined, and they change diurnally (over the course of a day) and seasonally.



F IGURE 3.1 Radio waves in the atmosphere and some external radiation sources.

The troposphere occupies the space closest to the earth's surface, extending upward to an altitude of 6 to 11 km (4 to 7 mi—roughly the cruising altitude of most jet airliners). The temperature

of the air in the troposphere varies with altitude, becoming considerably lower than temperatures at the earth's surface as altitude increases. For example, a $+10^{\circ}\mathrm{C}$ surface temperature could exist simultaneously with $-55^{\circ}\mathrm{C}$ at the upper edges of the troposphere.

The stratosphere begins at the upper boundary of the troposphere (6 to 11 km) and continues upward to the beginning of the ionosphere ($\approx 50\ \mathrm{km}$). Because the temperature in this region is relatively constant across a wide range of altitudes, the stratosphere is called an isothermal region.

The ionosphere occupies the region between altitudes of about 50 km (31 mi) and 300 km (186 mi). It is a region of very low gas density because it is the portion of our atmosphere where the earth's gravity has the least "pull" on individual molecules. Beyond the ionosphere, earth's gravity is so weak that air molecules that stray there ultimately drift off into space, never to be seen again.

Of particular interest to our study of high-frequency propagation is the fact that radiation from the sun has sufficient energy to strip electrons away from the gas molecules of the ionosphere. This is caused by a mixture of cosmic rays, electromagnetic radiation including ultraviolet light from the sun, and atomic particle radiation from space (most of these from the sun also) impinging on the outermost edges of our atmosphere. Free electrons are negatively charged, while the oxygen (O_2) and nitrogen (N_2) molecules that lost the electrons become positive ions. The density of the air is low at those altitudes, so each free electron can travel a long distance before bumping into a positive ion and neutralizing each other's electrical charge by recombining.

Ionization does not occur at lower altitudes—that is, in the troposphere and stratosphere—partly because much of the incoming radiation is blocked before it can reach that far and partly because the air density is so much greater at lower altitudes that the positive and negative ions are more numerous and closer together and recombination occurs rapidly.

Because a large percentage of the total radiation "raining" on our atmosphere comes from the sun, ionization levels in the ionosphere vary with the time of day, with the season (since the earth is farther from the sun during the northern hemisphere's summer), and with short-term and long-term solar radiation levels.

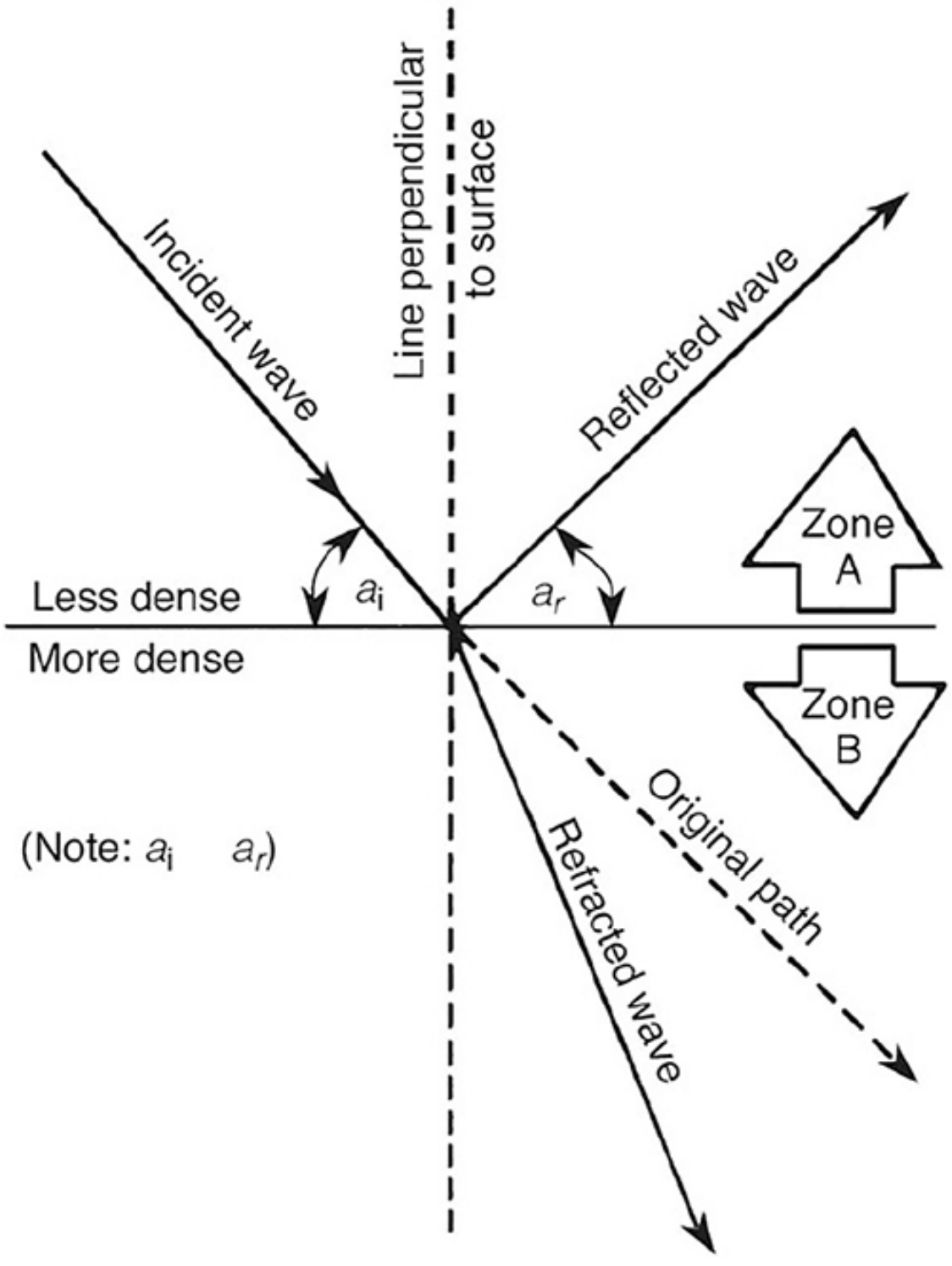
Ionization of the upper level of the earth's atmosphere by radiation from space causes the ionosphere to have electrical characteristics not shared by the lower levels. As electromagnetic waves propagate through the ionosphere, the fields associated with the wave interact with charged particles. These interactions between radio waves and the ionosphere cause waves to reflect and can change the phase and polarization characteristics of the waves. Not surprisingly, the effects of the interactions are frequency-dependent.

Another effect of radiation from space is to alter the characteristics of bands of magnetism (so-called magnetic belts) encircling our globe. When these bands are disturbed by excessively high cosmic radiation, they alter the magnetic fields surrounding the earth, causing disruption of normal EM propagation. As with ionization effects, some frequencies are affected more than others.

The ionization of our upper atmosphere is a major factor in the extreme variability of medium- and high-frequency radio-wave propagation. Ionospheric propagation effects can be almost magical for those of us who have been mesmerized by the unpredictability of long-distance terrestrial radio communications.

3.2 Reflection, Refraction, Diffraction, and Dispersion

The path and polarization of radio waves can be modified by the materials in which the waves propagate through reflection, refraction, diffraction, and dispersion. Figures 3.2 and 3.3 illustrate these wave behavior phenomena. All four effects play important roles in radio propagation.



F IGURE 3.2 Reflection and refraction phenomena.

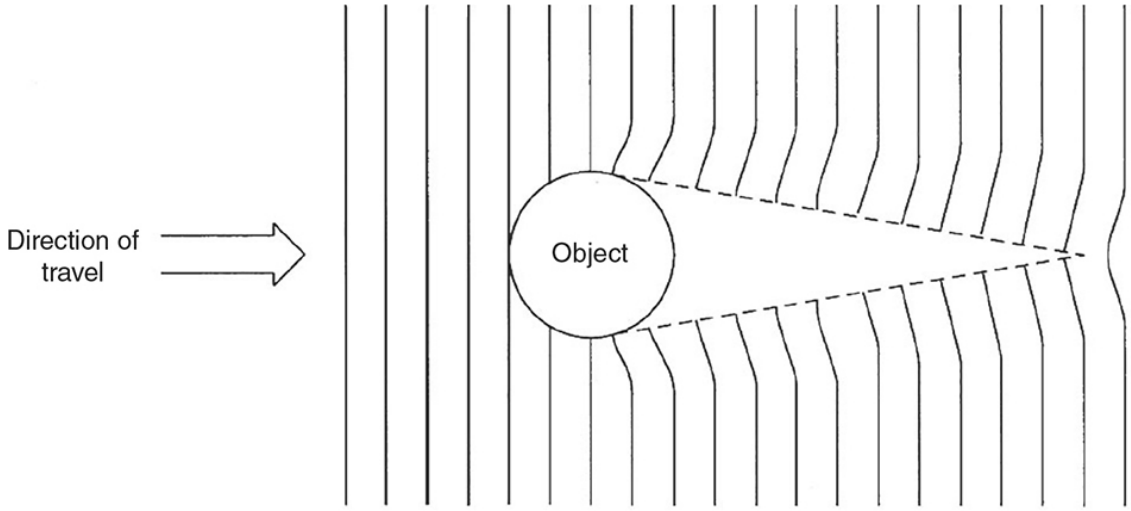


FIGURE 3.3 Diffraction phenomena.

Reflection and refraction are shown in Fig. 3.2. Reflection occurs when a wave strikes a denser reflective medium, such as when a light wave in air strikes a glass mirror. The incident wave (shown as a single ray) strikes the interface between less dense and more dense media at a certain angle of incidence (a_i), and is reflected at exactly the same angle, called the angle of reflection (a_r). Because these angles are equal, a light beam or radio signal undergoing pure reflection can often be traced back to its origin.

If the incoming light wave arrives at certain other angles of incidence at the boundary between media of two different densities, refraction is the result. The amount and direction of the change are determined by the ratio of the densities between the two media. If the material or medium in Zone B is much different from the material in Zone A, the bending is pronounced. In radio systems, the two media might be different layers of air with different densities. It is possible for both reflection and refraction to occur in the same system. Indeed, more than one form of refraction might be present. More on these topics later.

Diffraction is shown in Fig. 3.3. An advancing wavefront encounters an opaque object (e.g., a steel building). The shadow zone behind the building is not simply perpendicular to the wave but takes on a cone shape as waves bend around the object. The "umbra region" (or diffraction zone) between the shadow zone ("cone of silence") and the direct propagation zone is a region of weak (but not zero) signal strength. In practical situations, signal strength in the cone of silence rarely reaches zero. A certain amount of reflected signal scattered from other surfaces can sometimes fill in the shadow region. The degree of diffraction effect is a function of the object's electromagnetic properties and the wavelength of the signal relative to the dimensions of the object.

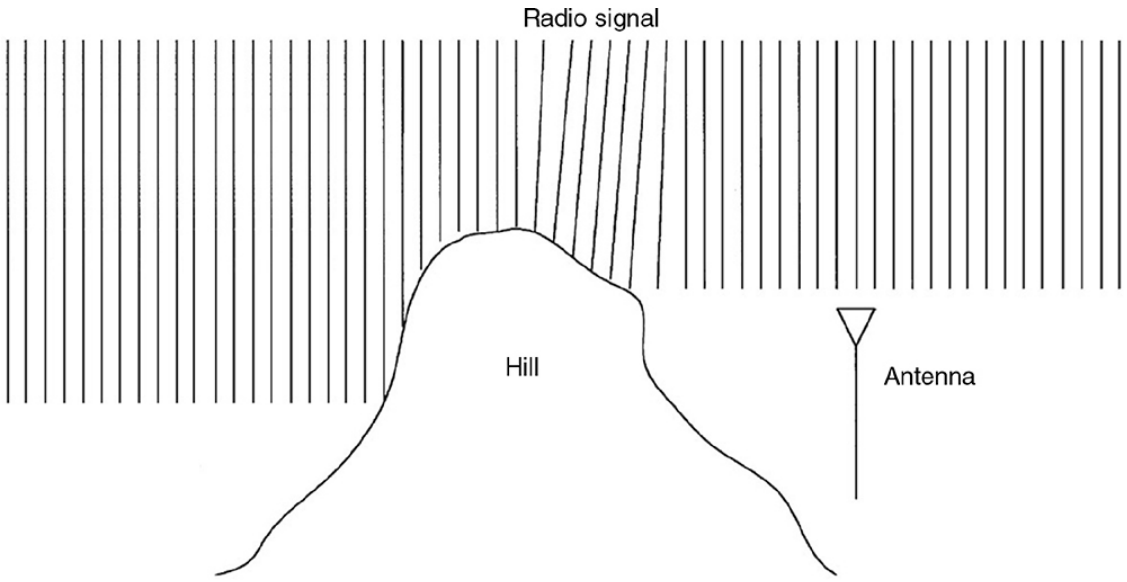
Dispersion can be seen in the spreading of a beam of white light into its separate colors, such as happens with a prism. It is a direct result of the angle of refraction for a medium varying with frequency. We will have occasion to discuss dispersion further when we discuss modes of fading over propagation paths.

3.2.1 Diffraction Phenomena

Electromagnetic waves diffract when they encounter an opaque object. The degree of diffraction and the resulting harm or benefit are frequency related. Above 3 GHz, wavelengths are so small (approximately 10 cm) compared to object sizes that large attenuation of the signal occurs. In addition, beamwidths for high gain antennas (a function of antenna size compared with wavelength) tend to be small enough above 3 GHz that blockage of propagation by obstacles is more significant than at lower frequencies.

Earlier in this chapter, large-scale diffraction around structures (such as buildings) was discussed in conjunction with Fig. 3.3, a top-down view of diffraction effects in the horizontal plane. Diffraction also occurs in the vertical plane. Terrain or man-made objects intervening in the path between UHF microwave stations cause diffraction (Fig. 3.4), along with the concomitant signal attenuation. There is a minimum clearance required to prevent

severe attenuation (more than 20 to 30 dB, say) from diffraction. Calculation of the required clearance comes from Huygens–Fresnel wave theory.



F IGURE 3.4 Terrain masking of VHF and higher-frequency signals.

As shown in Fig. 3.5, a wave source A, which might be a transmitter antenna, radiates a wavefront to a destination C (receiver antenna). At any point along path *A-C*, you can look at the wavefront as a partial spherical surface ($B_1 - B_2$) on which all wave rays have the same phase. This plane can be called an isophase plane.

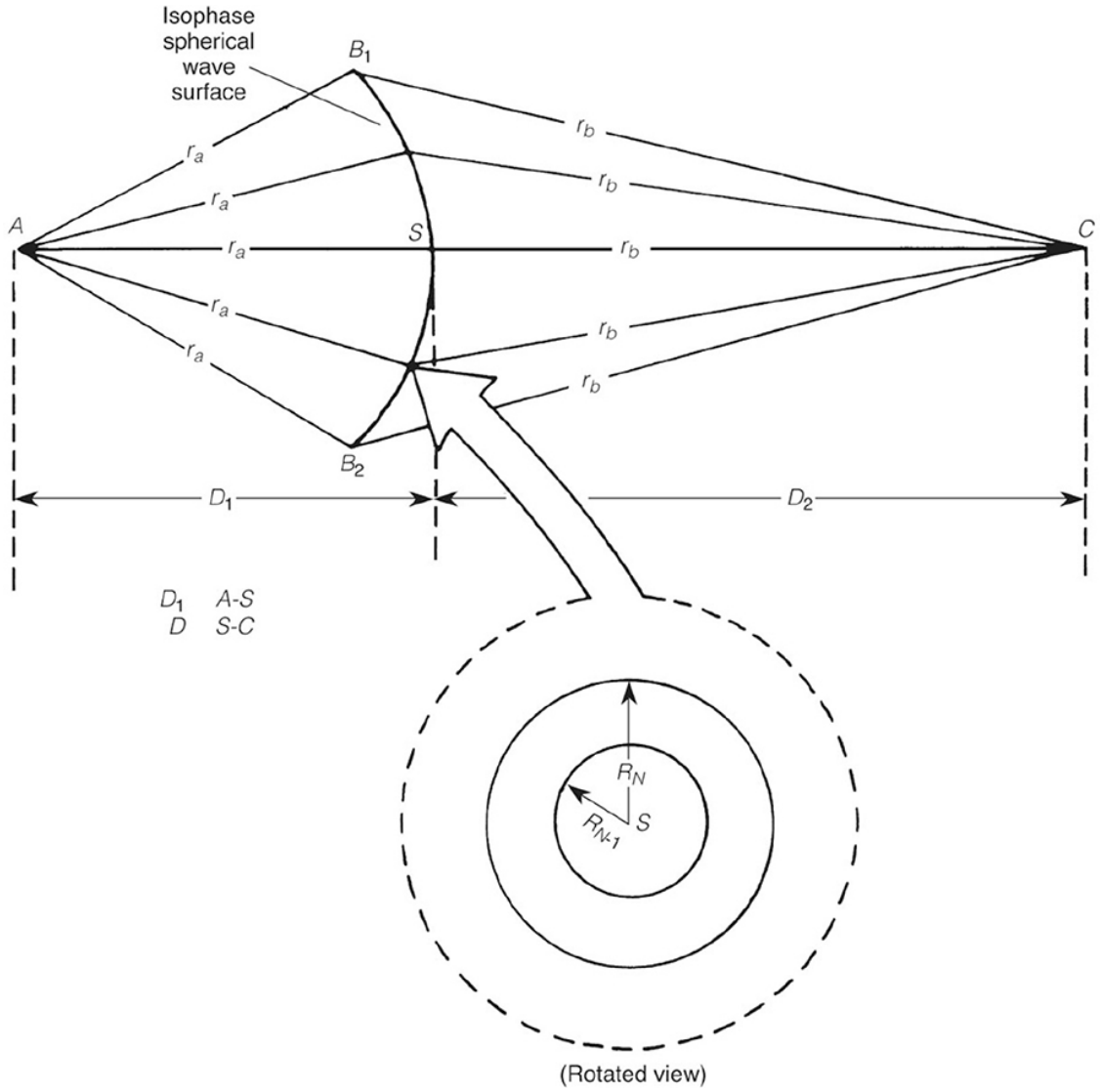


FIGURE 3.5 Fresnel zone geometry.

Using ray tracing, we see rays r_a incoming to the surface [$B_1 - B_2$] and rays r_b outgoing from plane [$B_1 - B_2$]. The signal seen at *C* is the algebraic sum of all rays r_b

. The signal pattern will have the form of an optical interference pattern with wave cancellation occurring between
r_b
waves that are a half wavelength apart on [
$B_1 - B_2$
]. The ray impact points on plane [
$B_1 - B_2$
] form radii
R_n
called Fresnel zones. The lengths of the radii are a function of the frequency and the ratio of the distances
D_1
and
D_2
(see Fig. 3.5). The radius of the *n*th Fresnel zone is

$$R_n = M\sqrt{\frac{n}{f}\left(\frac{D_1 D_2}{D_1 + D_2}\right)} \quad (3.1)$$

where *f* is the frequency in GHz,
D_1
is the distance from the source to plane
$B_1 - B_2$
,
D_2
is the distance from the destination to plane
$B_1 - B_2$
, *n* is an integer (1, 2, 3 . . .), and *M* is 17.3 if
R_m
is in meters and
D_1
,
D_2
are in kilometers or 72.1 if
R_m

is in feet and D_1, D_2 are in statute miles. If you first calculate the radius R_1 of the critical first Fresnel zone, you can calculate the *n*th Fresnel zone from

$$R_n = R_1\sqrt{n} \quad (3.2)$$

Example 3.1 Calculate the radius of the first Fresnel zone for a 2.5-GHz signal at a point that is 12 km from the source and 18 km from the destination.

Solution

$$R_1 = M\sqrt{\frac{N}{F}\left(\frac{D_1 D_2}{D_1 + D_2}\right)} = (17.3)\sqrt{\frac{1}{2.5}\left(\frac{12 \times 18}{12 + 18}\right)} = 29.4 \text{ m}$$

For most terrestrial microwave systems, an obstacle clearance of 0.6 R_1 is required to prevent diffraction attenuation under most normal conditions, but there are conditions in which the clearance should be more than one Fresnel zone.

3.3 Propagation Paths

There are four major propagation paths for a radio wave emanating from a terrestrial source: surface wave, space wave, tropospheric, and ionospheric. The ionospheric path is a major factor in medium-wave (MW) and HF propagation. For VHF, UHF, or microwave propagation to/from space-based antennas it serves primarily only as a potential source of loss. At these frequencies, ionization and

recombination phenomena in the ionosphere add to the noise level. In satellite communications, there are additional trans-ionospheric effects.

The space wave and surface wave are both ground waves, but they behave differently enough to warrant separate consideration. The surface wave travels in direct contact with the earth's surface and it suffers severe frequency-dependent attenuation caused by continual contact with and absorption by the earth.

The space wave is also a ground-wave phenomenon, but it is radiated from an antenna a wavelength or more above the surface. No part of the space wave normally travels in contact with the surface. VHF, UHF, and microwave signals are usually space waves. In many cases, there are two components of the space wave: direct and reflected (see Fig. 3.6).

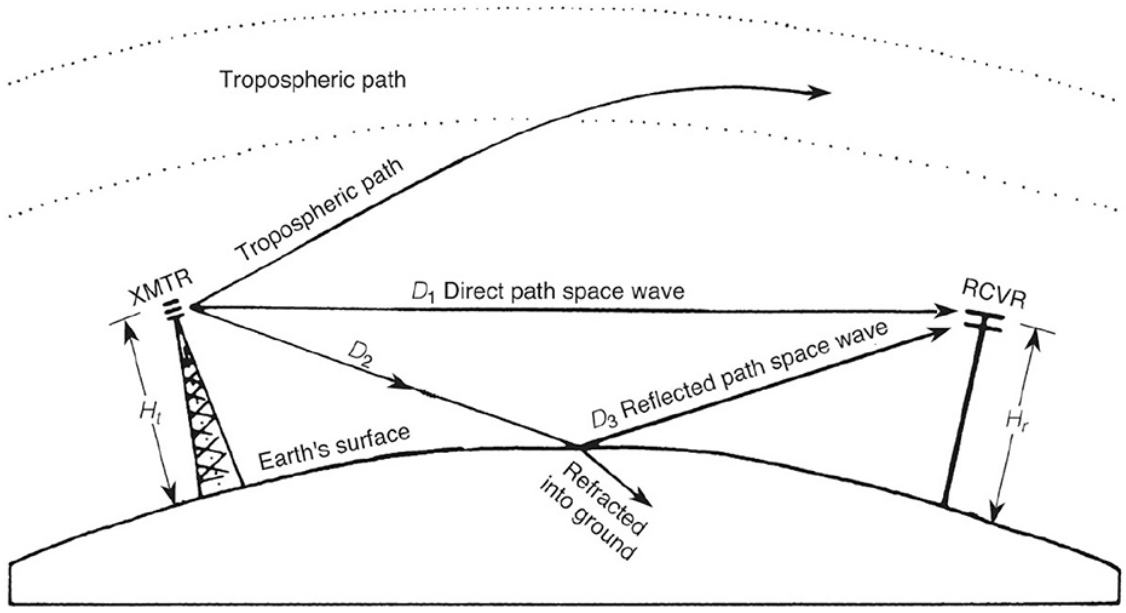


FIGURE 3.6 Space-wave propagation.

The tropospheric wave is sometimes lumped with the direct space wave, but it has unique properties in certain practical situations. The troposphere is the region of the earth's atmosphere between the surface and the stratosphere, or about 4 to 7 mi above the surface. Most forms of ground wave propagate in the

troposphere. Because some propagation phenomena (caused mostly by weather conditions) occur only at higher altitudes, tropospheric propagation should be differentiated from other forms of ground waves.

3.4 Ground-Wave Propagation

The ground wave, naturally enough, travels along the ground, or at least in close proximity to it. There are three basic forms of ground waves: space wave, surface wave, and tropospheric wave. The space wave does not actually touch the ground. As a result, space-wave attenuation as a function of distance in clear weather is about the same as in free space (except above about 10 GHz, where H_2O and O_2 absorption increases dramatically). Above the VHF region, weather conditions add attenuation not found in outer space.

The surface wave is subject to the same attenuation factors as the space wave, but it also suffers ground losses. These losses are caused by ohmic resistive losses in the conductive earth. Simply stated, the signal heats up the ground. AM band broadcast stations utilize ground-mounted vertically polarized antennas to radiate a surface wave that provides good signal strength throughout their local coverage area. At the same time, ground losses actually help to avoid interference with other broadcast stations on the same frequency in communities farther away.

Surface wave attenuation increases rapidly as frequency increases. For both forms of ground waves, reception is affected by the following factors:

- Wavelength
- Height of both the receiving and the transmitting antennas
- Distance between antennas

- Terrain along the transmission path
- Weather along the transmission path
- Ground losses (surface wave only)

Figure 3.7 is a nomograph that can be used to calculate the line-of-sight distances in miles over a spherical earth, given a knowledge of both receiving and transmitting antenna heights. Similarly, Figs. 3.8 and 3.9 show power attenuation with frequency and distance (Fig. 3.8), and power attenuation in terms of field intensity (Fig. 3.9).

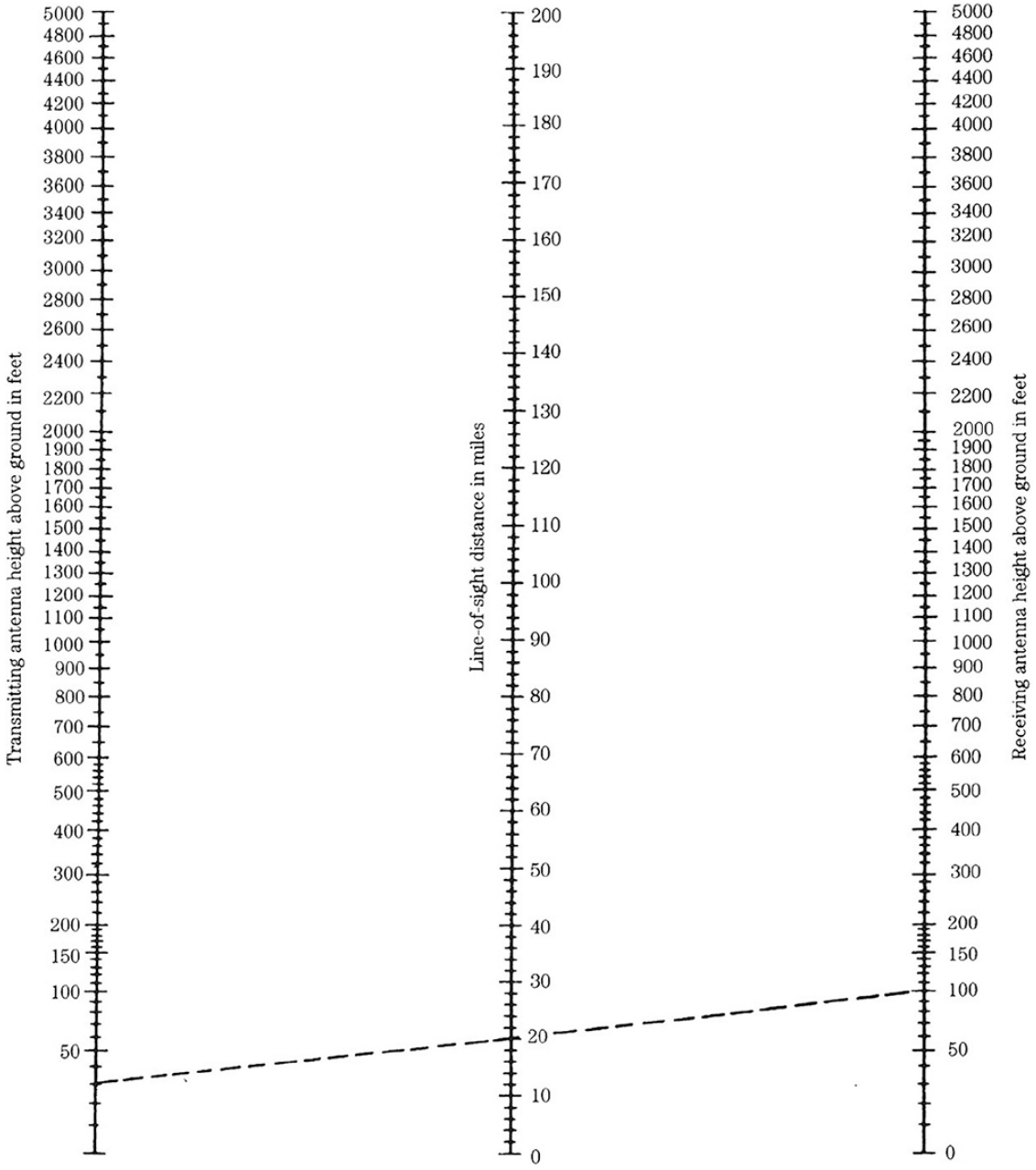


FIGURE 3.7 Nomograph of line-of-sight transmission distance as a function of receiving and transmitting antenna heights.

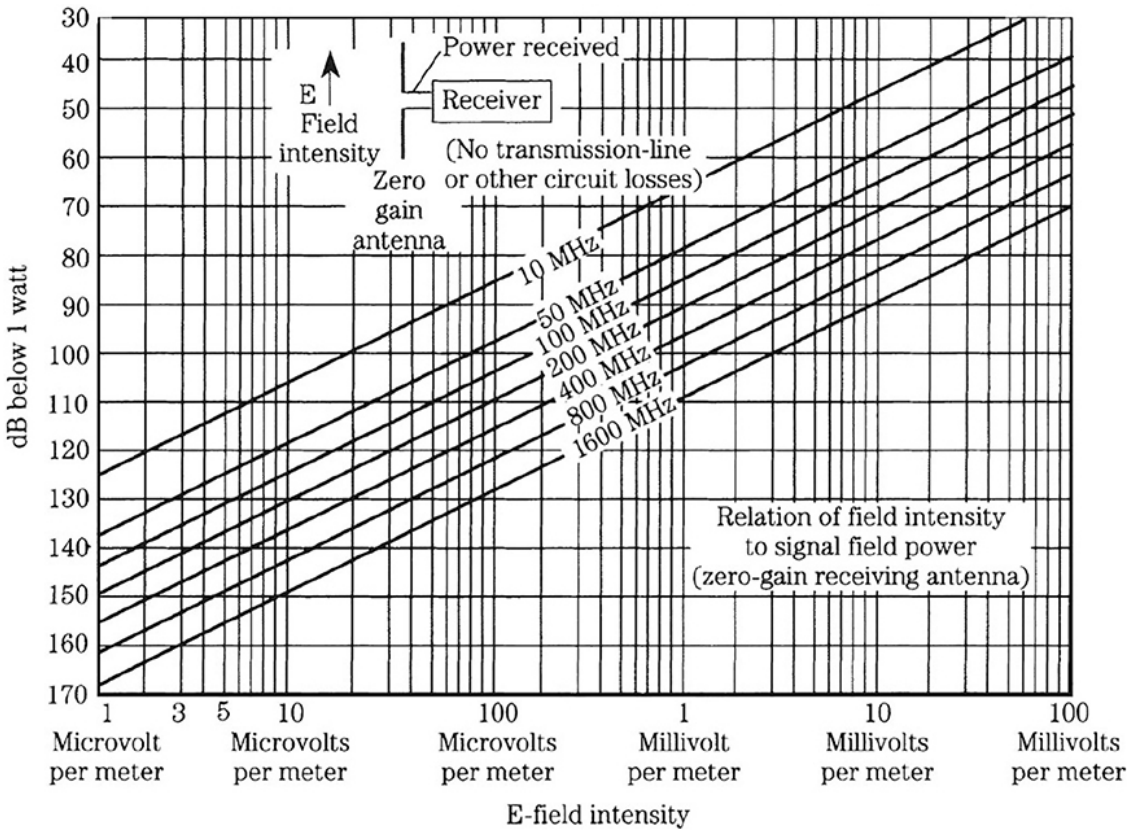


FIGURE 3.9 Relation of field strength to signal field power.

Ground-wave communication also suffers another difficulty, especially at VHF, UHF, and microwave frequencies. The space wave is like a surface wave, but it is radiated from an antenna at least a wavelength above the earth's surface. It consists of two components: the direct and reflected waves (see Fig. 3.6 again). If both of these components arrive at the receiving antenna, they will add algebraically (more accurately, vectorially) to either increase or decrease signal strength. There is always a phase shift between the two components because the two signal paths have different lengths (i.e., D_1 is less than $D_2 + D_3$). In addition, there may possibly be a 180-degree (π radians) phase reversal at the point of reflection (especially if the incident signal is horizontally polarized). If the overall phase shift

experienced by the reflected wave as a result of path length differences and polarization shifts is an odd multiple of 180 degrees at the receiving antenna, as shown in Fig. 3.10, the net signal strength there will be reduced (destructive interference). Conversely, if the direct and reflected signals arrive in phase, the resulting signal strength is increased (constructive interference). The degree of cancellation or enhancement depends on the relative amplitudes of the two signal components as they arrive at the receive antenna.

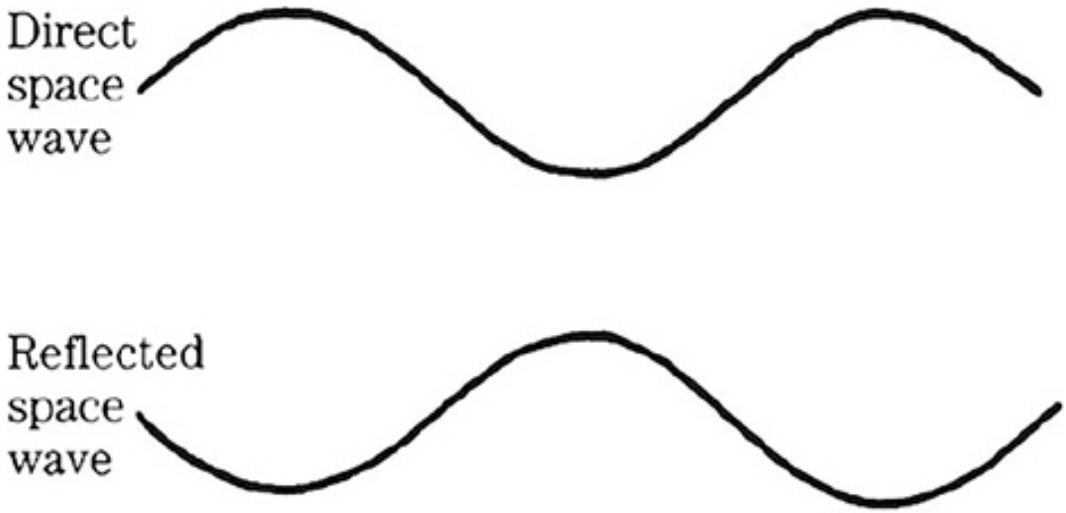


FIGURE 3.10 Out-of-phase direct and reflected waves cancel and lead to destructive interference.

The loss of signal due to destructive interference between the direct and reflected paths can be characterized using the signal loss coefficient

$$\eta = \frac{P_r}{P_f} \quad (3.3)$$

where
P_r
is the signal level at receiver in the presence of the ground reflection component and

P_f
is the free-space signal strength over path
D_1
if no reflection takes place. The signal loss coefficient can be estimated using

$$\eta = 2 \sin\left(\frac{2\pi h_t h_r}{\lambda \mathrm{D}_1} \right) \quad (3.4)$$

where
h_t
and
h_r
are the heights of the transmitting and receiving antennas, respectively.

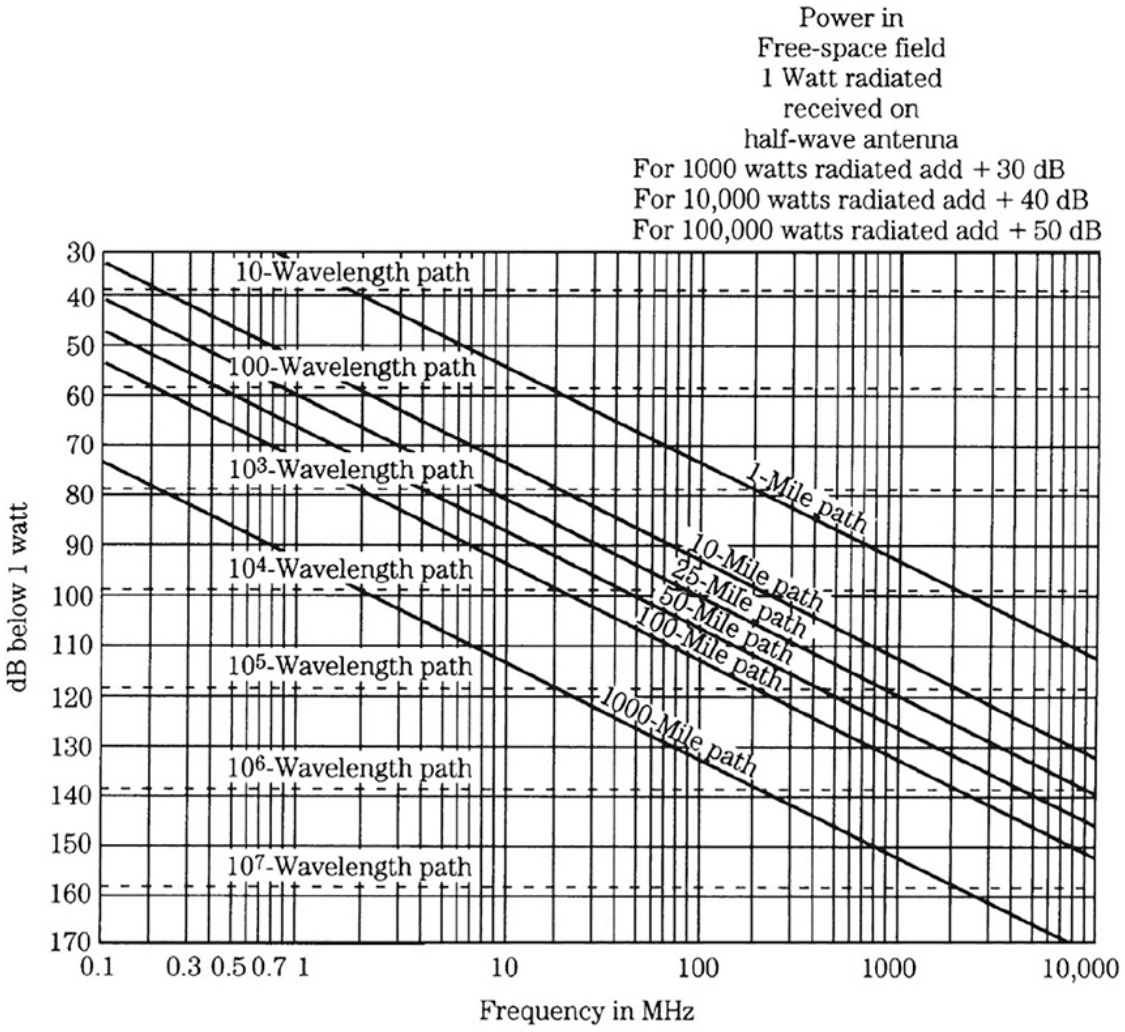


FIGURE 3.8 Power in a free-space field (normalized to 1 W).

The reflected signal contains both amplitude change and phase change. The phase change is typically π radians (180 degrees). The amplitude change is a function of frequency and the nature of the reflecting surface. The reflection coefficient can be expressed as

$$\Gamma = pe^{j\phi} \quad (3.5)$$

where p is the amplitude change, ϕ is the phase change, and $j = \sqrt{-1}$
is the imaginary unit. For smooth, high-reflectivity surfaces and a horizontally polarized microwave signal that has a shallow angle of incidence, the value of the reflection coefficient is close to -1.

The phase change of the reflected signal at the receiving antenna is at least π radians because of the reflection. Added to this change is an additional phase shift that is a function of the difference in path lengths. The phase shift can be expressed in terms of the two antenna heights and path length:

$$s = \pi + \frac{4\pi h_t h_r}{\lambda \mathrm{D}_1} \quad (3.6)$$

Multipath reception problems exist because of interference between the direct and reflected components of the space wave. Multipath caused ghosting of over-the-air analog television signals, although newer broadcast digital television is more immune to multipath effects. Multipath events can be transitory (as when an aircraft flies over the direct transmission path), while others are permanent (as when a large building or hill reflects the signal). In mobile communications, multipath phenomena contribute to reception dead zones and cause picket fencing. A dead zone exists when destructive interference between direct and reflected (or multiply reflected) waves drastically reduces signal strengths. This problem is most often noticed at VHF and above when the vehicle is stopped; the solution is to move the antenna about a half wavelength (which at VHF and UHF may be a matter of a few inches). Picket fencing occurs as a mobile unit moves through successive dead zones and signal enhancement (or normal) zones and with audio transmissions can sometimes sound like a series of short noise bursts.

At VHF, UHF, and microwave frequencies, the space wave is limited to line-of-sight distances. The horizon is roughly the limiting distance for communications, but the radio horizon is actually about 15 percent farther than the optical horizon (Fig. 3.11). This phenomenon is caused by refractive bending of the radio signal in the atmosphere around the curvature of the earth, and it makes the geometry of the situation look as if the earth's radius is 4⁄3 of its actual radius.

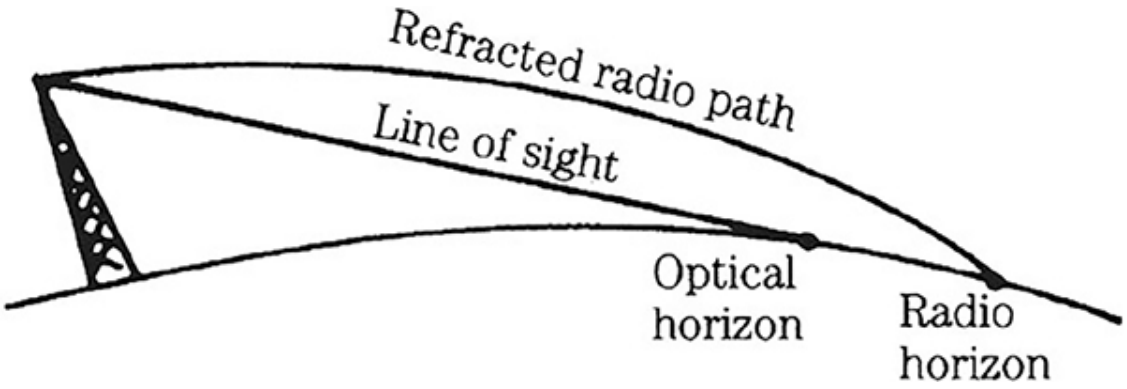


FIGURE 3.11 Radio horizon versus optical horizon.

3.4.1 Ray Path Curvature

Refraction occurs at VHF through microwave frequencies due to water vapor and variations in atmospheric pressure. The differential properties between adjacent zones in the atmosphere is expressed by the index of refraction *n*. The degree of curvature along any given path is characterized by the *K* factor, also called the effective earth radius factor or ray path curvature.

When a signal passes across boundaries between adjacent regions of distinctly different properties (as occurs during temperature inversions, for instance), the index of refraction is the ratio of the signal velocities in the two regions. In a homogeneous region *n* can be expressed as the ratio of the free-space velocity *c* to the actual velocity *v* in the atmosphere:

$$n = \frac{c}{v} \quad (3.7)$$

At the surface, near sea level, under standard temperature and pressure conditions, the value of *n* is approximately 1.0003, and in homogeneous atmospheres it decreases by
4×10^{-8}
per mile of altitude. Since the index of refraction for the atmosphere is close to 1, the UHF and microwave communities use a derivative parameter, *N*, called the refractivity of the atmosphere:

$$N = (n - 1) \times 10^6 \quad (3.8)$$

The refractivity of the atmosphere is altitude-dependent and typically varies from about 280 to 320. In nonhomogeneous atmospheres (the usual case), these parameters will vary approximately linearly for several tenths of a kilometer. Most microwave relay systems can assume an approximately linear reduction of *n* and *N* with increasing altitude, but for airborne radios and radars the linear approximation breaks down. There are two methods for calculating *N*:

$$N = \left(\frac{77.6}{T}\right)\left(P + \frac{481 e_s H_{\text{rel}}}{T}\right) \quad (3.9)$$

and

$$N = \frac{77.6}{T}\left(\frac{3.73 \times 10^5 e_s}{T^2}\right) \quad (3.10)$$

where *P* is the atmospheric pressure in millibars
$(1\ \text{torr} = 1.3332\ \text{mbar})$
, *T* is the temperature in Kelvin,
e_s
is the saturation vapor pressure of atmospheric water in millibars, and
H_{rel}
is the relative humidity expressed as a decimal fraction (rather than a percentage). For higher altitudes above the linear zone, an exponential model can be used:

$$N_a = N_s e^{-\ C_e(h_r - h_t)} \quad (3.11)$$

where
N_1
is the refractivity at 1 km altitude,
h_r
is the height of receive antenna,
h_t

is the height of transmit antenna,
$C_e = L_n(N_s/N_1)$
,
N_a
is the refractivity at altitude, and
N_s
is the refractivity at the earth's surface.

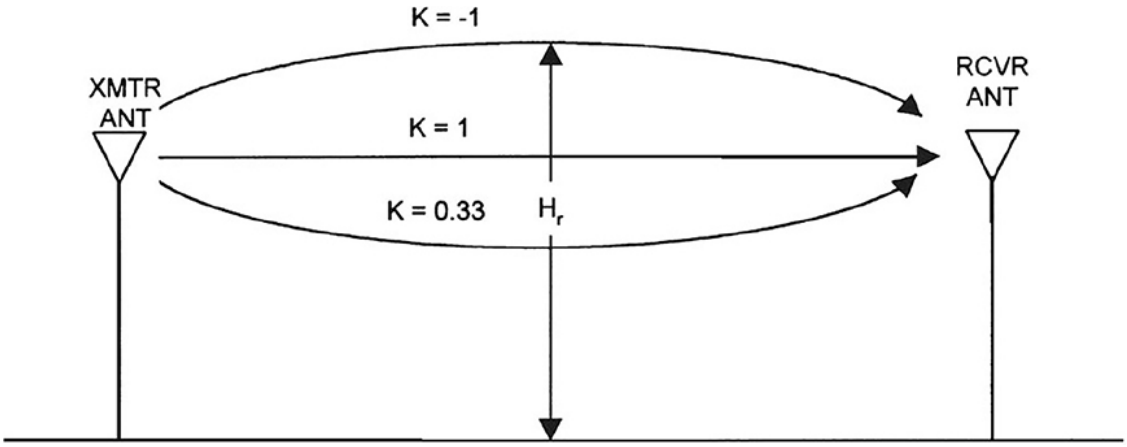
The ray path curvature *K* is defined by the relationship of two hypothetical concentric spheres, both centered on the earth's center. The first sphere is an idealization of the earth's surface, which has a nominal radius

$r_0 = 3440$

nautical miles (nmi) or 6370 km. The second sphere is larger than the first by the curvature of the signal ray path, and has a radius *r*. The value of *K* is then approximately

$$K = \frac{r}{r_0} \quad (3.12)$$

The ray path curvature *K* caused by atmospheric gradients is related to the change in the index of refraction *n* or the refractivity *N* with altitude. The more rapidly the refractivity changes, the greater the ray path curvature. A value of
$K = 1$
indicates a straight path (Fig. 3.12); a value of
$K > 1$
indicates a positively curved path (refraction); and a value of
$K < 1$
indicates a negatively curved path (subrefraction).



F IGURE 3.12 Refraction changes path length.

The value of *K* varies with local weather conditions, so one can expect variation not only between locations but also seasonally. In the Arctic regions, *K* varies approximately over the range 1.2 to 1.34. In the "lower 48" of the United States, *K* varies from 1.25 to 1.9 during the summer months (especially in the south and southeast), and from 1.25 to 1.45 in the winter months. In the absence of additional data, for the near-surface region, the value of *K* is 1.33. For most terrestrial microwave paths, this value ($K = 1.33 \simeq 4/3$) is called the standard refraction.

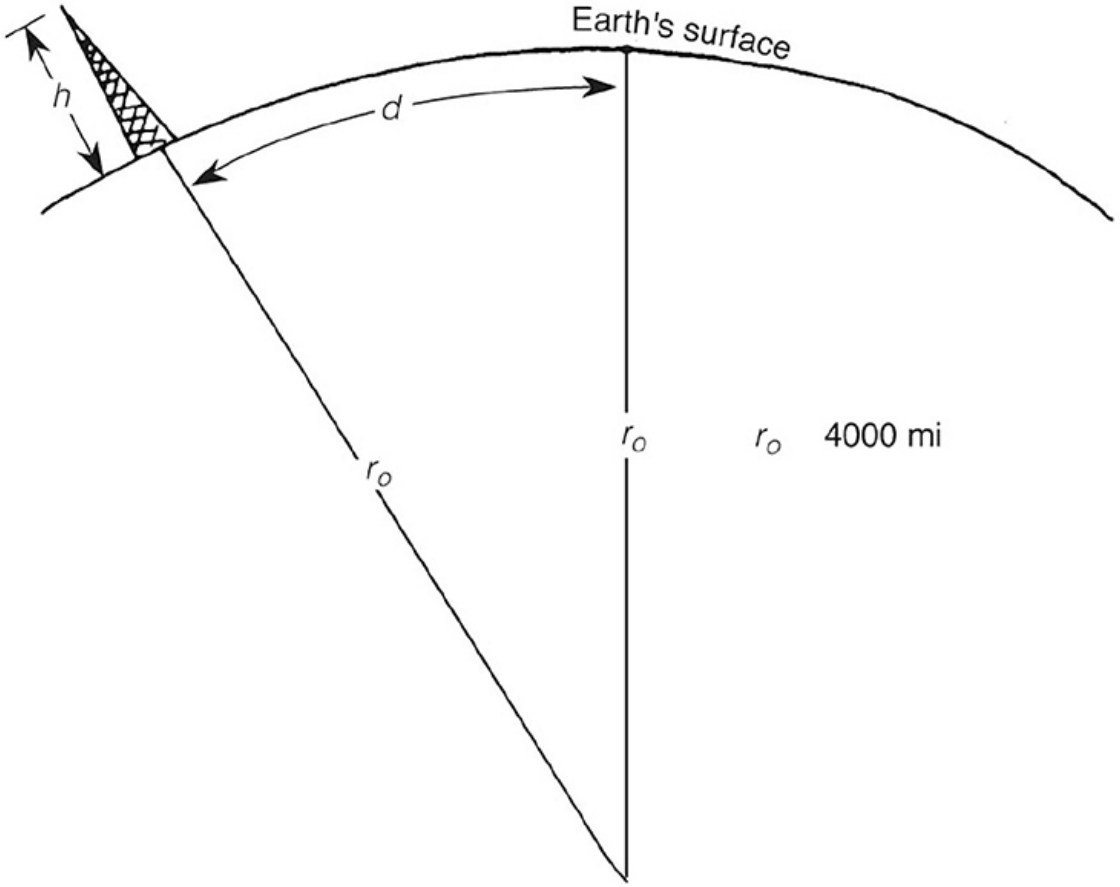


FIGURE 3.13 Geometry for calculating radio line-of-sight distances.

3.4.2 Horizon Distance

For models close to the surface, use the geometry shown in Fig. 3.13, where distance *d* is a curved path along the surface of the earth. Because the earth's radius is much larger than any practical antenna height *h*, the horizon distance *d* can be approximated by

$$d = \sqrt{2rh} \quad (3.13)$$

where *d* is the distance to the radio horizon and *h* is the antenna height. With the earth's radio radius approximated as $4/3$ $(K = 1.33)$
of its physical radius, as discussed previously, and in convenient units, this becomes

$$d(\text{mi}) = 1.414\sqrt{h(\text{ft})} \quad (3.14)$$

Example 3.2 A radio tower has a UHF radio antenna that is mounted 150 ft above the surface of the earth. Calculate the radio horizon for this system.

Solution

$$\begin{aligned} d(\text{mi}) &= 1.414[h(\text{ft})]^{1/2} \\ &= (1.414)(150\ \text{ft})^{1/2} \\ &= 17.3\ \text{mi} \end{aligned}$$

3.4.3 Surface Waves

The surface wave travels in direct contact with the earth's surface. Consequently, it suffers a severe frequency-dependent attenuation from absorption by the ground. The surface wave extends to considerable heights above the ground level, but its intensity drops off rapidly at the upper end. Above the ground, the surface wave is subject to the same atmospheric attenuation factors as the space wave. Additional ground losses are caused by ohmic (resistive) losses in the conductive earth and by the dielectric properties of the earth. In short, the signal heats up the ground.

Horizontally polarized waves are not often used for surface-wave communications because the earth tends to short-circuit the E-field component. (A perfect conducting plane has no voltage between any two points on it, so no tangential E-field can exist on the plane.) For vertically polarized waves, the earth offers electrical resistance to the E-field and returns currents to subsequent waves (Fig. 3.14). The conductivity of the soil determines how much energy is returned. Table 3.1 shows the typical conductivity values for several different forms of soil, as well as saltwater and freshwater.

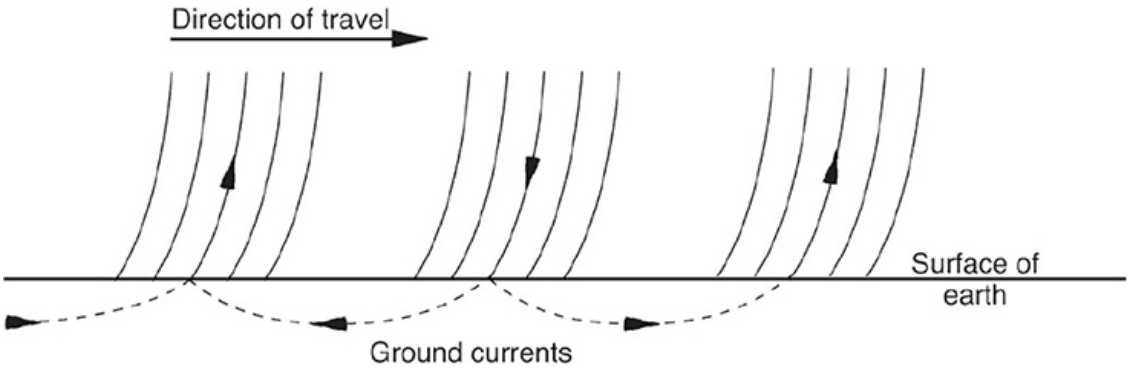


FIGURE 3.14 Distortion of vertically polarized electric field by lossy ground resistance.

TABLE 3.1 Sample Soil Conductivity Values

Type of Soil	Dielectric Constant	Conductivity (Siemens/Meter)	Relative Quality
Saltwater	81	5	Best
Freshwater	80	0.001	Very poor
Pastoral hills	14–20	0.03–0.01	Very good
Marshy, wooded	12	0.0075	Average/poor
Rocky hills	12–14	10	Poor
Sandy	10	0.002	Poor
Cities	3–5	0.001	Very poor

The wavefront of a surface wave is tilted because of the losses in the ground that tend to retard the lower end of the wavefront (also shown in Fig. 3.14). The tilt angle is a function of the frequency, as shown in Table 3.2.

TABLE 3.2 Tilt Angle as a Function of Frequency

Frequency (kHz)	Tilt Angle Ratio	Earth/Seawater (Degrees)
20	207	4.3/0.021
200	104	13.4/0.13
2,000	64	32.3/0.5
20,000	25	35/1.38

3.4.4 Ground-Wave Propagation Frequency Dependence

The surface-wave behavior of a radio signal is influenced by its frequency. In the low-frequency (LF) band (30–300 kHz), ground losses are small for vertically polarized signals, so medium-distance communication up to several hundred miles is possible. In the medium-wave (MW) band (300–3000 kHz, which includes both the AM broadcast band and the 160-m amateur band), distances of a few hundred miles are common. In the high-frequency (HF) band (3000–30,000 kHz), ground losses are typically greater, so the surface-wave distance reduces drastically. It is possible in the upper end of the HF band for surface-wave signals to die out within a few dozen miles. This phenomenon is often seen in the 15- and 10-m amateur radio bands, as well as the 11-m (27-MHz) citizens band. Stations only 20 mi apart may not be able to communicate while—surprisingly— both can talk to a third station across the continent via ionospheric skip.

3.5 Tropospheric Propagation

The troposphere is the portion of the atmosphere between the surface of the earth and the stratosphere or about 4 to 7 mi altitude. Some older texts group tropospheric propagation with ground-wave propagation, but modern practice requires separate treatment. The older grouping overlooks certain tropospheric

propagation phenomena that don’t happen with space or surface waves.

Refraction is the mechanism for most tropospheric propagation phenomena. The dielectric properties of the air, which are set mostly by the moisture content (Fig. 3.15), are a primary factor in tropospheric refraction. Refraction occurs when a wave passes between media of differing densities. The wave path bends in approximate proportion to the difference in densities.

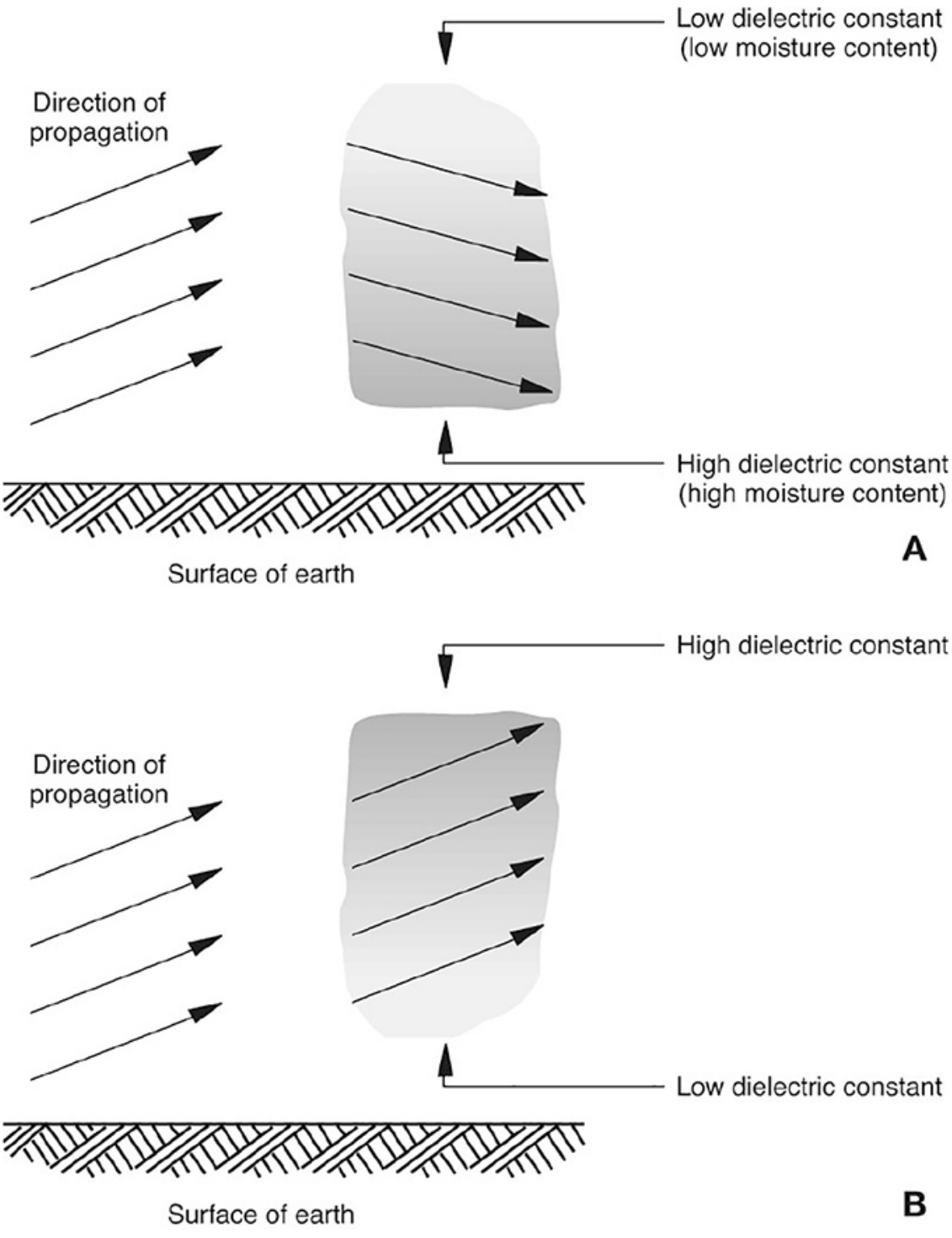


FIGURE 3.15 Refraction in the troposphere.

Two general situations are typically found—especially at UHF and microwave frequencies. First, because air density normally decreases with altitude, the top of a beam of radio waves typically

travels slightly faster than the lower portion of the beam. As a result, those signals refract a small amount. Such propagation provides slightly longer surface distances than are normally expected from calculating the distance to the radio horizon. This phenomenon is called simple refraction, and was discussed in a preceding section.

A special case of refraction called superrefraction occurs in areas of the world where warmed land air flows out over a cooler sea (Fig. 3.16). Typically these areas have deserts adjacent to large bodies of water. The Gulf of Aden, the southern Mediterranean, and the Pacific Ocean off the coast of Baja California are examples. Communications up to 200 mi at VHF/UHF/microwave frequencies are frequently reported in such areas.

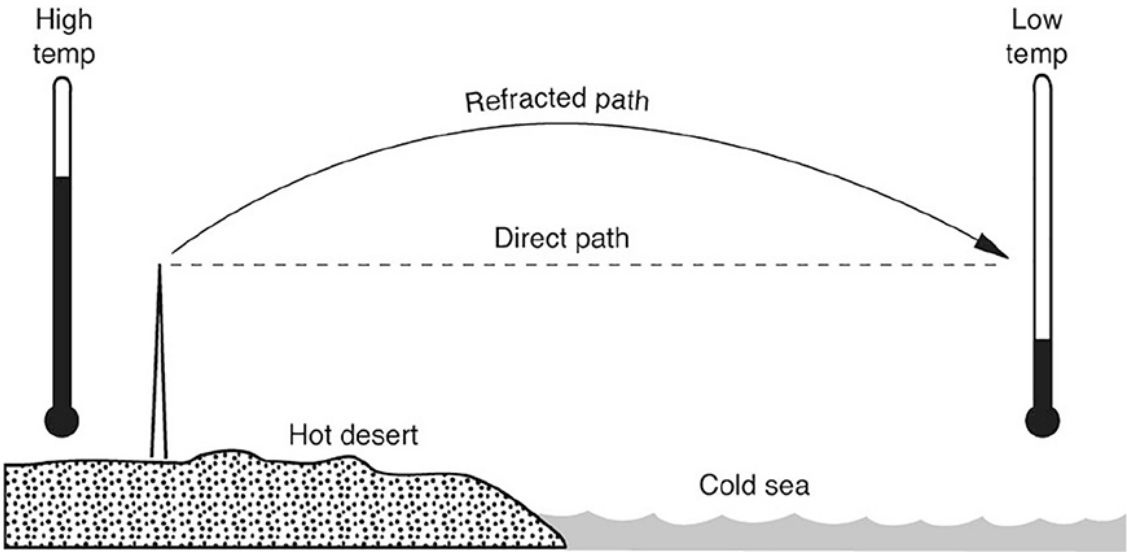


FIGURE 3.16 Superrefraction phenomena.

The second form of refraction is weather-related. Called ducting, this form of propagation (Fig. 3.16) is a special case of superrefraction. Evaporation of seawater causes layered air masses in which the air temperature is greater than in the layers below it, or temperature inversion regions, to form in the atmosphere. Air temperature normally decreases with altitude, but at the boundary with an inversion region it increases. The inversion layer forms a "duct" that acts like a waveguide. In Fig. 3.17, the distance

D_1
is the normal radio horizon distance, and
D_2
is the distance over which communications can occur though ducting.

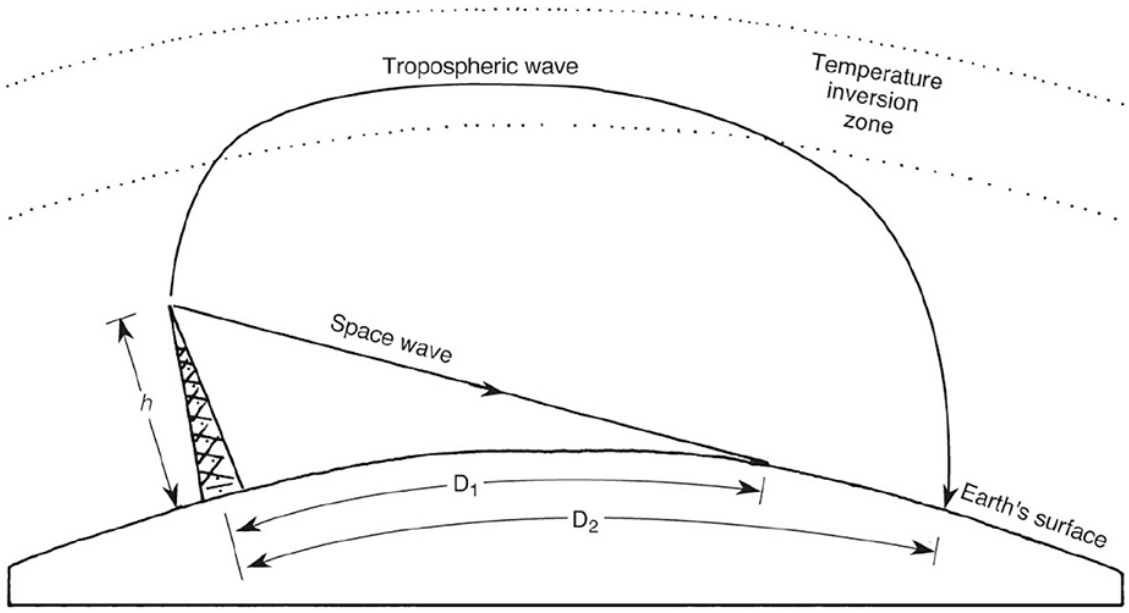


FIGURE 3.17 Ducting phenomenon.

Ducting allows long-distance communications from lower VHF through microwave frequencies, with 50 MHz being a practical lower limit and 10 GHz being a rough upper limit. Airborne radio transmitters can experience ducting at higher microwave frequencies.

To use ducting as a propagation method, antenna placement is critical. The receiving and transmitting antennas must be either (1) inside the duct physically (as in airborne cases) or (2) able to radiate at an angle such that the signal is trapped inside and received from the duct. Distances up to 2500 mi or more are possible through ducting. Certain paths where frequent ducting occurs have been identified:

- Great Lakes to Atlantic seaboard

- Newfoundland to Canary Islands
- Florida to Texas across the Gulf of Mexico
- Newfoundland to the Carolinas
- California to Hawaii
- Ascension Island to Brazil

A related condition occurs in polar regions, where colder air from the land mass flows out over warmer seas, causing subrefraction (Fig. 3.18). This phenomenon bends EM waves away from the earth's surface, reducing the radio horizon by about 30 to 40 percent.

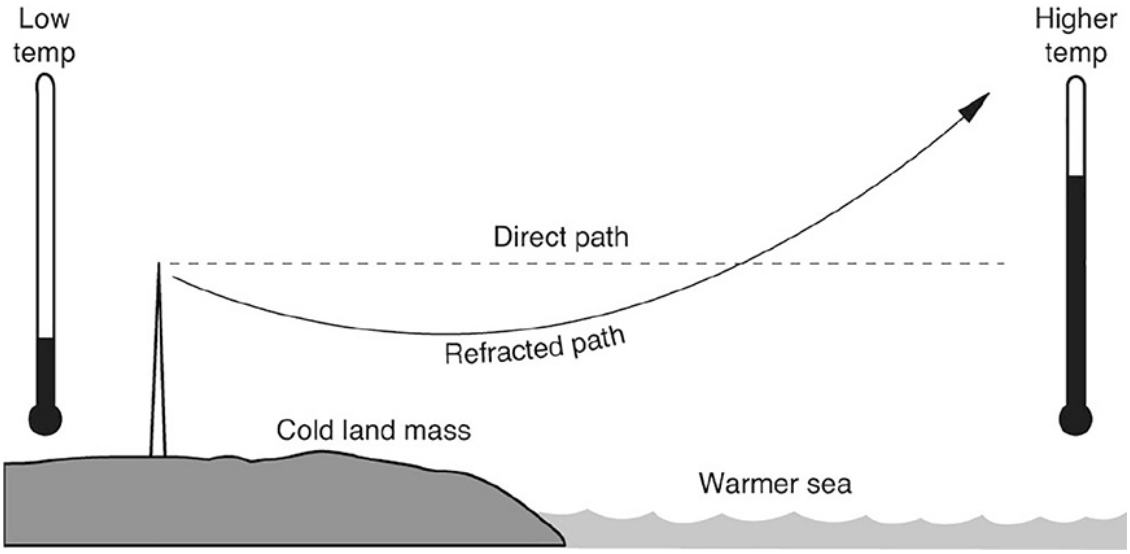


FIGURE 3.18 Subrefraction phenomena.

Tropospheric propagation is dependent upon air-mass temperatures and humidity and changes over the course of the day with the local rising and setting of the sun. These variations are called diurnal. Received signals may vary by 20 dB in strength over a 24-hour period. These tropospheric phenomena explain how TV, FM broadcast, and other VHF signals can propagate great distances at certain times—especially along seacoast paths—while being weak or undetectable at others.

3.6 Fading Mechanisms

Variations over time in received amplitude and/or phase are called fading. Fading originates in changing propagation conditions between the transmitter and the receiver. Common sources include weather-related path loss, multipath, vehicle motion, and variations in the height and density of the ionospheric layer of the atmosphere.

Fading can be experienced by line-of-sight radio relay links. Fading of 20 dB is relatively common and can reach levels of 30 dB. Fading phenomena in the VHF range and higher frequencies can last several hours, with some periods of several days reported on rare occasions.

Fading is generally the result of varying path loss or phase shift on at least one path between the transmitting (TX) and receiving (RX) antennas. In Fig. 3.19, Ray A represents the direct path between source and destination. Ray B is an example of energy from the same transmitting antenna traveling through an elevated layer or other atmospheric anomaly capable of providing refraction or subrefraction of the transmitted wave. Ray C adds signal via a ground reflection path. Ray D is an example of yet another wave from the transmitting antenna arriving at the receiver via subrefraction fading.

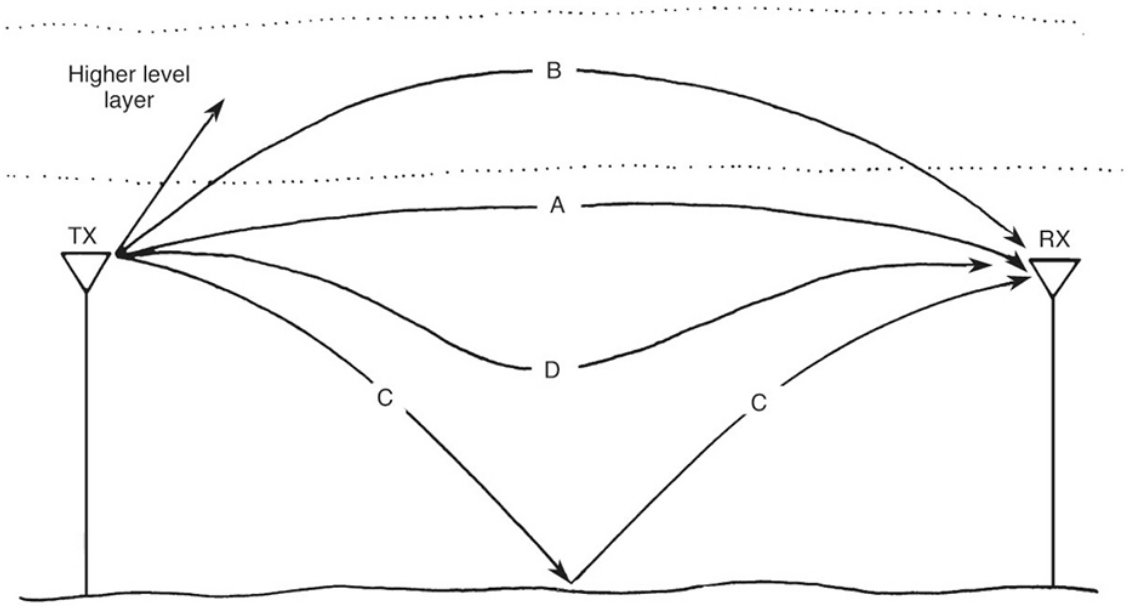


FIGURE 3.19 Multiple signal paths between transmitter and receiver.

All of the multipath rays will have amplitudes and phases that differ from those of the direct path. Other rays may be out of phase with the received voltage from Ray A. If one of the rays varies over time, then the received signal, which is the sum of all rays, will change due to constructive or destructive interference, and fading occurs. A moving vehicle midway between the transmitting and receiving sites may interrupt path C. While the moving vehicle is in the ground reflection path, it may create a new and different geometry for reflection if the incoming signal hits its roof and reflects off in the direction of the receive antenna. The ray paths are subject to weather-related changes due to wind, temperature or pressure gradients, or precipitation.

Fading mechanisms are usually frequency-sensitive, so one frequency may have high signal loss whereas a nearby frequency may be strong. A possible countermeasure for fading is frequency diversity. In many cases, hopping over a 5 percent frequency range will help eliminate or reduce the fading. If other system constraints and/or local spectrum usage prevent a 5 percent range, 2 or 3 percent may be adequate.

Over oceans or other large bodies of water fair weather surface ducting is sometimes encountered. These ducts form in the midlatitudes, starting about 2 to 3 km from shore, up to heights of 10 to 20 mi, with wind velocities in the 10- to 60-km/h range. Power fading is caused by the presence of the duct along with surface reflections (see Fig. 3.20). Fading alone can occur when there is a superrefractive duct elevated above the surface. The duct has a tendency to act as a waveguide and focus the signal (Fig. 3.21). Although the duct shown is superrefractive, subrefractive ducts are also possible.

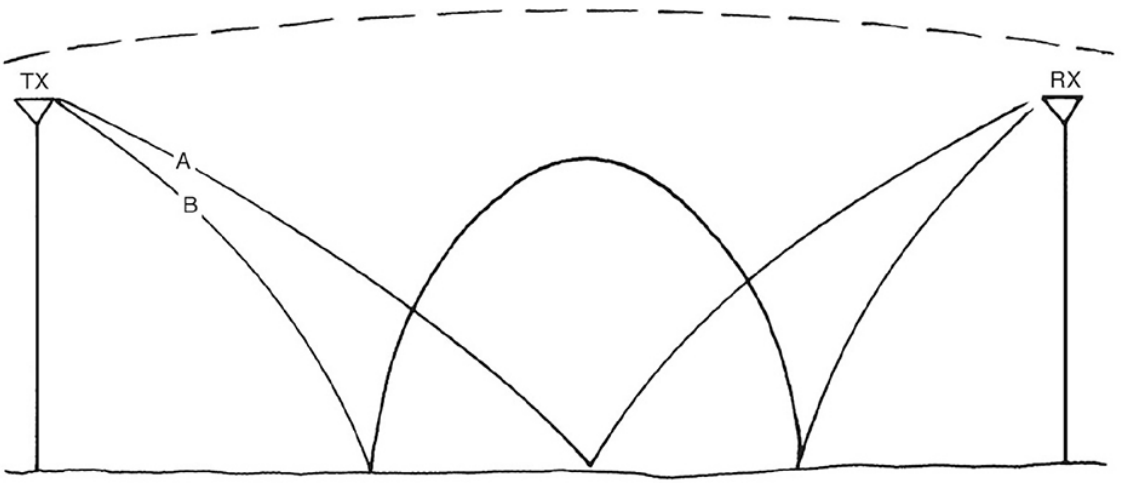


FIGURE 3.20 Multihop interference.

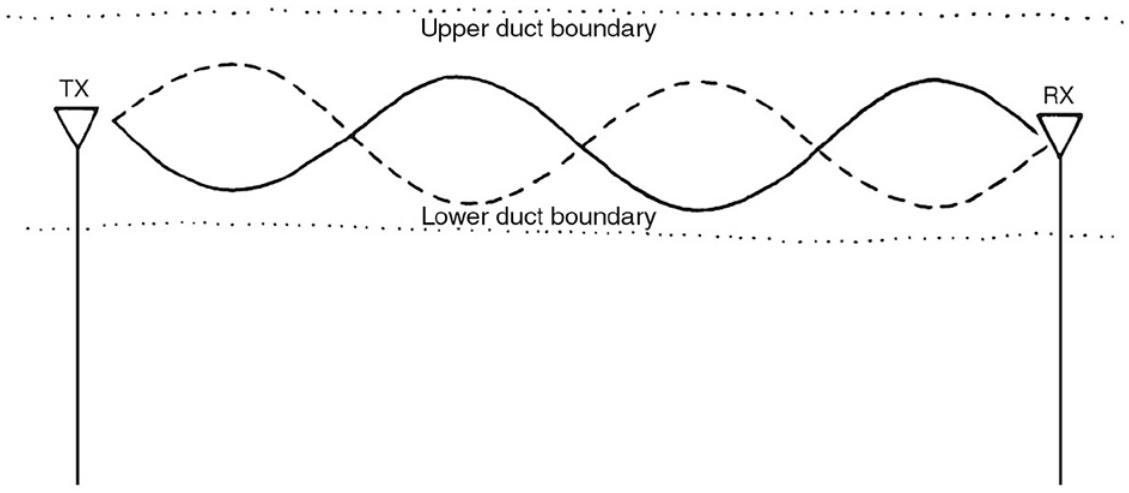
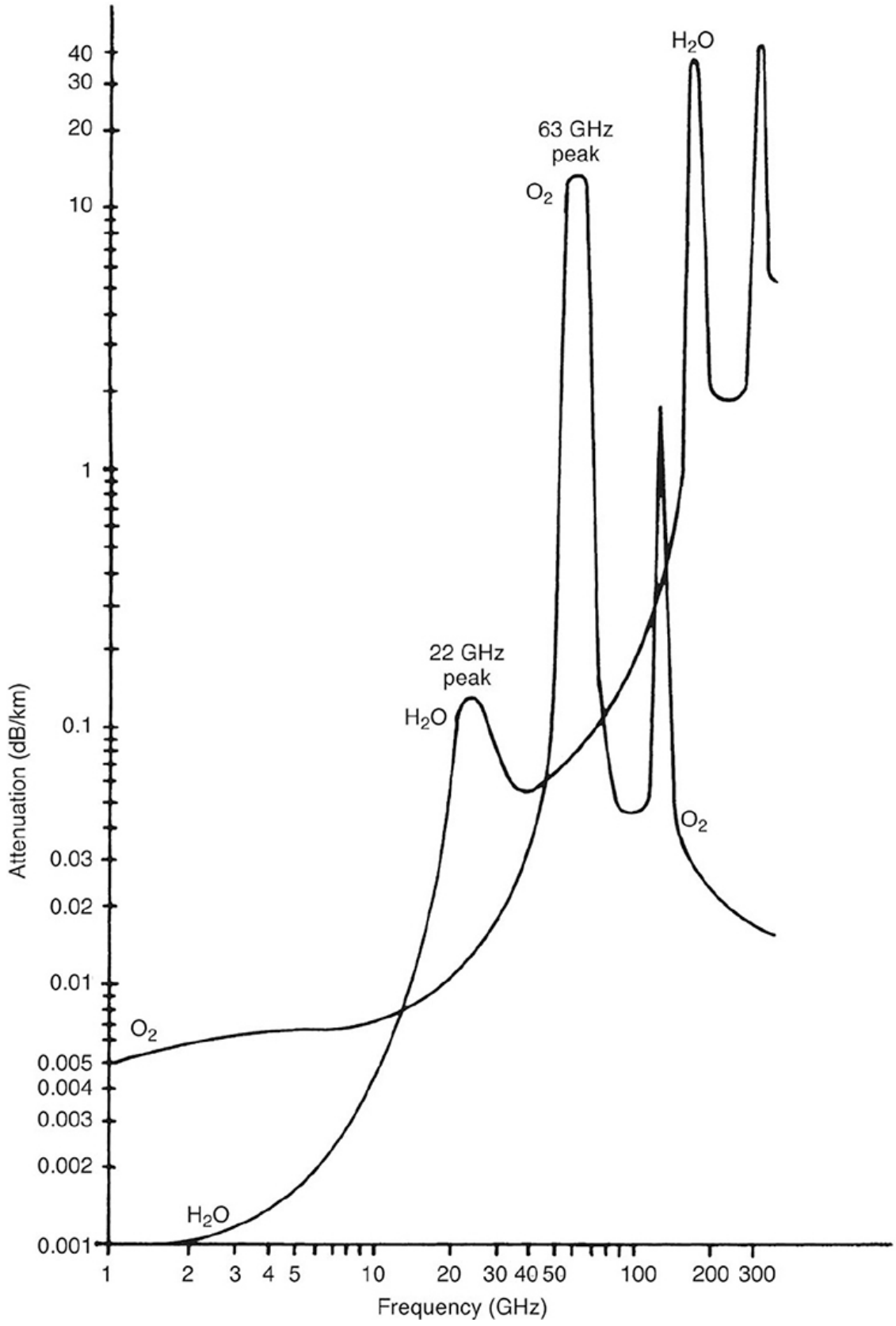


FIGURE 3.21 Duct focusing.

Microwave communication paths above about 10 GHz suffer increasingly severe attenuation caused by water vapor and oxygen in the atmosphere. Figure 2.25 shows the standard attenuation in decibels per kilometer for microwave frequencies. In addition to a general upward slope to the graph there are strong peaks at several specific frequencies. Selecting a system frequency within any of these ranges will cause poor communications or will require a combination of more transmit power, better receiver sensitivity, and better antennas at both receiving and transmitting locations. Further, the curves of Fig. 3.22 are based on certain assumed standard conditions. Rain and other forms of airborne precipitation can severely increase the attenuation of signals beyond the amounts shown in the chart.


Attenuation (dB/km)
40
30
20
10
1
0.1
0.03
0.02
0.01
0.005
0.004
0.003
0.002
0.001
1
2
3
4
5
10
20
30
40
50
100
200
300
Frequency (GHz)
H2O
63 GHz peak
O2
22 GHz peak
H2O
O2
O2
H2O

F IGURE 3.22 Atmospheric absorption of radio signals at microwave frequencies.

CHAPTER 4
Ionospheric Propagation

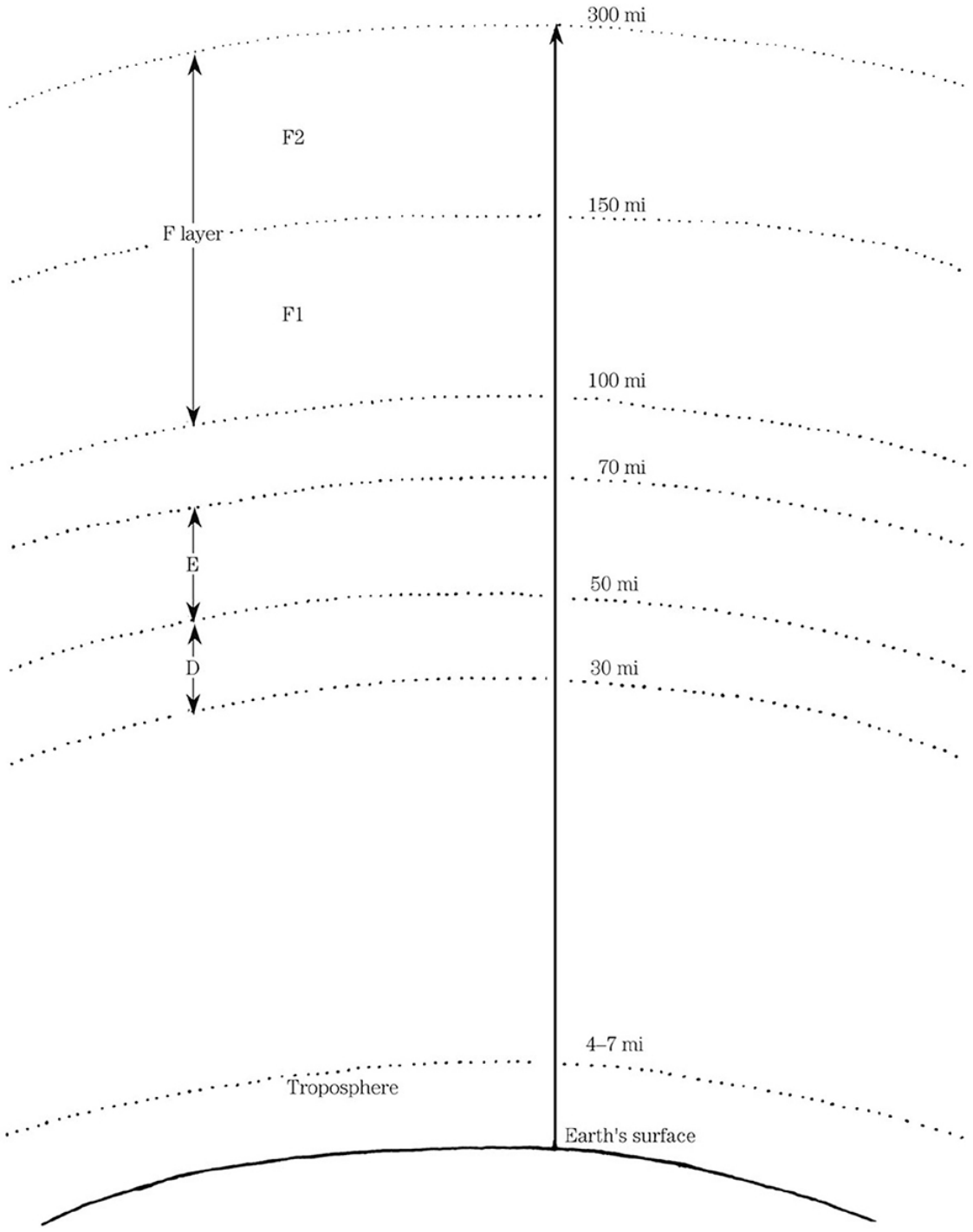
After Hertz, Marconi, and other pioneering radio scientists of the late 1800s, the early 1900s found experimenters from all walks of life constructing radio transmitting and receiving equipment. By 1909, radio clubs had been formed at a few colleges around the United States, and by 1914, the American Radio Relay League (ARRL) had been founded. Commercial, government, and amateur stations shared the airwaves in willy-nilly fashion until around 1912 when the United States began to regulate radio. In those days, "good DX" might mean the ability to communicate between, say, Boston and New York City. The prevalent assumption was that the longer wavelengths were superior for extending those distances. Forced to shift from wavelengths as long as 1000 m (well below the bottom of today's AM broadcast band) to 200 m (roughly around 1500 kHz, near the top of the AM broadcast band) by the new regulations, amateurs grumbled a bit but dug in and figured out how to attain better DX distances. The distances achieved by persistent amateur radio enthusiasts increased and experts were finding it harder and harder to explain how this was possible.

World War I brought a halt to all amateur transmissions. When the war was over the US government was inclined to keep it that way. The ARRL convinced Congress and various federal agencies of the merits of radio amateurs, and in the fall of 1919 amateurs returned to the airwaves. In 1921 and again in 1922, the ARRL sponsored transatlantic receiving tests, and in both events too many stations were heard for it to be a fluke: 30 or so the first year and 10 times that number the following year!

Spurred on by these encouraging results, the ARRL and others began to investigate what it would take to convert these successes into two-way communication across the ocean. Despite the almost

unanimous opinion of the best scientists of the day that the shorter wavelengths would prove to be useless, some hardy experimenters plunged below 200 m and, within a year, had shown that the shorter the wavelength, the better the long-distance transmission.

Textbooks and magazines of the era were rife with theories to explain these unexpected results. Ultimately, the answer turned out to be one advanced in 1901 when Arthur Kennelly of the United States and Oliver Heaviside of England independently proposed the existence of a "reflecting" layer circling the earth some miles above it. Soon named the Kennelly-Heaviside layer, it is now known as the E layer (Fig. 4.1) of the ionosphere.



F IGURE 4.1 Classification of the earth's atmosphere for radio propagation.

Thus, intercontinental radio communications were made possible on the MF, HF, and lower VHF bands (between roughly 0.3 and 70

MHz) by the ionosphere. Included in the range of frequencies for which ionospheric propagation is possible are the AM broadcast band, many traditional maritime frequencies, and amateur bands from 160 to 6 m. Especially at the upper end of the range, during sunspot maxima or during brief periods of unusual solar activity long communication range are frequently observed.

Kennelly and Heaviside described the unseen layer as a reflecting layer, but the sparseness of the ionization and the variation of ionization with altitude actually make it a refracting layer. As can be seen from Fig. 4.2, the refracting action of the layer beginning at B-C can be approximated by pretending there is a reflecting surface at altitude C′.

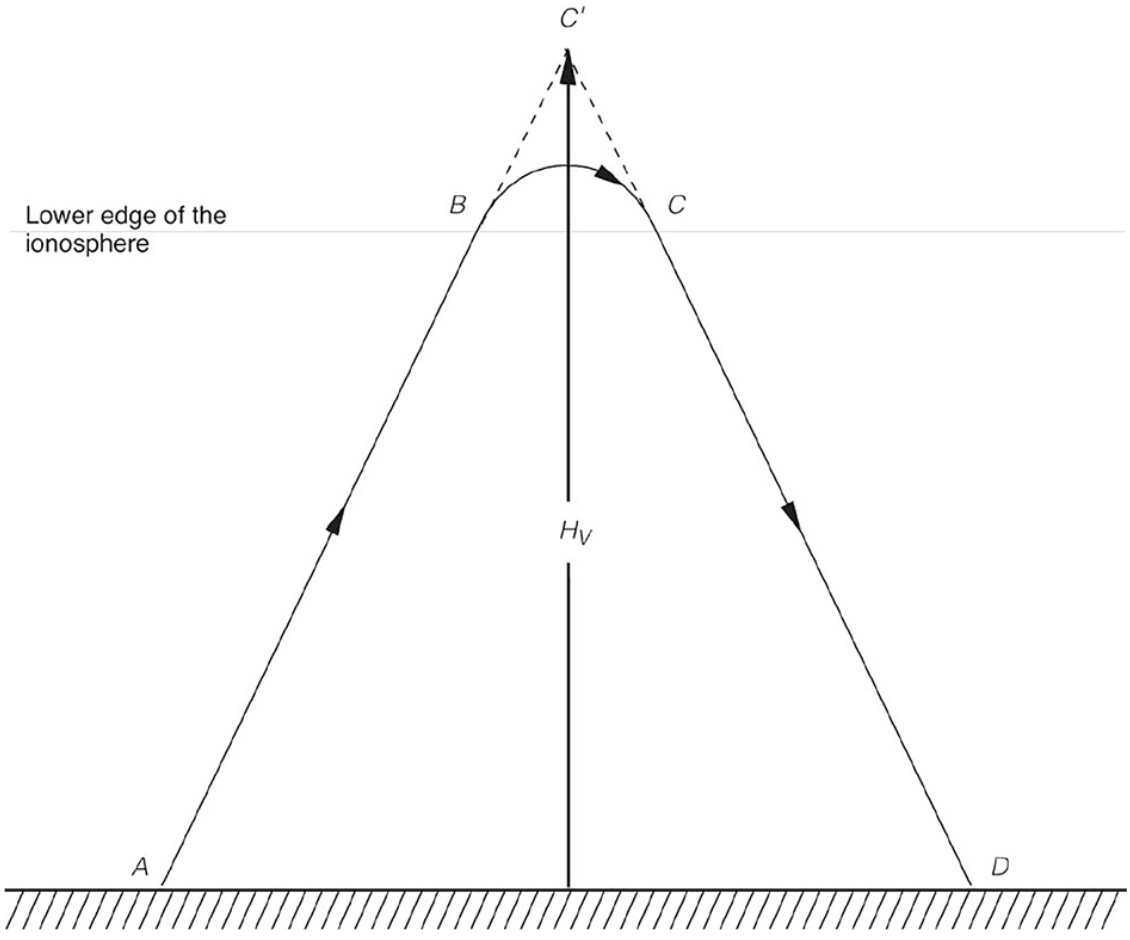


FIGURE 4.2 Finding the virtual height of the ionosphere.

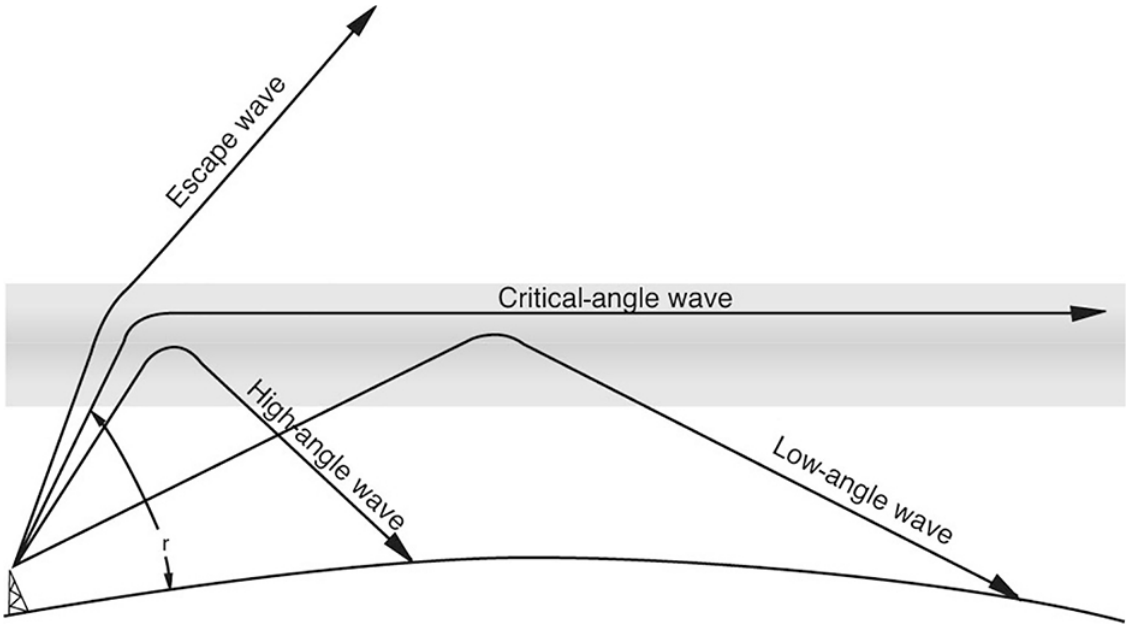
The ability of the ionospheric layer to return a perfectly vertical signal back to earth is frequency-dependent. Above a certain frequency, called the critical frequency, or

f_C

, signals sent straight up continue into space. If transmitted signals hit the layer at an angle to the vertical (Fig. 4.3), at some transmitted takeoff angle (

α_r

) the transmitted signal no longer escapes from the ionosphere but remains within it. This is called the critical angle, and it changes over frequency.



F IGURE 4.3 Sky-wave propagation as a function of antenna radiation angle.

At takeoff angles greater than the critical angle, the transmitted wave either stays within the ionospheric layer or continues on into space. For angles between the critical angle and 90 degrees (pure vertical incidence), the ionosphere does not return the transmitted signal to earth and does not support earth-to-earth communications links at this frequency. The distance on the earth's surface from the transmitting antenna out to the closest point of returned signal

from the ionosphere is called the skip distance and locations closer to the transmitter than that distance are said to be inside the skip zone.

Most importantly, takeoff angles smaller than the critical angle will be refracted from the layer and returned to earth. We see from Fig. 4.3 that there is a range of takeoff angles for the transmitted signal that will be "reflected" back to earth at some distance from the source. This is the magic of the ionosphere.

4.1 The Sun's Effect on the Ionosphere

Several sources of energy contribute to ionization of the upper atmosphere. Cosmic radiation from outer space creates some degree of ionization. The majority is caused by solar energy. Events on the surface of the sun sometimes cause the radio mirror to be almost perfect, and this situation makes spectacular propagation possible. At other times solar disturbances (Fig. 4.4) can disrupt radio communications for days at a time.



FIGURE 4.4 Solar event affecting radio propagation on the earth.

There are two principal forms of solar energy that affect shortwave communications, short-wavelength electromagnetic radiation and charged particles. Most of the radiation is above the visible spectrum in the ultraviolet and X-ray/gamma-ray region. Because electromagnetic radiation travels at the speed of light, solar events that release radiation cause changes to the ionosphere about eight minutes later. Charged particles have a finite mass and travel at a considerably slower velocity, requiring two or three days to reach earth.

Multiple sources of both radiation and charged particles exist on the sun. Solar flares release huge amounts of both radiation and particles. These events are unpredictable and sporadic. Some layers of the sun rotate with a period of approximately 27 days, causing specific sources of radiation to face the earth once every 27 days. Many components of solar radiation levels and the corresponding ionospheric activity also tend to repeat every 27 days.

At some radio frequencies, solar and galactic noise establish a practical lower limit on the reception of weak signals. Solar noise can affect radio propagation and act as a harbinger of changes in propagation patterns. Solar noise can be demonstrated by using an ordinary radio receiver and a directional antenna operating in the VHF/UHF regions of the spectrum (150–152 MHz frequently is used). Simply aim the antenna at the sun on the horizon at either sunset or sunrise and a dramatic change in background noise will be noted as the sun rises or sets across the horizon.

4.1.1 Sunspots

Sunspots are the most well-known contributor to the sun's effect on our ionosphere. Sunspots (Fig. 4.5) can be as large as 70,000 to 80,000 mi in diameter and often appear in clusters. The number of sunspots typically visible on the face of the sun on any given day is quite variable in the short term, but as shown in Fig. 4.6, it exhibits a fairly consistent and pronounced long-term cycle over a period of approximately 11 years. This is a very rough number; actual cycles since 1750 (when records were first kept) have varied from 9 to 14 years. The sunspot number is reported daily as the statistically massaged Zurich smoothed sunspot number, or Wolf number. The monthly smoothed sunspot number (SSN) is based on a weighted average of a 12-month period centered on the month being reported. A century of observations has confirmed the strong relationship between the SSN and general levels of ionospheric radio propagation. The lowest monthly SSNs calculated during the early radio era hovered around 1.5 in mid-1913 and did not go that low again until mid-2008. The highest recorded monthly SSN was 201 (in March of 1958). Daily sunspot numbers vary considerably and can remain at zero for weeks on end (as they did throughout much of 2008 and 2009) or rise to values of 150 percent or more of the corresponding monthly SSNs. NOAA's National Geophysical Data Center Web site (www.ngdc.noaa.gov/stp/SOLAR) is an excellent source of background information on the methodology involved.

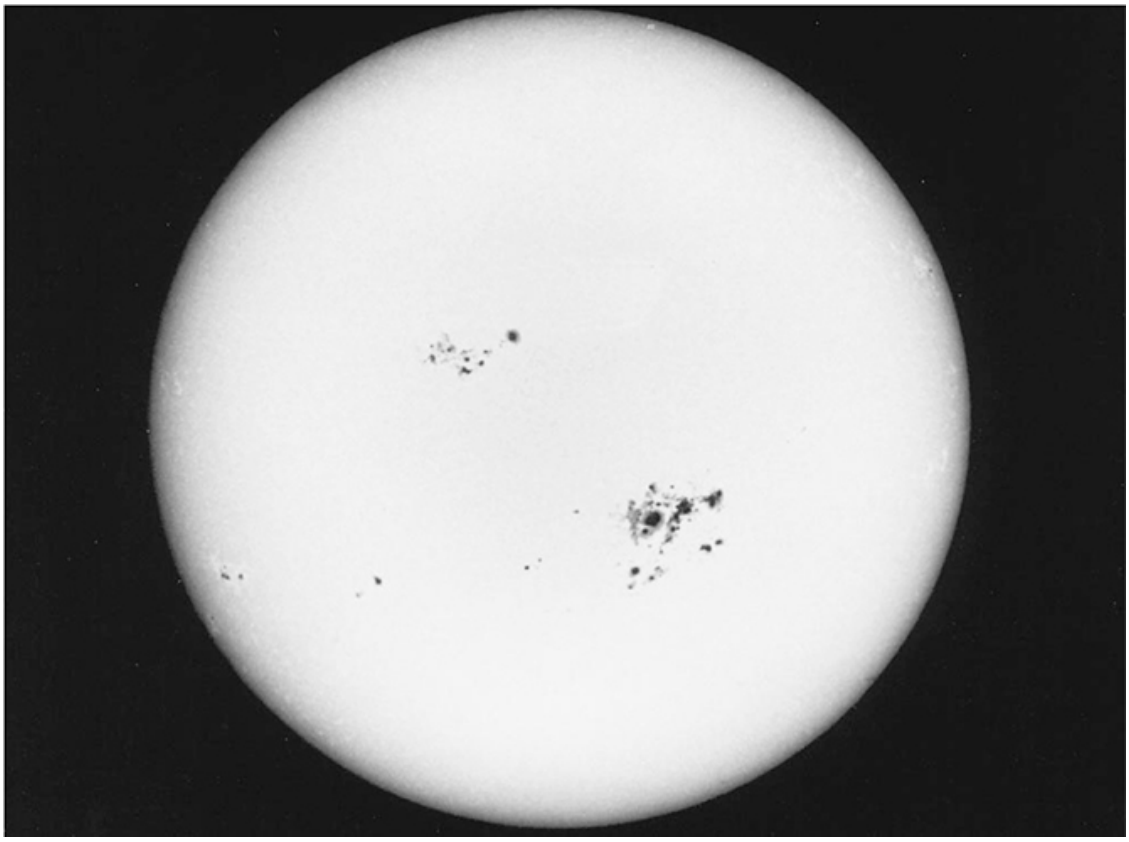
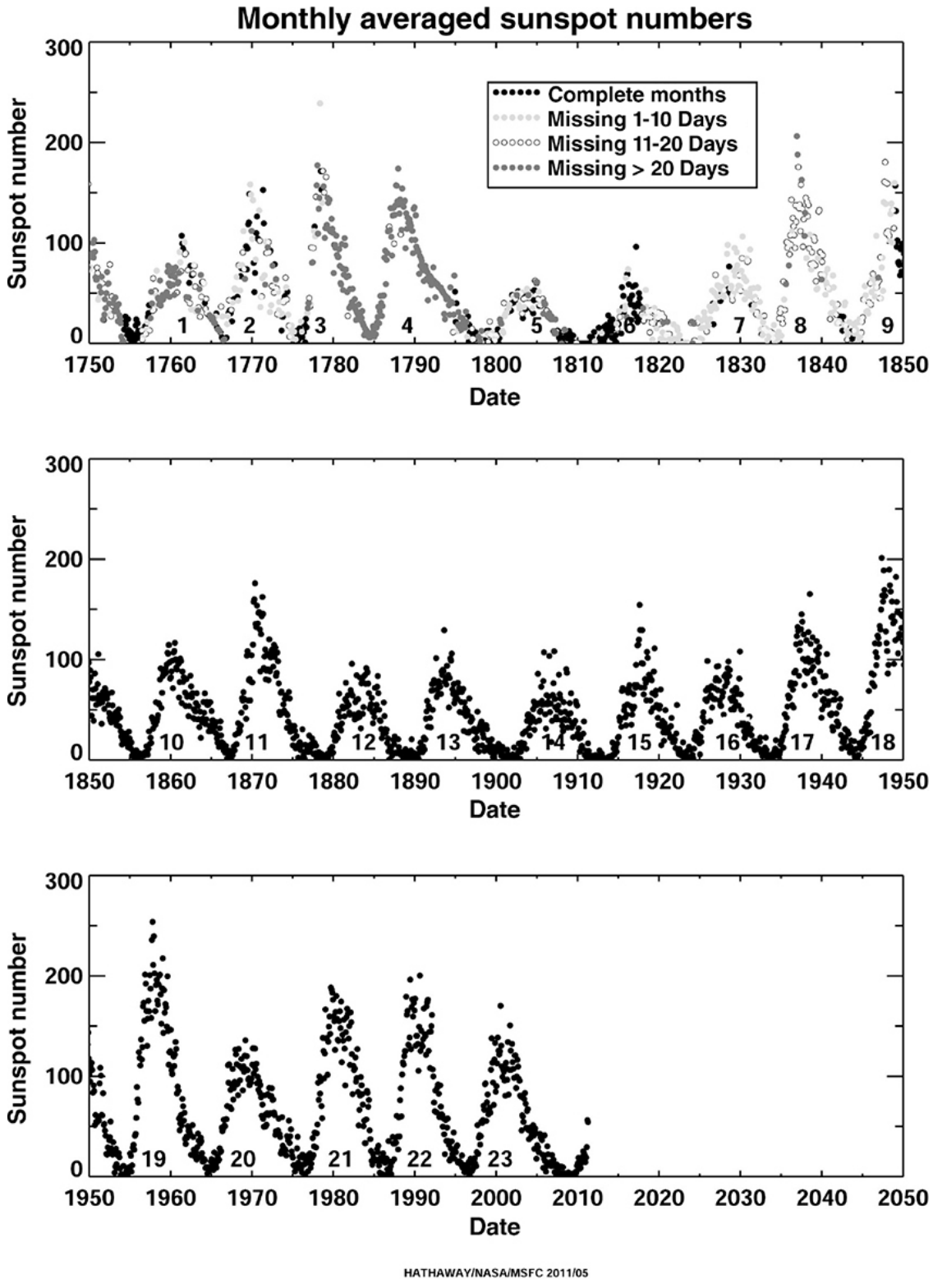


FIGURE 4.5 Sunspots.



F IGURE 4.6 Monthly averaged sunspot numbers, 1750–2011.

For many decades, the smoothed sunspot number was the best available indicator of ionospheric propagation potential. With the availability of satellites, balloons, and other extraterrestrial probes, other measures now augment SSNs. One such indicator is the solar flux index (SFI). This measure is taken in the microwave region (at a wavelength of 10.2 cm, or 2.8 GHz), at 1700 UTC at Ottawa, Canada. The SFI is reported in an over-theair announcement hourly by the National Institute for Standards and Technology (NIST) radio stations WWV (Fort Collins, Colorado) and WWVH (Maui, Hawaii) and is also available from many additional sources via the Internet. Two short-term propagation forecasting tools, the A-index and the K-index, relate to how ionospheric disturbances and are included in NIST announcements. Many other solar and ionospheric monitoring tools are available and can be found online.

4.2 Ionospheric Layers

The ionosphere is divided, for purposes of radio propagation studies, into multiple layers. Like the boundary between the ionosphere and the stratosphere, the boundaries between these layers are approximate. Layers do not have sharply defined boundaries and blend one into another. Ionospheric layers (Fig. 4.1) are designated D, E, and F, with F being further subdivided into F1 and F2 sublayers.

4.2.1 D Layer

The D layer is the lowest layer in the ionosphere, at approximately 30 to 50 mi above the surface. This layer is not ionized as much as the higher layers because all forms of solar energy that cause ionization are attenuated by the higher layers before the energy reaches the D layer. The D layer is denser than the E and F layers, causing any positive and negative ions to recombine and form electrically neutral atoms more quickly than in the higher layers.

The extent of D-layer ionization is proportional to the elevation of the sun and achieves maximum intensity at midday or shortly

thereafter. The D layer exists mostly during the warmer months of the year because of both the greater height of the sun above the horizon and the longer hours of daylight. The D layer almost completely disappears after local sunset, although some observers have reported sporadic incidents of D-layer activity for a considerable time past sunset. The D layer typically is a strong absorber of medium-wave signals to such an extent that signals below 3 or 4 MHz are completely absorbed by the D layer. Because of this, most AM broadcast stations during daylight hours enjoy only local ground-wave coverage.

4.2.2 E Layer

The E layer extends from approximately 50 to 70 mi above the earth's surface and is considered the lowest region of the ionosphere that supports, rather than absorbs, ionospheric radio communications. Like the D layer, this region is ionized only during the daylight hours, with ionization levels peaking at midday. The ionization level drops off sharply in the late afternoon and almost completely disappears after local sunset.

For most of the year, the E layer is absorptive and does not reflect radio signals. During the summer months E-layer propagation occurs frequently. A phenomenon called short skip (i.e., less than 100 mi for medium-wave and 1000 mi for shortwave signals) occurs in the E layer during the summer months and in equatorial regions at other times.

Sporadic E propagation is another phenomenon associated with the E layer. This mode is made possible by scattered zones of intense ionization in the E-layer region of the ionosphere. Sporadic E propagation affects the upper HF and lower VHF region. It is observed most frequently in the lower VHF spectrum (30 to 150 MHz) but it is also sometimes observed at higher frequencies. Skip distances on VHF can reach 500 to 1500 mi on one hop, especially in the lower VHF region (including the 6-m band).

4.2.3 F Layer

The F layer of the ionosphere is the primary support for long-distance shortwave communications. This layer is located between 100 and 300 mi above the earth's surface. Unlike the lower layers, the air density in the F layer is low enough that ionization levels remain high all day and decay slowly after local sunset. Minimum levels are reached just prior to local sunrise. Because of its height, the F layer can support skip distances up to 2500 mi on a single hop.

During the day, the F layer often splits into two identifiable and distinct sublayers, designated F1 and F2. The F1 layer is found approximately 100 to 150 mi above the earth's surface, with the F2 layer above the F1, extending up to the 270- to 300-mi limit. Beginning at local sundown the lower regions of the F1 layer begin to deionize as positive and negative ions recombine. After local sunset the F1 and F2 layers merge to become a single reduced layer beginning at about 175 mi.

The height and degree of ionization of the F2 layer varies with local sun time, with the season of the year, and with the 27-day and 11-year sunspot cycles. The F2 layer begins to form shortly after local sunrise and reaches maximum shortly after noon. During the afternoon, the F2-layer ionization begins to decay in an exponential manner until, for purposes of radio propagation, it disappears sometime after local sunset. During sunspot maxima, F-layer ionization does not completely disappear overnight, and the 20-m amateur band may be open for long-distance communication day and night.

4.3 Measures of Ionospheric Propagation

Characteristics of the ionosphere can be measured and used to make predictions of radio activity and long-distance propagation. The critical frequency (f_C) and maximum usable frequency (MUF) measure the state of ionization and long-range communications ability.

4.3.1 Critical Frequency

The critical frequency, f_C, is the highest frequency that is reflected when a signal strikes the ionosphere as a vertical (90 degrees with respect to the reflecting surface) incident wave. Values of f_C can be as low as 3 MHz during the nighttime hours and as high as 10 to 15 MHz during the day.

4.3.2 Virtual Height

Radio waves below the critical frequency are refracted by the ionosphere and return to earth. Such waves appear to have been reflected from an invisible radio "mirror." The height of this apparent mirror is called the virtual height of the ionosphere.

4.3.3 Maximum Useable Frequency (MUF)

The MUF is the highest frequency at which communications can take place via the ionosphere over a given path. Normally the MUF for a given transmitter site is found to be at the farthest distance(s) the signal can graze the ionosphere in the direction of the sun. As a rough rule of thumb, the MUF is approximately three times higher than the critical frequency.

Both the MUF and the critical frequency vary geographically, and they become higher at latitudes close to the equator. The critical frequency is determined by bouncing a signal off the ionosphere directly overhead; the measured MUF, on the other hand, will depend on the state of the ionosphere, which is 2500 mi away from the site where f_C is measured, so the two indices do not track perfectly.

The best propagation occurs on average at frequencies just below the MUF. The frequency of optimum traffic (FOT) is approximately 85 percent of the MUF. Near the FOT, signal

strengths are optimum because ionospheric absorption is at a minimum, and signal-to-noise ratios (SNRs) are highest because atmospheric noise generally decreases with increasing frequency.

4.3.4 Lowest Useable Frequency (LUF)

At low frequencies, ionospheric absorption and atmospheric noise conspire to hinder and even prevent radio communications. The lowest useable frequency (LUF) is a measure intended to help identify the lower frequency limit for successful communications. Unlike the MUF, the LUF is not dependent solely on atmospheric physics. The LUF of a system can be varied by controlling the signal-to-noise ratio of the transmitter-receiver path. Although many of the factors that contribute to SNRs are beyond our control, the effective isotropic radiated power (EIRP) of the transmitter is one that often can be changed. A 2-MHz decrease in LUF is available for every 10-dB increase in the EIRP of the transmitter. This can be accomplished through the use of higher power, lower transmission line losses, higher gain antennas at both ends of the circuit, or narrower receiver filter bandwidth.

4.4 Regular Ionospheric Variations

The ionosphere is an extremely dynamic region of the atmosphere, especially from a radio operator's point of view, because it significantly alters radio propagation on a minute-by-minute basis. It is convenient to divide the dynamics of the ionosphere into regular variations and disturbances. There are several forms of variation seen on a regular basis in the ionosphere:

- Diurnal (daily)
- 27-day (monthly)
- Seasonal
- 11-year and 22-year cycles

4.4.1 Diurnal (Daily) Variation

Because the sun is the principal source of ionization of the upper atmosphere, diurnal variations occur in the ionosphere. The E and D layers merge at night. The height of the F2 layer increases until midday and decreases until evening, when it disappears or merges with other layers. As a result of higher absorption in the E and D layers, lower frequencies are not as useful during daylight hours. At the same time, the F layers support reflection at higher frequencies during the day. As a rough rule of thumb, for ionospheric skip communications purposes, frequencies below 10 MHz can be thought of as primarily "nighttime" bands and the frequencies above 10 MHz are primarily "daylight" bands.

4.4.2 27-Day Cycle

Radiation from the sun exhibits a 27-day cycle caused by the rotational period of the sun's surface. Sunspots form or come to the surface of the sun near its poles then drift toward the equatorial region. During this latitude migration they tend to rotate in synchronism with the surface of the sun as a whole, so they will face the earth only during a portion of each month. As new sunspots are formed, they are not visible on the earth until their region of the sun rotates earth-side. Sunspots are believed to be regions of swirling high-intensity magnetic fields and plasmas that somehow stimulate greater overall solar radiation. Paradoxically, as seen in Fig. 4.5, sunspots are darker than the rest of the sun's surface, indicating that they are cooler.

4.4.3 Seasonal Variation

The earth's tilt varies the angle of exposure of our planet to the sun on a seasonal basis. The earth's annual orbit is elliptical, so the intensity of the sun's energy that ionizes the upper atmosphere varies with the seasons of the year. Ion density in the F2 layer tends to be highest in winter and less in summer. During the summer, the distinction between F1 and F2 layers is less obvious.

4.4.4 11-Year and 22-Year Cycles

The number of sunspots varies on an approximately 11-year cycle. The ionospheric characteristics that affect radio propagation also vary on an 11-year cycle. Skip propagation on the HF and lower VHF bands is most prevalent when the smoothed sunspot number is highest. During the peak years (1957–61) of the strongest sunspot cycle of the radio era to date, true F-layer skip propagation was frequently enjoyed on 6 m (50–54 MHz) and higher, affecting even some VHF TV channels and possibly even the low end of the FM broadcast band. By comparison, during a recent minimum (2007–10), many populated areas of the midlatitudes went weeks or months with no evidence of skip activity above 20 MHz.

Sunspots are regions of high magnetization, and their polarity on either side of the sun's equator switches with every sunspot cycle. Thus, the true cycle of the sun's activity is a 22-year cycle, but the effect of sunspots on ionospheric propagation is unrelated to the magnetic polarity of the spots, so, for all practical purposes, we experience only the 11-year cycle.

4.5 Ionospheric Disturbances

Disturbances in the ionosphere can have a profound effect on radio communications— and most of them (but not all) are negative. This section briefly examines some of the more commonly encountered types.

4.5.1 Sporadic E Layer

A reflective cloud of ionization sometimes appears in the E layer of the ionosphere. This layer is sometimes called the sporadic *E*, or E_S, layer. Sporadic E propagation is normally thought of as a VHF phenomenon with most activity between 30 and 100 MHz and decreasing activity up to about 200 MHz. At times sporadic E propagation is observed on frequencies as low as 10 or 15 MHz. Reception over paths of 1400 to 2600 mi is possible in the 50-MHz region when sporadic E is present. In the northern hemisphere, the

months of June and July are the most prevalent sporadic E months. When sporadic E is present, it typically lasts only a few hours. Sporadic E layer activity is related to weather and radio users sometimes track the progress of weather fronts in their part of the world.

4.5.2 Sudden Ionospheric Disturbances (SIDs)

Ionospheric disturbances that occur rapidly and without warning are sudden ionospheric disturbances (SIDs) or Dellinger fades. An SID can last from a few minutes to many hours. SIDs often occur in conjunction with solar flares that emit an immense amount of ultraviolet radiation over a short period. When this UV radiation reaches the earth, the SID causes a tremendous increase in D-layer ionization and abnormally high levels of absorption result. The ionization is so intense that MF/HF receivers on the earth's sunlit hemisphere experience profound loss of signal strength above about 1 MHz. Received signal strengths can drop by 60 to 90 dB in a few minutes at the onset of the SID, and it is not uncommon for amateurs, shortwave listeners, and others monitoring the HF bands to think the cause is a malfunction in their receivers or antennas.

SIDs are often accompanied by variations in terrestrial electrical currents and magnetism levels. There is good correlation between the occurrence of an SID and aurora borealis or aurora australis in the southern hemisphere for a few hours or days thereafter. Because solar flare activity is far higher during periods of high sunspot activity, SIDs are most commonly observed during the peak years of the strongest sunspot cycles.

4.5.3 Ionospheric Storms

An abnormally large influx of particles in the upper atmosphere can produce an ionospheric storm. These storms are often preceded by SIDs 18 to 24 hours earlier and by an unusually large collection of sunspots crossing the solar disk 48 to 72 hours earlier. The storms occur most frequently and with greatest severity in the higher latitudes, decreasing toward the equator. When such a storm

commences, shortwave radio signals may begin to flutter rapidly and then drop out altogether. The upper ionosphere becomes chaotic, turbulence increases, and the normal stratification into layers is disrupted.

Radio propagation may come and go over the course of the storm, but it is mostly "dead." Unlike SIDs, which affect only the sunlit side of the earth, ionospheric storms are global. Both the MUF and the critical frequency tend to drop rapidly as the storm commences.

An ionospheric disturbance observed on November 12, 1960, was preceded by about 30 minutes of extremely good, but abnormal, propagation. At 2000 hours Greenwich mean time (GMT), European stations were heard in the United States with S9+ signal strengths in the 7000- to 7100-kHz region of the spectrum—an extremely rare midafter-noon occurrence when sunspot levels are high. After about 30 minutes, the bottom dropped out and even AM broadcast band skip later that evening was nonexistent. This disturbance occurred on the weekend of the annual ARRL "Sweepstakes" contest, a competitive operating event that brings out thousands of amateurs in the United States and Canada. The ARRL's summary of results in the May 1961 issue of QST led off with the editorial comment, "Universal lament: [propagation] conditions that first weekend!"

4.6 Ionospheric Sky-Wave Propagation

When signals in the ionosphere are refracted sufficiently to be bent back toward the earth's surface, sky-wave propagation occurs. The skip distance is the shortest distance along the surface of the earth from the transmitter site (*A* in Fig. 4.7) to the nearest of all points on the earth (point *C* in Fig. 4.7) for which the ionosphere is capable of refracting the transmitted signal back to earth. At shorter distances, the higher-angle radiation from the transmitted signal either remains in the ionospheric layer or continues into outer space. The ground-wave zone is the distance from the

transmitter site (*A* in Fig. 4.7) to the locus of points where the ground wave fades to an unusably low level (point *B* in Fig. 4.7). The skip zone is the distance from the outer edge of the ground-wave zone to the skip distance, or the distance from *B* to *C* in Fig. 4.7. Only one of these distances, the ground-wave zone, is strongly dependent upon transmitter power level (or, more precisely, effective radiated power). The other distances are a function primarily of transmitter frequency and the state of the ionosphere.

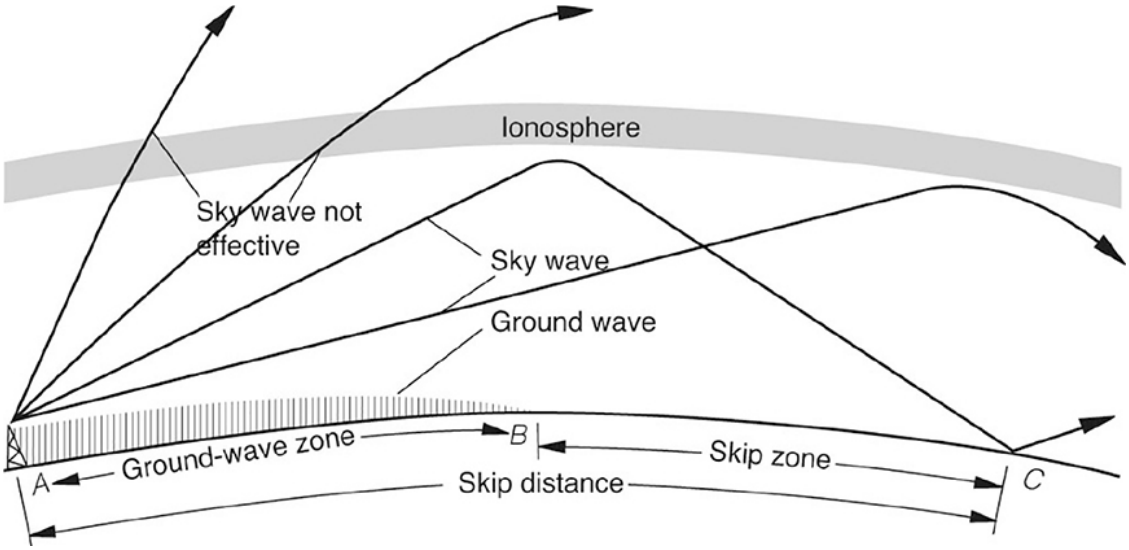


FIGURE 4.7 Ground-wave zone and skip zone.

At some frequencies and power levels, the sky wave and the ground wave may interfere with each other. When this happens, the sky wave arrives at the receiver with a signal whose amplitude and phase relative to the ground-wave signal depend on the specifics of its path to the ionosphere and back. Since the ionosphere is constantly undulating, the sky wave will arrive with randomly varying amplitude and phase relative to the ground wave. The sky wave can selectively strengthen or cancel the ground wave, giving rise to a class of fading mechanisms discussed further in Sec. 4.6.4.

4.6.1 Angle of Incidence and Radiation Angle

One of the factors determining skip distance is the angle at which the radio wave enters the ionosphere. If the transmit antenna is relatively omnidirectional, it illuminates the lower ionospheric boundary with energy at all angles. As discussed earlier, the radio frequency (RF) energy in angles nearly perpendicular to the boundary may or may not be refracted back to the earth, depending on the relationship between the transmitter frequency and
f_C
, the critical frequency. For a low enough frequency, transmitted energy will be refracted from the ionosphere at all antenna takeoff angles from perfectly vertical to slightly below horizontal (i.e., tangent to the earth's surface).

If the transmitted frequency is at
f_C
or below, the transmitted energy for all takeoff angles will be returned to the earth, eliminating the skip zone. When that occurs, the ionosphere supports communication over a wide range of distances from the transmit antenna, potentially all the way out to where the refracted wave hits the earth some 2500 mi (on average) distant.

This effect is illustrated Fig. 4.8. High-angle waves from the transmitting antenna are returned to the earth closer to the transmitter than low-angle waves are. When they hit the earth's surface (whether land or water) they are again refracted (or reflected) and once again toward the ionosphere. Although Fig. 4.8 shows signal paths for only two specific takeoff angles, there is an infinite number, all of which potentially bounce off the earth's surface and go back to the ionosphere. For extremely high angle waves, this process can occur tens of times before the wave ultimately reaches the receiving antenna far away.

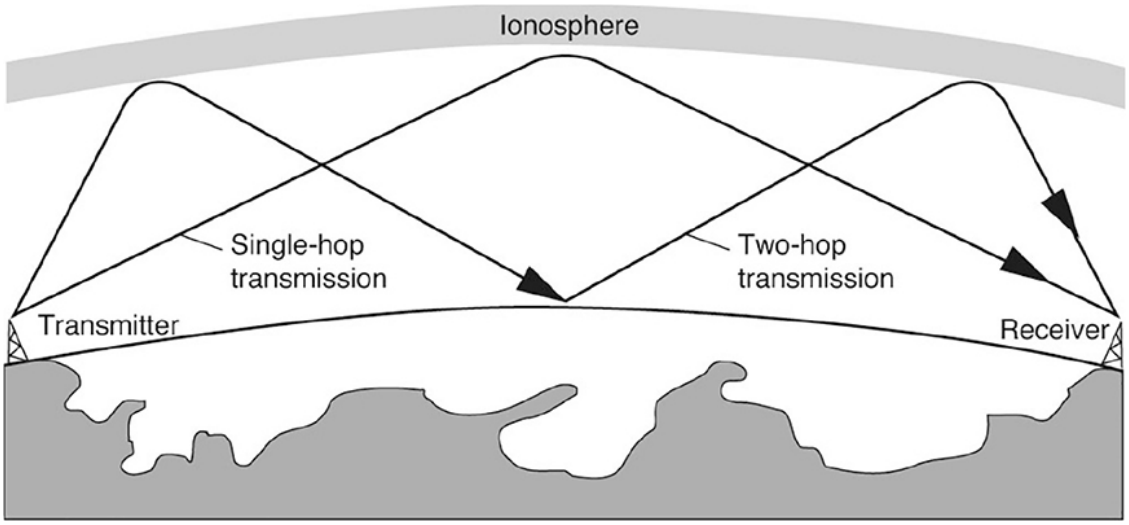


FIGURE 4.8 Single-hop and multihop skip communications.

To calculate probable signal strengths at the receiving site for various propagation modes, many propagation prediction utilities are available. The user can often determine how many hops were employed and from which layers the reflections (or refractions, to be precise) likely occurred. Fewer hops are normally better than more, but high-angle waves travel through less of the ionosphere during the process of refraction, so the path loss per reflection in the ionosphere is often lower for a high-angle wave. Under normal conditions, the ground reflection loss is large enough to favor takeoff angles leading to fewer hops, but more hops can be better over reflective sea water, especially if the ionosphere is disturbed (as indicated by higher A- and K-values).

Another factor in determining which takeoff angle will result in the strongest received signal is the relationship between the operating frequency and the MUF along the path between the transmitter and the receiver. The MUF varies with geographical location. For sites located 2500 mi or less from each other, the simplest mode is onehop. For the ionosphere to support single-hop propagation between those two sites, the MUF at the midpoint of the path, 1250 mi or less from the transmit site, must be greater than the operating frequency. The MUF at either end of the path is not the parameter of interest for single-hop propagation, since the

only region where the radio wave is in contact with the ionosphere is near the midpoint of the path. As a rule, the MUF is highest where average temperatures are highest—namely, in the equatorial regions—so higher-frequency communication is often possible on transequatorial paths when polar paths and midlatitude east-west paths are nonexistent.

For distances greater than 2500 mi, or for those occasions when the height of the F layer is lower than normal, multihop propagation will normally be required. For the two-hop path of Fig. 4.8, there are two MUFs that control whether the transmitted signal is heard at the receiving site. Both MUFs must be greater than the operating frequency; if either one drops below it, the path will fail. If the sketch in Fig. 4.8 represents a transatlantic 20-m path in the northern hemisphere, most likely the MUF on the right side of the page (closer to Europe) will fail first, since the sun sets over the eastern Atlantic at least two or three hours before it does so on the western side.

As additional propagation modes, employing increasing numbers of hops, are analyzed, increasing numbers of MUF locations (often called MUF control points) must be examined. These additional modes require predicted MUFs for a series of locations between the transmitter and the receiver. The specifics of the transmitting and receiving antenna installations play a role in determining how many hops may provide the best signal over a given path. For ionospheric propagation paths (i.e., frequencies up through the lower VHF range), both the antenna itself and its height above local ground contribute to determining the relative amounts of energy radiated at each possible takeoff angle or received at each possible arrival angle. An antenna such as a ground-mounted vertical monopole has good low-angle radiation and minimal high-angle radiation. Although an infinite number of simultaneous propagation modes may exist between the two sites, the relatively strong low-angle radiation from the vertical will tend to make modes with the minimum number of hops the most probable paths to success. In contrast, a horizontal dipole suspended $\lambda/8$ above ground will send most of its radiated energy nearly straight up, thus favoring

multihop modes employing relatively high takeoff and arrival angles.

Multihop modes are generally weaker and more subject to fading than a single-hop path. There is no one best propagation mode or antenna radiation pattern. Some high-frequency bands (including the 11-m, or 27-MHz, citizens band) exhibit predominantly high-angle "short skip" during the summer and longer skip during the other months. When using higher angles of radiation at the higher frequencies, there will be a critical angle wave and escape angle waves that are not returned to the earth. These combinations are useless for terrestrial broadcasting or two-way communications.

The ideal angle of radiation comes from the accumulated knowledge of an inquisitive band of dedicated DXers and contest participants in the amateur fraternity. Using highly directive antennas at multiple heights in combination with sophisticated switching systems in their radio rooms to allow rapid switching between antennas and combinations of antennas, these groups have compiled databases that make it clear no single angle of radiation is best. Even on relatively short DX paths between, say, Europe and the east coast of North America, the optimum radiation angle varies throughout the day. Around sunrise and sunset it is not uncommon to hear stations from the other continent with very high antennas (on the order of 2λ or higher) that clearly have superior low-angle performance. By midday, the best antennas often become the low ones (typically

$\lambda/2$

to λ in height), with their superior pattern response at much higher takeoff angles.

If long-haul communication is not the objective, different choices in antennas and/or their heights may result. A resident of Virginia who wants to communicate with stations in the Carolinas or New York might select a higher angle of radiation with the help of Fig. 4.9 so that his or her signal has the best chance of being heard in those regions. On 80 m, that preference might cause the Virginian to eschew a vertical in favor of a low dipole.

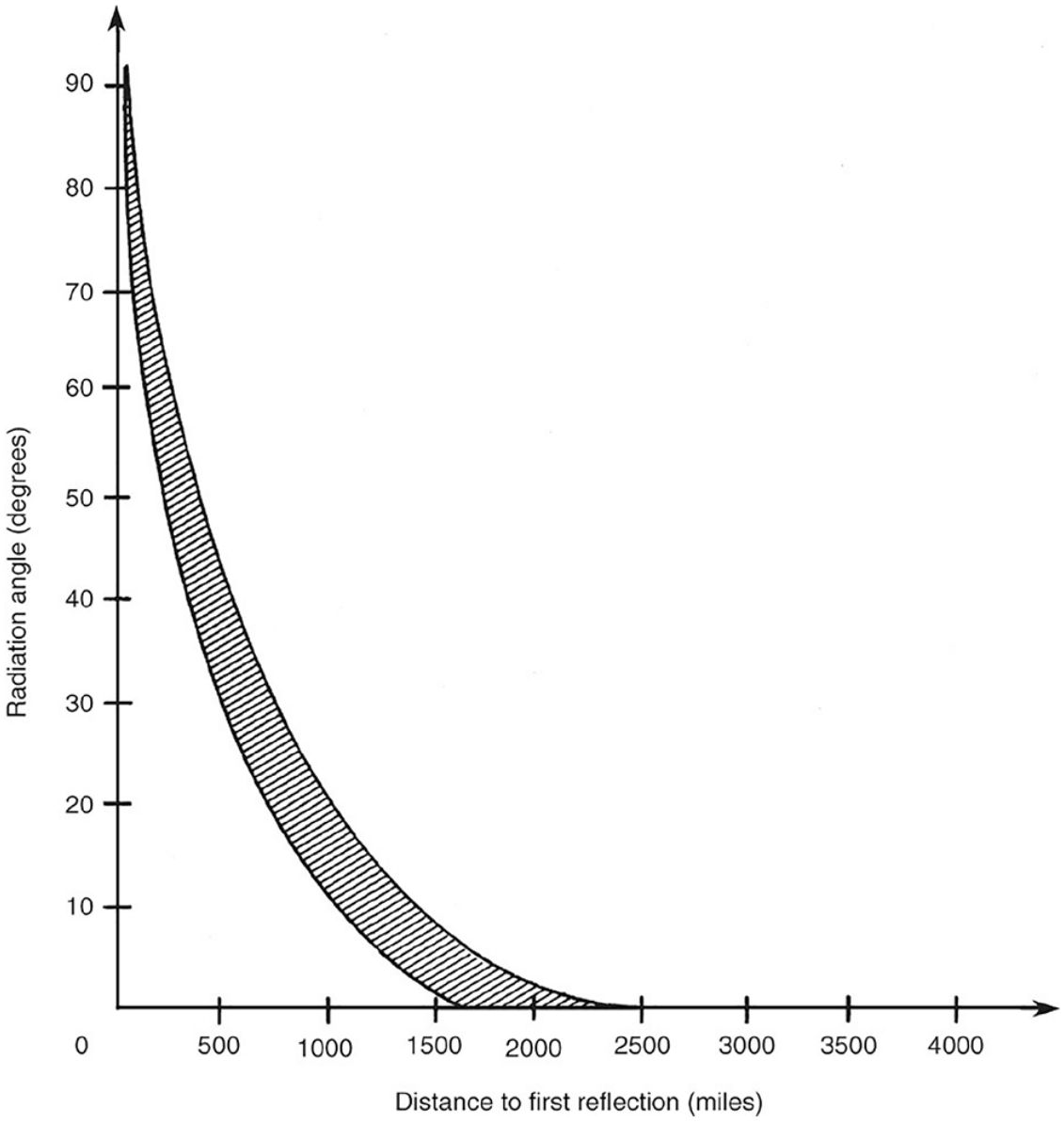


FIGURE 4.9 Takeoff angle versus skip distance.

4.6.2 Geography and the MUF

As Figs. 4.7 and 4.8 make clear, the distance covered on each hop by a transmitted signal is a function of the angle of radiation, often called the takeoff angle. Figure 4.9 shows the angle of radiation relative to the horizon or ground of the antenna versus distance to the first skip zone. As the angle of radiation is increased, shorter distances along the earth's surface are traversed with each hop. At

an angle of about 30 degrees, for example, the typical distance per hop supported by the F layer is about 500 mi.

Since the ionosphere is found at different heights throughout the 24-hour day/night cycle and throughout the course of a year, the distance to first reflection for a given radiation angle varies. This is shown as a shaded zone in Fig. 4.9. As a general guideline, in the absence of disturbed ionospheric conditions you can expect to attain between 1500 and 2500 statute miles per hop in the HF bands at low angles of radiation. For a radiation angle that is only a degree or two above the horizon the maximum skip distance is determined by the geometrical equations associated with the curvature of the earth's surface.

At distances greater than those shown in Fig. 4.9, a signal makes multiple hops. With an antenna having decent low-angle radiation and the ionosphere at an altitude capable of supporting one-hop skip communication out to 2500 mi, covering a distance of 7500 mi requires a minimum of three hops. Depending on the instantaneous characteristics of the ionosphere at the distant reflection points and the relative gain of the transmitting and receiving antennas at different vertical angles, a greater number of shorter hops—involving a higher takeoff angle—may provide a stronger signal 7500 mi away than the three-hop path.

For communicating over short distances (a few hundred miles, say) at MF and HF, we are concerned with the characteristics of the ionosphere in the vicinity of the transmitting station. As the desired distances increase, transmission range depends on the characteristics of the ionosphere at all locations where the transmitted signal will come in contact with it. At a minimum, we are interested in the MUF at the midpoint of each hop on the path between the transmitting and receiving sites.

For the 7500-mi path example, we will normally want to be sure the MUF is above our chosen transmitting frequency in the vicinity of at least three places spaced 2500 mi apart from each other, each corresponding to the midpoint of one of the three hops minimally required to establish communications over that particular path.

Just as you can't determine the weather in New York City by sticking your head out the window in Los Angeles, you can't draw any conclusions about your ability to communicate with an amateur in China by the signal strength of an amateur in a neighboring state. As discussed earlier, the MUF at each specific location is a function of many environmental factors. Here are two examples:

- MUFs are greatest near the equator and decline with increasing latitude, both north or south. Noon hour one-hop transequatorial communications having the midpoint of the hop close to the equator will generally be possible on higher frequencies than can be supported by paths elsewhere on the globe. Although multihop links are possible, the typically lower MUFs at the midpoints of the outer hops will force use of a lower frequency.
- MUFs usually follow the sun, peaking shortly after noon. To maximize the likelihood of being able to communicate on a typical midlatitude east/west path, such as between Maine and Oregon or between California and Hawaii, choose midafternoon time at the eastern end of the circuit and mid- to late-morning at the western end. For best signal-to-noise ratios, pick the highest open operating frequency.

For nighttime bands, you typically want to select a time when the MUFs at the hop midpoints or control points are as low as possible, yet still above the desired operating frequency. This minimizes signal loss from D-layer absorption.

4.6.3 Great Circle Paths

The shortest path around the globe is a great circle. Looking at a map, one might think that a jet plane would travel due east from Washington, DC to Lisbon, Portugal. In truth, if you head due east from Washington, DC, across the Atlantic, the first landfall will be West Africa, somewhere near Zaire or Angola, far south of Portugal. Similarly, if we operate a Portuguese language shortwave broadcast

station near Washington, DC, and want to maximize the signal we deliver to listeners in Lisbon, our conventional wall map might cause us to beam our signal due east when the correct direction is northeast. The geometry of spheres governs both air travel and radio waves.

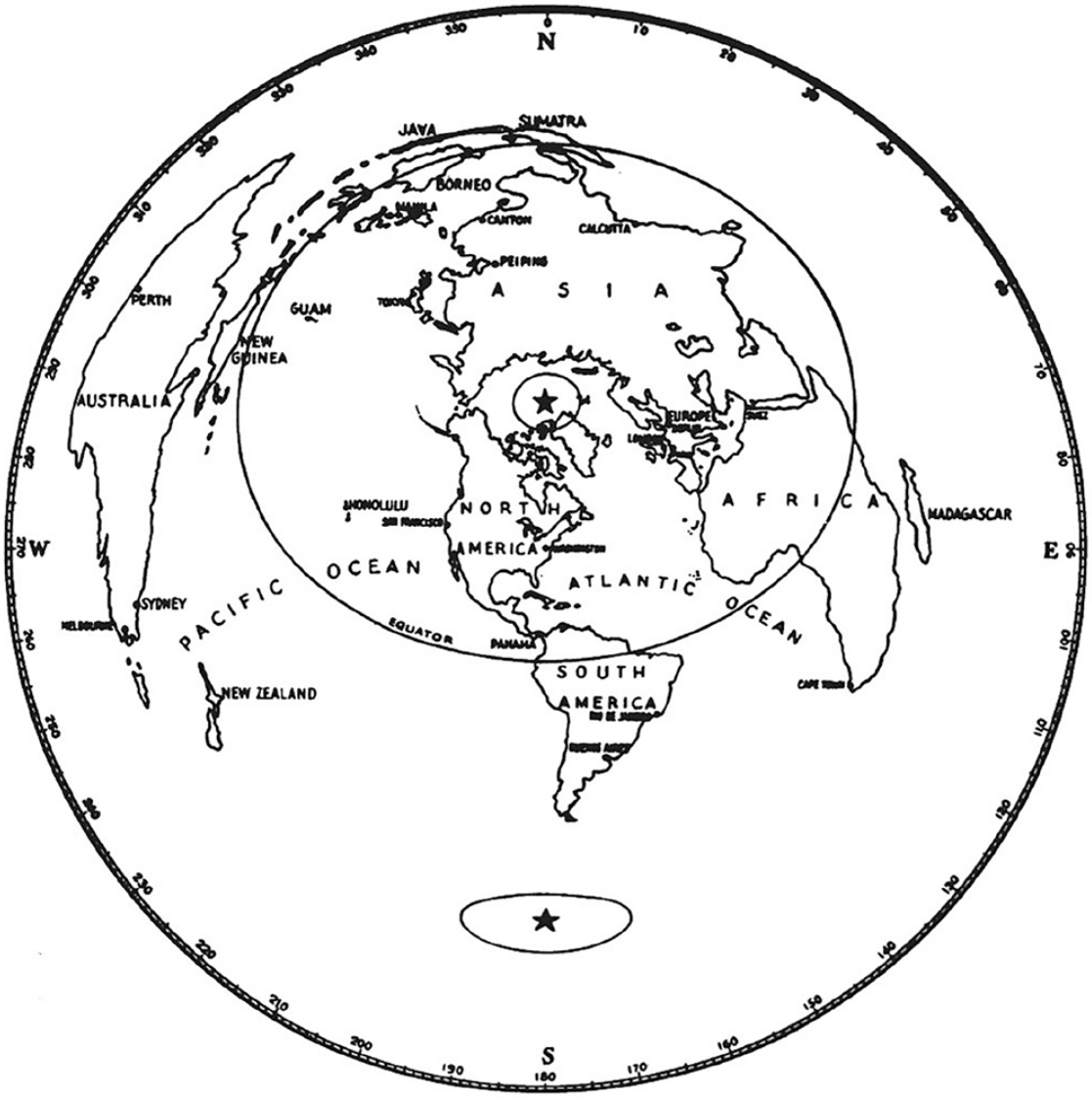
The problem of accurately representing the surface of a sphere on a flat two-dimensional surface such as a sheet of paper has confounded cartographers for centuries. There is no one projection that is best for all possible uses. For long-distance terrestrial radio communication, the most important map projection of our earth is the azimuthal equidistant projection, or, more simply, the great circle map.

Great circles are circumferences of a sphere. Only one great circle can normally be drawn between two points on a sphere. Both a short path and a long path connect those two points. There is one special case where this breaks down. If the two points are halfway around the globe from each other, they are called antipodes of each other, and the short path is identical in length to the long path. Antipodes are connected by an infinite number of great circles.

On a great circle map, any straight-line radius from the center of the map (usually the user's location) to its periphery corresponds to a great circle on the earth. Thus, by tracing the single radius line that runs between the center of the map and any other location in the world, the user is assured of following the shortest possible path to that location. For travel, one clearly selects the shorter of the two great circle segments connecting the two points of interest, but that is not always the case for radio communications.

Older antenna manuals published by the ARRL included a family of great circle maps centered on selected population centers around the world. Now, software for these maps is readily available, either stand-alone or as part of a larger log-keeping and station control package. For radio amateurs, the best centers for great circle maps are usually their own transmitting and receiving sites. Maps having their centers a short distance away are useable except for the most precise measurements and for very short distances

from the center. Figure 4.10 is a great circle map drawn on Washington, DC. For HF DXing, it is adequate for users located up and down the east coast of the United States.



F IGURE 4.10 Azimuthal map centered on Washington, DC. (*Courtesy of The ARRL Antenna Book*.)

Later in this book we take up directional antennas—antennas deliberately designed to maximize radiation in one direction or a small range of directions in exchange for accepting reduced radiation in other directions. With the exception of broadcast

stations and point-to-point stations, directional antennas are rotated either electronically or mechanically. Some highly directive MF and HF antennas for transmitting (the rhombic) or receiving (the Beverage) are too large to be rotated. Instead, one or more of each is constructed with its direction of maximum signal fixed on a desired region of our globe. An amateur low-band DXer in Maryland might choose to install a Beverage receiving antenna aimed at Europe. By referring to the great circle map of Fig. 4.10, he would conclude (correctly) that his Beverage should be oriented so as to provide maximum pickup of incoming signals centered on a true compass bearing of somewhere between 45 and 50 degrees. An amateur in Japan (roughly the same latitude as Washington, DC, and Lisbon), on the other hand, would not aim due east or west to favor Europe on his receiving antenna; instead, he would first consult a great circle map centered on Tokyo and then aim the antenna 25 to 30 degrees west of north.

4.6.4 Ionospheric Fading and Diversity

Due to fluctuations in the ionosphere, signals vary in strength and phase at the receiver site. This is referred to as ionospheric fading.

An audibly dramatic form of fading, selective fading, is often heard with analog voice transmissions on the MF and lower HF bands at night. At distances beyond the local ground-wave reception zone, the received signal often consists of ground wave, sky wave, and multiple reflections of the sky wave from the ionosphere. Due to the un-ceasing motion of the ionosphere, the received signal is an ever-changing mix of the transmitted signal arriving by multiple paths, each having its own time-varying amplitude and phase relative to original signal. The most distinctive aspect of this fading is caused by the fact that the ionosphere is dispersive and its reflection characteristics are different for even slightly different frequencies. An analog HF voice transmission consists of frequency components spread over a 3- to 20-kHz range around the carrier frequency. When the multipaths combine at the receiver, some frequency components add constructively and others

interfere destructively. The received signal is an eerie distortion of the original, similar to the flanger feature of sound mixers. It is such a distinctive sound that it was replicated in a music single released in 1959—"The Big Hurt" by Toni Fisher—which went all the way to number three on Billboard magazine's top 100.

Fading from any cause is a serious problem that must be overcome to achieve reliable communications. This can be done by sending the signal in multiple forms and selecting the one with best signal level. This is known as diversity. Three forms of diversity are popular: frequency diversity, spatial diversity, and polarization diversity.

In the frequency diversity system of Fig. 4.11, the transmitter delivers RF with identical modulating information to two or more frequencies simultaneously. Because the two frequencies will fade differently, one will generally be stronger.

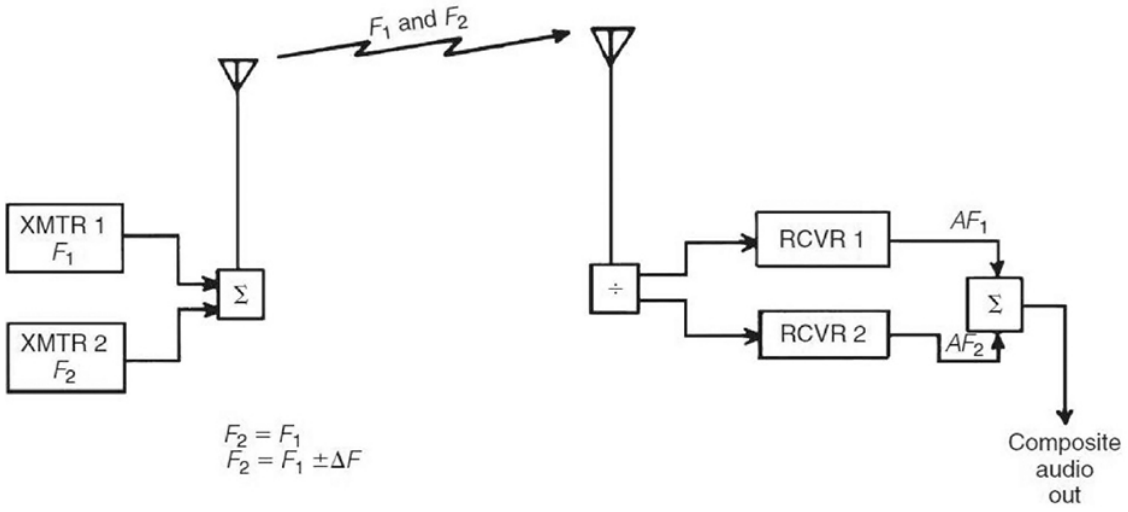


FIGURE 4.11 Frequency diversity reduces fading.

The spatial diversity system of Fig. 4.12 employs a single transmitter frequency but two or more receiving antennas are used, typically spaced one-half wavelength apart in the direction of the transmitter. Path lengths are different from the transmitter to each receiving antenna, so if the signal is in a fade at one antenna it may be stronger at another. Analog audio mixing or digital

combining techniques based on the relative strengths of the signals keeps the audio output relatively constant while the RF signal at any one antenna fades in and out.

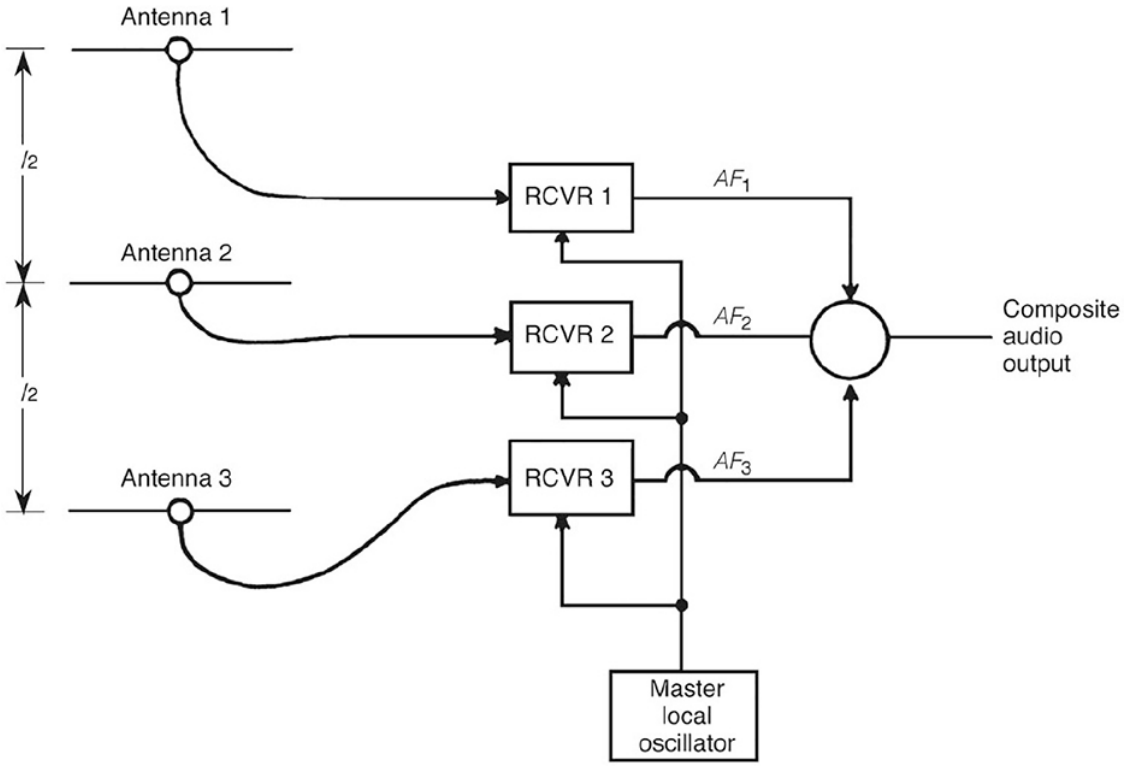


FIGURE 4.12 Spatial diversity.

Polarization diversity (Fig. 4.13) uses vertically and horizontally polarized antennas to enhance reception. This form of diversity is based on the fact that the ionosphere shifts the polarization of the transmitted signal as it refracts to the earth. As in the space diversity system, the outputs of the vertical and horizontal receivers are combined to produce a constant level output.

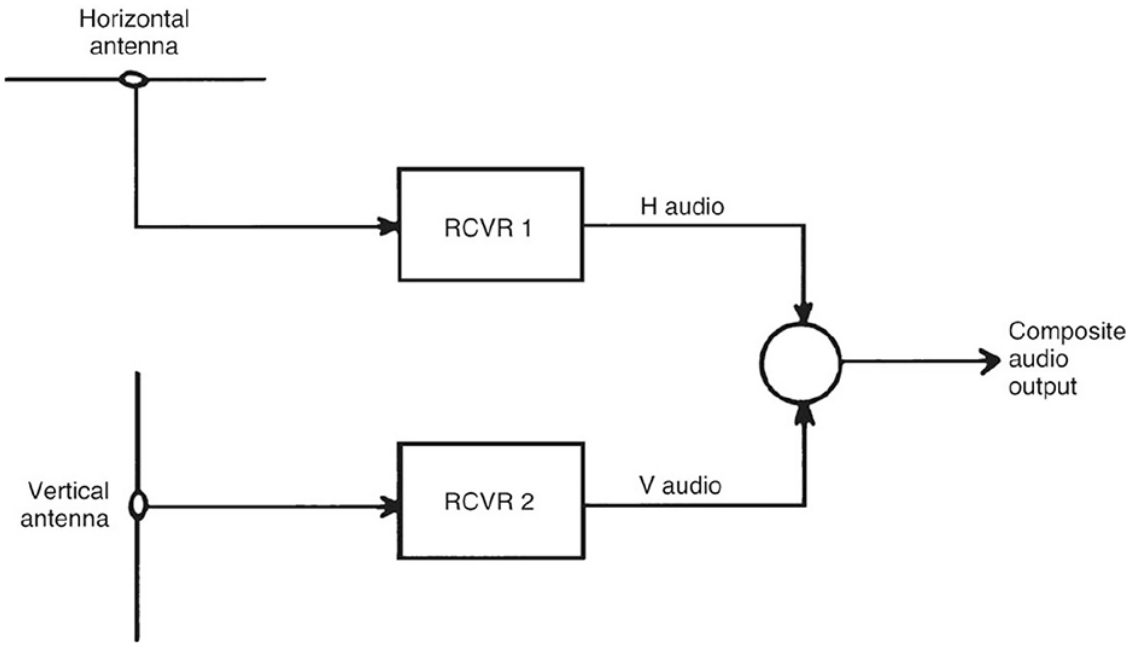


FIGURE 4.13 Polarization diversity.

4.7 Using the Ionosphere

Refraction of MF, HF, and some lower VHF radio signals back to the earth via the ionosphere is the backbone of long-distance terrestrial radio communications. There is a complex interplay between the three major ionospheric layers and their individual reactions to bombardment by the sun. Path loss considerations, absorption, and the ever-changing position of the earth with respect to the sun make determining the best frequencies and times for communicating challenging. To help readers more easily select a frequency or a time of day for best communications success, here are some guidelines and rules of thumb:

- Because the MUF responds to the total level of radiation impinging on the ionosphere, it follows the sun. The higher the desired operating frequency, the greater the solar radiation levels needed to support ionospheric propagation. As a result, the 6-m amateur band at 50 to 54 MHz only rarely opens for F-layer DX propagation during the course of an average sunspot cycle. The 10-

and 12-m amateur bands may lie dormant for a few years of each sunspot minimum and go through periods of spotty openings on either side of the maximum. The same will be true to a lesser degree with the 15- and 17-m bands. All these bands are considered to be daylight bands, even though signals arriving by other propagation modes are sometimes heard on them during dark hours. Because the critical frequency, f_C, is below many or all of these bands for much of a typical sunspot cycle, skip zones are almost always observed, although an outbreak of sporadic E propagation may occasionally make it appear that 20 m has no skip zone.

- Over a cycle of sunspot activity, 20 m will most likely swing between being open for just a few hours each midday near sunspot minimum and bustling with activity around the clock near the maximum. Part of the popularity enjoyed by the 20-m band is that it is seldom completely "dead"; signals are almost always evident on the band, regardless of when activity is randomly checked.
- Frequencies below about 10 MHz are considered to be nighttime bands. High daytime absorption by the D layer is the primary reason they suffer a poor reputation when the sun is up, but this is largely a phenomenon of extended periods of extremely high sunspot counts such as was experienced in the 1957–61 time frame. Moderate daytime absorption can be overcome or ameliorated by a combination of higher transmitter power and higher gain antennas at both ends of the circuit. Throughout most of an average sunspot cycle, both 80 and 160 m provide reasonable daytime propagation out to 1000 mi or so, and both expand their reach during the hours extending from shortly before sunset to shortly after sunrise to encompass worldwide DXing capability. During sunspot minima, f_C often drops

below 3.5 MHz during the evening hours, with 80 m becoming useless for short skip contacts. On occasion, f_C has been known to drop below 1.8 MHz and even into the AM broadcast band.

- Forty meters is a transition band. During periods of high sunspot activity, f_C rises above 7.5 MHz and the band serves as a great short- and medium-haul communications medium for much of each 24-hour period. During these periods, there are many daylight hours each day when there will be no skip zone on 40. During sunspot minima, however, 40 is much more like 20 m—exhibiting a very pronounced skip zone (even at midday), providing long-haul DX communications during daylight hours and shutting down completely at night as the MUF routinely drops below 7 MHz.

Because the characteristics of the MF and HF bands depend on radiation from the sun, choosing a band requires thought about the position of the sun with respect to the great circle path and the station at the other end. Paths that go through roughly equal segments of darkness and daylight normally are the most difficult to use, especially during periods of low sunspot activity. When sunspot activity is low, attempts by North American amateurs to contact DXpeditions to the Himalayas, Afghanistan, and other remote central Asian regions have often failed. In contrast, when the sun is active the MUF (even in the far northern or southern latitudes) may stay above 14 MHz long enough that 20 m may well be open to nearly all parts of the world simultaneously for at least a few hours out of each 24-hour period.

4.7.1 Skew Paths

Because the energy absorbed from the sun by the ionosphere is greatest near the equator and during the midday hours, MUF values at any given time vary substantially from point to point on the globe. In addition to causing the higher HF bands to be open only

during daylight hours, sometimes equatorial MUFs are high enough to allow north/-south propagation across the equator on a specific band while midlatitude or polar MUFs fail to support east/west or polar paths. If the 10-m band appears to be open between the United States and Chile or Argentina some noontime, communications from the United States to Europe or Asia will usually require dropping down a band or two to 12 m or 15 m, where midlatitude and polar MUFs are high enough to support communications along those paths.

An alternative for east/west paths when the MUF is not high enough at midlatitudes or in the polar regions is to aim your antenna at a point on the globe that is open to you and that you have reason to believe is also open to the person or country you're trying to contact. Many times when the 10-m band is open only for north/south and transequatorial paths, it is still possible for stations in the United States and Europe to establish contact by aiming directional antennas at a region of the Atlantic Ocean near the equator and midway between their respective longitudes. In this situation, the normal rules of aiming along the shorter great circle path do not apply, since the MUF in that direction is not high enough to support midlatitude paths. Instead, each station points its antenna at an angle to the other. This propagation mode is called side scatter.

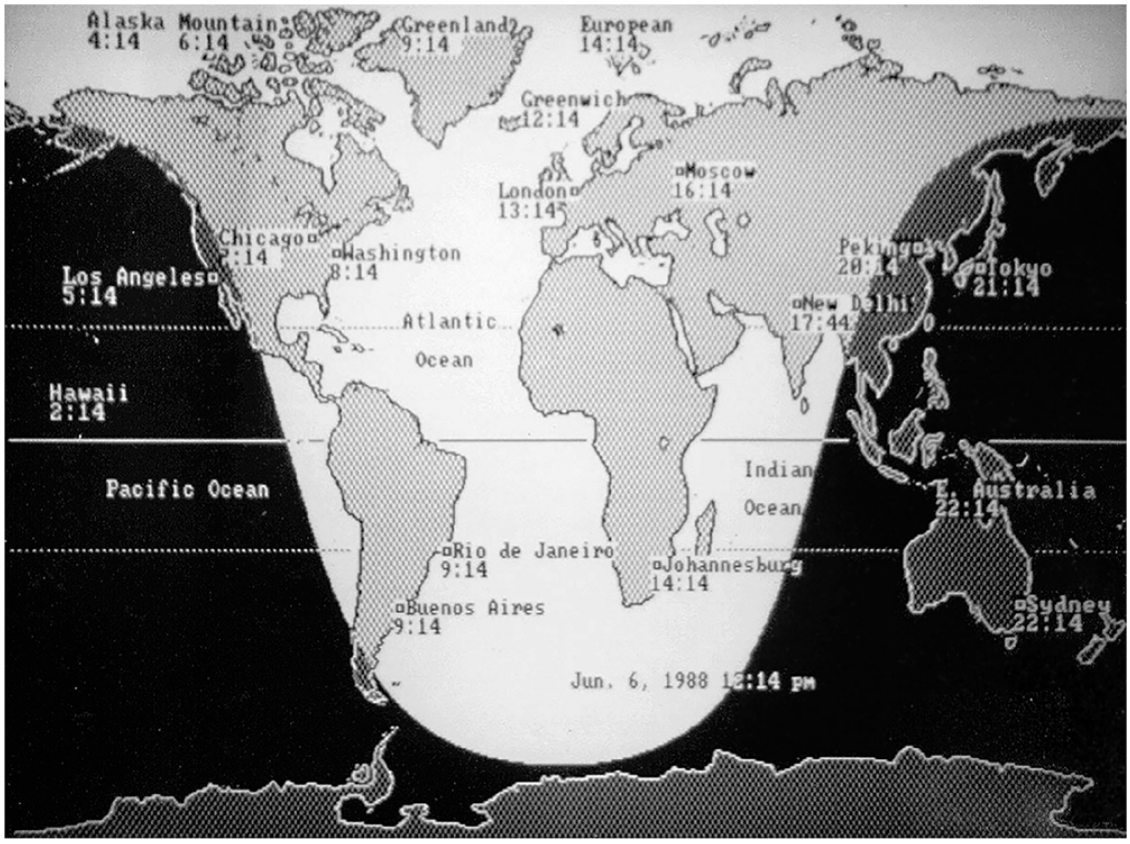
One way to determine exactly where to aim your antenna is to find a clear frequency, run full break-in ("QSK") with one of today's transceivers, and transmit (call "CQ DX," for example) while rotating your antenna and listening for backscatter from your transmitted signal. When you are aimed at the right area you'll be able to hear the last few milliseconds of your own transmissions (especially if it's international Morse code, generally referred to as continuous wave, or "CW"). By making small adjustments to the position of your antenna you'll be able to zero in on the best beam heading. Then you hope that your faraway friends are doing the same thing.

If you were successful in working distant stations in this fashion, you would tell others that you worked Europe on 10 m, for instance, via a skew path. Over the years, skew paths have helped

many DXers contact countries they might never reach otherwise. One of the most productive skew paths is obtained by aiming roughly due west from the northeastern United States and eastern Canadian provinces when trying to establish contact with stations in Japan, Korea, the Philippines, Southeast Asia, and the easternmost regions of Asiatic Russia—especially on 160 and 80 m. The reason for this can be seen by looking at the great circle map in Fig. 4.9. The great circle short path between the eastern part of the North American continent and the Far East goes right through the middle of the north auroral circle that surrounds the magnetic north pole (which wanders around Alaska and northwestern Canada). A perturbation of the earth's magnetic field increases signal absorption in the auroral regions and makes an otherwise difficult path almost impossible. In this example, the skew path is used to circumvent a region of high absorption rather than finding a way around an inadequate MUF.

4.7.2 Gray Line Propagation

The gray line is the twilight zone between the nighttime and daytime halves of the earth. This zone is also called the planetary terminator (Fig. 4.14). It tilts up to 23 degrees on either side of the north-south longitudinal lines during the course of a year, depending on the season. During the vernal and autumnal equinoxes it runs directly north/south, and it reaches its most pronounced tilts on the longest and shortest days of the year.



F IGURE 4.14 The planetary terminator ("gray line") provides some unusual propagation effects. (*Courtesy of MFJ Enterprises*.)

The gray line is of interest because signal levels between stations located on it are significantly enhanced—sometimes as much as 20 dB. As the gray moves with the sun, the effect is temporary but it can make the difference between a contact and no contact with a rarely heard station in a remote part of the world. Because of the gray line, many serious DXers make sure they are carefully monitoring the HF bands during the twin twilight periods of their local dawn and dusk. It is not unusual for a gray line opening between two specific locations to be only 10 or 15 minutes in duration.

Because the tilt of the terminator changes over the course of a year, the paths that can benefit from it change as well. In Fig. 4.14, which is a gray line projection for early June, a station near

Vancouver, British Columbia, Canada, might observe enhanced signal levels from stations in Bangladesh, Myanmar, or interior China. Three months earlier, in early March, that same Vancouver station might be more apt to hear surprisingly strong signals from Turkey, Saudi Arabia, and the horn of Africa.

4.7.3 Long Path

The great circle connecting any two points on the earth has two segments; for radio communications purposes we refer to them as the long path (major arc) and the short path (minor arc). If both paths are open, signal propagation between the two points is usually more efficient along the short path. But when the short path is not open the long path can sometimes deliver an adequate signal.

Figure 4.15 is a great circle map centered on New York, with night and day for 0200 Zulu (GMT) on April 1 shaded in. A 15-m transmitter in the northeastern United States would normally use a short path with beam heading 80 degrees from New York to Madagascar. When prolonged darkness over the Atlantic has caused the short-path MUFs between the two locations to fall below 21 MHz, there is a chance that the New York operator might be able to make contact by turning the directional antenna 180 degrees to compass heading 260 degrees and using the long path. The nearest important MUFs to the two stations on long paths are about 1250 mi or so from each end of the path. The first MUF of importance is about 1250 mi west of New York, in the southwestern United States, where the sun has just set and the MUF there may still be above 21 MHz.

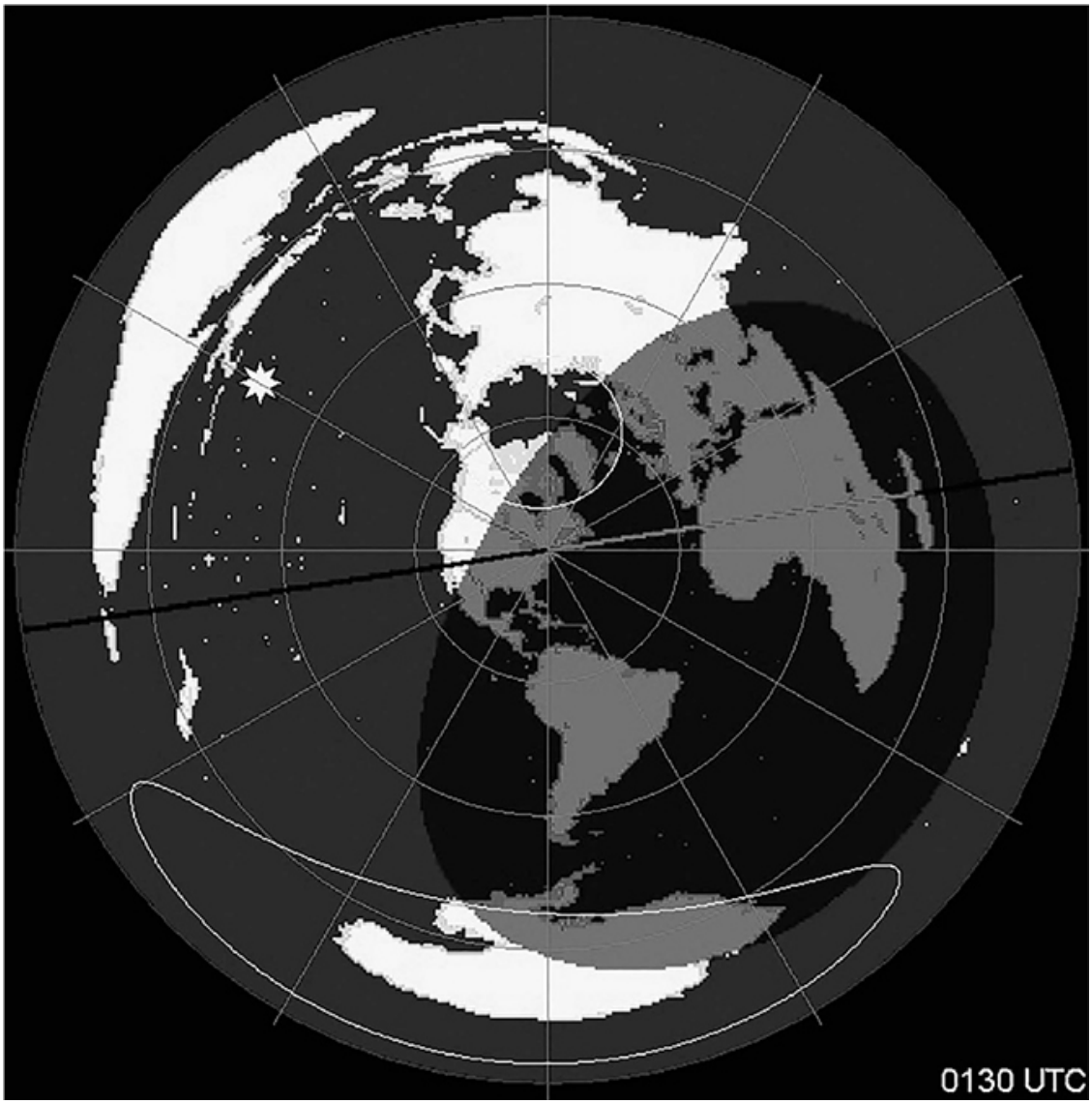


FIGURE 4.15 0200Z April 1 great circle map showing long and short paths between northeastern United States and Madagascar.

Similarly, the other end of the link has daylight an equal distance east of the station, directly opposite the short-path bearing to New York. The entire path between the first MUF control point and the US operator's control point is in daylight. Even though peak long-path signal strengths will likely be lower than short-path strengths, there is still a good chance that the operators will be able to make contact the long way around. By switching to long path, we have

put the path back into daylight where it counts—at the first MUF control point closest to each side of the link and beyond.

The opposite approach works well on 40 and 80 m between the northeastern United States and Japan during a winter afternoon on the US east coast. The objective is to find MUF control points that are in darkness, even though either or both of the stations involved may be in daylight at the time of the contact. For a DXer in the northeastern part of the United States whose short-path propagation to the Asiatic Pacific region is often blocked by unsettled polar conditions, there are few HF operating experiences that compare with aiming the 40-m beam southeast on a winter afternoon and hearing an operator in the Far East answer a "CQ." It's magic.

4.7.4 Ionospheric Scatter Propagation

Clouds of ions in the atmosphere can lead to ionospheric scatter propagation. These clouds can exist in both the ionosphere and the troposphere, although the tropospheric model is more reliable for communications. Figure 4.16 shows the mechanism for scatter propagation. Radio signals from the transmitter are reflected from the cloud of ions to a receiver location that otherwise might not receive them.

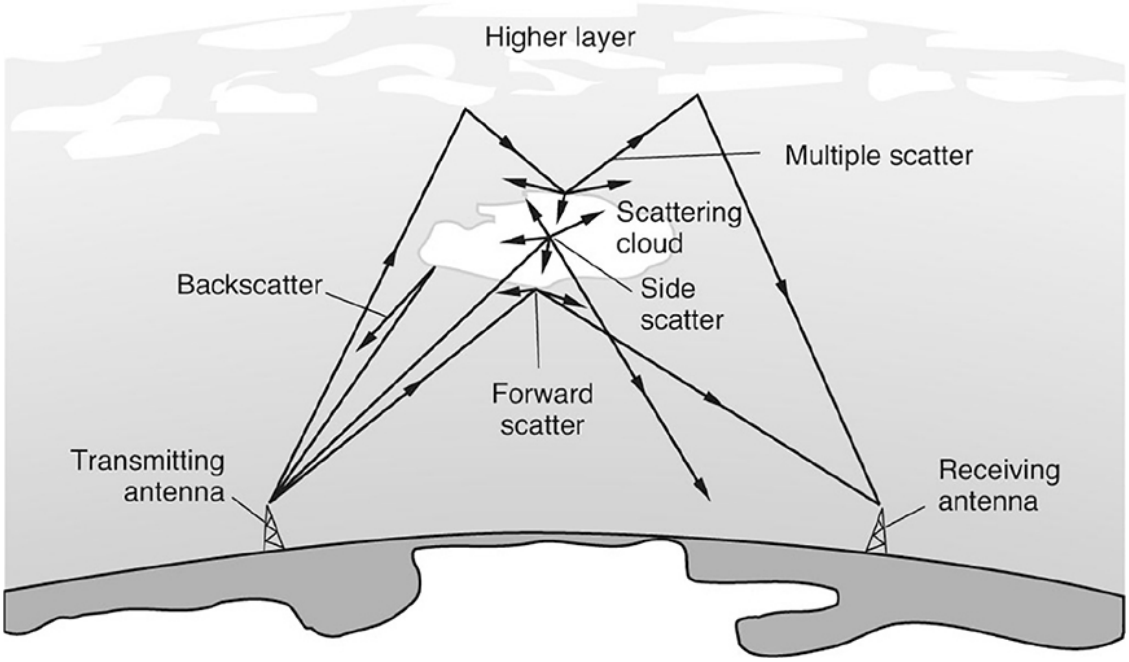


FIGURE 4.16 Various modes of scatter propagation.

There are at least three different modes of scatter from ionized clouds: backscatter, side scatter, and forward scatter. The backscatter mode is a bit like radar, in that the signal is returned to the transmitter site, or to regions close to the transmitter. Forward scatter occurs when the reflected signal continues in the same azimuthal direction (with respect to the transmitter), but is redirected toward the earth's surface. Side scatter is similar to forward scatter, but the azimuthal direction usually changes. Unfortunately, there are often multiple reflections from the ionized cloud (shown as "multiple scatter" in Fig. 4.16). When these reflections are able to reach the receiving site, the result is rapid and potentially deep fading.

4.7.5 Auroral Propagation

Bursts of particles released from the sun 18 to 48 hours earlier causes a luminescence in the upper atmosphere called aurora. The ionized regions of the atmosphere that create the lights form a reflective curtain at radio frequencies, especially at VHF and above (al-though auroral propagation is infrequently observed on

frequencies as low as 14 MHz). Auroral effects are normally seen at higher latitudes, although auroras have been seen as far south as Mexico. At those times, listeners in the southern tier of states in the United States are often treated to the reception of signals from the north being reflected from auroral clouds. Reflection off the aurora curtain allows radio communication out to 1500 mi or so. Numerous websites provide aurora alerting services based on the expected time delay between visual observance of a flare on the sun's surface and the subsequent disruption of the earth's magnetosphere by arriving solar particles.

4.7.6 Meteor Scatter Propagation

Scatter propagation has been exploited mostly at VHF from meteor trails. When meteors enter the earth's atmosphere, they leave transient paths of ionized particles. The ions act as a radio mirror that permits short bursts of reception—especially high-speed CW—between sites correctly situated. Meteor scatter reception is sporadic by nature, but it occasionally makes communication possible over extended paths that are not normally available.

4.8 Propagation Predictions

Terrestrial and satellite-borne sensors have increased our ability to monitor solar and atmospheric conditions. Sunspots and solar activity are monitored. Auroras can be predicted. We can follow weather fronts as they track across our continent and identify patterns likely to spawn clouds of sporadic E-skip in the next few hours. Even with this data, ionospheric propagation can be difficult to predict, but significant gains have been made in the last few decades.

Many software packages and tools are available for ionospheric signal propagation prediction:

W6ELProp predicts ionospheric (sky-wave) propagation between any two locations on the earth on frequencies between 3 and 30

MHz, focusing on conventional F-layer propagation.

HamCap is freeware offered by Alex Shovkoplyas, VE3NEA. It relies on a propagation prediction program called *VOACAP*, originally developed by NTIA/ITS for the Voice of America.

VE3NEA also offers *IonoProbe*, which monitors and displays a number of valuable space weather parameters on a near-real-time basis.

Other software includes *CAPMan* and its successor, *WinCAP*, *DXProp*, and *HF-Prop*. VHF/UHF specific packages include *Radio Mobile Deluxe, Ground Wave Prediction System*, and *RFProp*.

Software is also available to determine the effect of terrain on radio propagation. Examples include:

HFTA (High Frequency Terrain Assessment) is a program designed by R. Dean Straw, N6BV. Based on terrain contours, HFTA assesses the terrain that may stand between the user's antenna site and a clear shot at the ionosphere at a variety of takeoff and arrival angles, examining both reflection and diffraction modes of propagation.

4.9 The Ionosphere and the Magic of Radio

Because so many of the known modes of propagation depend on variable environmental factors, our ability to communicate—especially via the ionosphere—is statistical. Antenna improvements, understanding of subtle propagation modes, and careful analysis can help, but ultimately we are at the mercy of Mother Nature, who is not above throwing a curveball every now and then.

Yet, for many of us, it is this statistical uncertainty that is at the very heart of the joy we find in communication by radio. The unique sound of a weak signal fading in and out of the band noise is part of the wonder of radio. And there is a nearly infinite variety of possible band conditions awaiting us each time we sit down at our receiver and put on the headphones. The unpredictability of

ionospheric HF propagation is without a doubt one of the lifelong draws of radio communications for tens of thousands. It’s magic!

CHAPTER 5
Antenna Basics

An antenna converts energy from one form to another. A radiating antenna converts time-varying electrical currents that are confined and guided within a circuit or transmission line to a radiated electromagnetic wave varying at the same rate and propagating outward through space. It is not unreasonable to visualize this process as decoupling the electromagnetic fields created by the time-varying currents from the antenna structure and launching them into space as a wave. As we shall see in future chapters, the efficiency with which an antenna performs this conversion is one of the key measures of its goodness. Devices that accomplish this conversion efficiently share special physical and geometrical attributes.

In this chapter, we will summarize enough basic electronics and electromagnetic theory to provide a feeling for the physics of antennas. An in-depth knowledge of antenna theory requires a command of various advanced mathematics techniques, including calculus, coupled with completion of specialized antenna courses usually found in upper division electrical engineering or physics courses, but our treatment will be at a more intuitive level and accessible with an understanding of a few principles of electrical systems.

5.1 Circuit Fundamentals

In this chapter and throughout this book, we will assume and use some fundamentals of physics and electrical engineering as summarized in this section.

5.1.1 Superposition Principle

Unless otherwise noted, all components, circuits, systems, antennas, transmission lines, and propagation media discussed in this book are linear in their operation. If an input voltage of 1 V applied to a device results in an output of 1 A (ampere) elsewhere in the system, an input of 5.32 V will result in an output of 5.32 A. For linear systems, the output is directly and exactly proportional to input. If we apply the sum of two signals to the input, the output from the device will be the sum, or superposition, of the outputs that would be obtained individually from the two input signals. Example: We apply a 1-V signal oscillating at a frequency of 3 MHz and a 4-V signal oscillating at 3.1 MHz to the original input of the aforementioned device. At the output we measure a 5.32-V 3-MHz signal and a 21.28-V 3.1-MHz signal.

5.1.2 Electric Charge

The smallest unit of electric charge is the charge of a single electron. The total charge in a localized region is the sum of all the net, or excess, electrons in that region. Like charges repel [each other]. If two electrons are forcibly brought near each other and then released, they will separate and get as far from each other as the space will allow. If we add free electrons to a solid metal ball, they will all move quickly to spread themselves uniformly around the surface of the ball.

5.1.3 Current

Current represents the flow of moving charges. During his experiments with electricity, Benjamin Franklin arbitrarily defined the "plus" and "minus" terminals of batteries and "positive" and "negative" current flow in wires. Other researchers determined that the primary conduction mechanism in metals consisted of electrons being broken away from their parent atoms and moving along the wire, from atom to atom, in response to an applied voltage. Franklin's conventions turned out to be at odds with the electron model of the atom developed a century later. Even though we speak of positive current leaving the positive terminal of a battery

and flowing through a circuit and back into the negative terminal, electrons exit the negative terminal, flow through the external circuit in the opposite direction to the current, and reenter the battery at its positive terminal.

5.1.4 Resistors

A resistor (Fig. 5.1) is a two-terminal circuit element with the idealized characteristic

$$V = IR \quad (5.1)$$

which is equivalent to saying the voltage V (in volts) across the resistor is equal to the current I (in amperes) flowing through the resistor multiplied by the resistance R (in ohms) of the resistor. A useful analogy is the flow of water through a pipe. The water pressure (analogous to voltage) causes the flow of water (current) through pipes (resistance). Resistance is proportional to the total length and inversely proportional to the cross-sectional area. A constant voltage across the resistor terminals will result in a steady current flow through the resistor until the voltage is removed. The power dissipated in a resistor is given by



FIGURE 5.1 Circuit symbol for a resistor.

$$P = VI = \frac{V^2}{R} = I^2 R \quad (5.2)$$

The resistance of an ideal resistor does not depend on frequency, but real resistors become capacitative or inductive and the value of the resistance can change at high frequencies.

5.1.5 Capacitors

A capacitor (Fig. 5.2) is a two-terminal component capable of storing electrical energy in the form of electric charge. The fundamental unit of capacitance is the Farad (F). The total energy stored in a capacitor is
$\frac{1}{2}CV^2$
, where *C* is the capacitance and *V* is the voltage (in volts) across its terminals. The capacitor current-voltage law is

$$I = C\frac{\Delta V}{\Delta t} \quad (5.3)$$

where I (in amperes) is the current through a capacitor of value $C$when the voltage across the capacitor terminals changes by some amount ΔV (in volts) over a short time interval Δt (in seconds). Capacitors usually consist of metallic plates or sheets separated by a thin insulator.

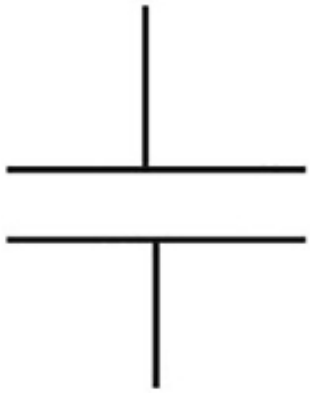


FIGURE 5.2 Circuit symbol for a capacitor.

A capacitor is said to exhibit capacitive reactance. Reactance works much like resistance to impede the flow of current in a circuit but the two are very different:

- An ideal capacitor does not dissipate any power.
- The reactance of a capacitor is frequency dependent.
- The capacitive reactance is

$$X_C = -\frac{1}{2\pi f C} \quad (5.4)$$

where f is the frequency in hertz and C is the capacitance in Farads. Alternatively, f can be in megahertz (MHz) and C in microfarads (μF). The minus sign is important in distinguishing the effect of capacitors from that of inductors and relates to the phase relationship between the voltage across a capacitor and the current through it.

5.1.6 Inductors

An inductor is a two-terminal component capable of storing magnetic energy associated with current flowing through the device (Fig. 5.3). The unit of inductance is the Henry (H). The energy stored in an inductor is
$\frac{1}{2}LI^2$
, where *L* is the inductance of the device and I is the current in amperes through it. The inductor current-voltage law is

$$V = L\frac{\Delta I}{\Delta t} \quad (5.5)$$

where *V* (in volts) is the voltage across the inductor of value *L* when the current through the inductor changes by some amount ΔI over a short time interval Δt. A simple inductor can be made from a length of insulated wire wrapped into a coil. The inductance increases if the coil is wrapped around an iron rod or other ferromagnetic core.



FIGURE 5.3 Circuit symbol for an inductor.

As discussed for capacitors, an ideal inductor does not dissipate power. The current through the inductor and the voltage across it are out of phase. The inductive reactance is

$$X_L = 2\pi f L \quad (5.6)$$

where *f* is the frequency in hertz and *L* is the inductance in henries (or megahertz and microhenries). Inductive reactance X_L is positive, in contrast to capacitive reactance in Eq. (5.4).

5.1.7 Phasors

If we apply a sinusoidal wave of frequency *f* across two terminals or nodes in a circuit consisting of interconnected resistors, inductors, and capacitors, the voltage across and the current through the

circuit is sinusoidal at frequency *f*, but the current is not in phase with the voltage. If the circuit was only resistive, the current and voltage would be in phase. The capacitors and inductors cause phase shifts of varying degrees between currents and voltages in different parts of the circuit.

When dealing with sinusoidal (also called time-harmonic) voltages, currents, and fields, it is convenient to use a complex phasor notation. The complex phasor notation allows us to keep track of phase shifts for voltages and currents as they pass through circuit components with capacitative or inductive reactance. The phasor voltage is defined by

$$V(t) = \mathrm{Re}\left[\widetilde{V} e^{j\omega t}\right] \quad (5.7)$$

where *V*(*t*) is the time-varying voltage,
$\widetilde{V}$
is the complex phasor, and
$\omega = 2\pi f$
is the angular frequency of the sinusoidal voltage in radians per second. Phasor currents and electromagnetic field quantities are defined similarly.

5.1.8 Complex Impedance

We can describe the relationship between the phasor voltage and current at a given frequency in terms of a resistance and reactance combined into a complex impedance given by

$$Z = R + jX \quad (5.8)$$

The formula includes the imaginary unit
$j\left(j = \sqrt{-1}\right)$
to represent reactances. The phasor voltage across a circuit with complex impedance *Z* is related to the current through the circuit by
$V = ZI$

. This is the complex or phasor form of Ohm's law.

The real part of the impedance, *R*, is the resistance of the device and the imaginary part, *X*, is the reactance. Depending on whether the net reactance seen appears to be inductive or capacitive, *X* may be positive or negative. *R* and *X* have units of ohms and are functions of frequency. Their values depend in general on the values of all components in the circuit. Regardless of how the components in a circuit are connected, we can represent the two-terminal impedance with an impedance of the form of Eq. (5.8). The resistance, capacitance, and inductance of any real components or structures depend on frequency. The impedance *Z* for some circuit may be

$50 + j47\ \Omega$

at one frequency and

$35 - j17\ \Omega$

at another frequency.

5.1.9 RLC Circuit

A circuit consisting of an inductor, capacitor, and resistor is an RLC circuit. The impedance of an RLC circuit with the components connected end to end or in series is

$$Z = R + j(X_L + X_C) \quad (5.9)$$

Since the reactance of the capacitor,
X_C
, is a negative number, this amounts to subtracting the magnitude of the capacitive reactance from the inductive reactance. Looking into the circuit, we see the net reactance. Parallel RLC circuits are also used and are sometimes called tank circuits.

5.1.10 Resonance

When the inductive and capacitative reactance in an RLC circuit are equal in magnitude and opposite in sign, the circuit is resonant or in resonance. At the resonant frequency,

$X_C = -X_L$
, and the second term or the imaginary part of Eq. (5.9) is zero. Using the formulas for inductive and capacitative reactance, we find that resonance occurs at the frequency

$$f = \frac{1}{2\pi\sqrt{LC}} \quad (5.10)$$

This is one of the most important formulas of radio and electronics design. We will have occasion to use this relationship at numerous points in this book. When a radio transmitter or receiver is tuned to a given frequency, we are often changing capacitance or inductor values in the circuit to make the system resonant at the desired frequency. Antennas store energy in electric and magnetic fields around the antenna, and can be modeled as resistor, inductor, capacitor (or RLC) circuits and are often designed to resonate so that the net reactance looking into the antenna is zero at the antenna's desired operating frequency.

5.1.11 Quality Factor

An RLC circuit with a small valued resistance in parallel with the inductor and capacitor has a narrow resonant behavior. Close to the resonant frequency, the current through the reactances is large, but if the frequency of a driving source is moved slightly away from resonance, the current decreases significantly. The narrowness of the current or voltage characteristic of a resonant circuit is measured using the quality factor, Q factor, or simply the Q of the circuit. A high Q circuit has a narrow resonance and a low Q circuit has a wideband resonant behavior. For sharp tuning, a high Q circuit is desirable, but for wideband applications, low Q is needed. As mentioned, if the resistance in an RLC circuit is in parallel with the reactances, a large resistor means a high Q, but if the resistance in an RLC circuit is in series with the inductor or capacitor, then a low resistance corresponds to a high Q circuit. The quality factor of an RLC circuit or other resonant system is

approximately equal to the half-power bandwidth of the resonant characteristic of the system divided by the resonant frequency.

5.1.12 Sources

Electronic circuits, antennas, and other systems and devices require a supply of electrical energy. Devices designed for this specific purpose are called power supplies, batteries, oscillators, generators, or collectively, sources. The source may be direct current (DC) as in Fig. 5.4, like a battery, or the source may supply alternating current (AC) or a sinusoidal wave at some frequency. If the frequency of an AC source is high, we refer to it as a radio frequency (RF) source. Depending on the circuit or antenna we are examining, we may prefer to power it with a voltage source or a current source. An ideal voltage source delivers a specified voltage to a pair of terminals on a circuit or antenna regardless of the current through the source. Similarly, an ideal current source passes a specified current between two terminals regardless of the resulting voltage across the source. In real life, no source is ideal but most are good enough that these approximations are valid for the range of practical currents and voltages we will deal with in this book.

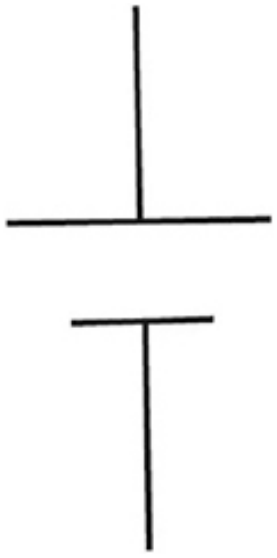


FIGURE 5.4 Circuit symbol for a DC source.

5.1.13 Units

The fundamental unit values (Farads and Henries) for capacitors and inductors are larger than is encountered with most real-world RF circuits and antennas. For convenience, we use smaller units, such as millihenries, microhenries, and nanohenries, and microfarads, nanofarads, and picofarads, where these prefixes have the following meanings:

milli m
10^{-3}
or 0.001 times

micro μ
10^{-6}
or 0.000 001 times

nano n
10^{-9}
or 0.000 000 001 times

pico p
10^{-12}
or 0.000 000 000 001 times

The fundamental unit of resistance (the ohm) is often too small, and we usually use kilohms or megohms:

kilo k
10^{3}
or 1 000 times

mega M
10^{6}
or 1 000 000 times

Example 5.1 In common terminology, a 0.001μF capacitor is called a "double zero one." The micro-farad unit is implied. A 0.001-μF capacitor is the same as a 1-nF (nanofarad) or 1000-pF (picofarad)

capacitor (obtained by moving the decimal point either three or six places to the right, respectively).

5.2 Fields

Fields were conceived as a way of understanding and predicting the observed effect of charges and magnets on other objects in the absence of any direct physical connection between the source and the affected objects. If a voltage is applied between two points (in space, in a circuit, etc.), an electric field is said to exist between those points. If a current flows in a wire or other conducting medium, a magnetic field is said to surround the conductor.

A common physics experiment illustrates this. As shown in Fig. 5.5, currents flowing in the same direction in two parallel wires will cause the two wires to be attracted and unrestrained sections of the two wires move toward each other. Reversing the direction of current flow in one of the wires will push the wires apart. The force between the wires is proportional to the product of the magnitudes of the two currents and inversely proportional to the distance between them. The magnetic force is at right angles to the direction of current flow.

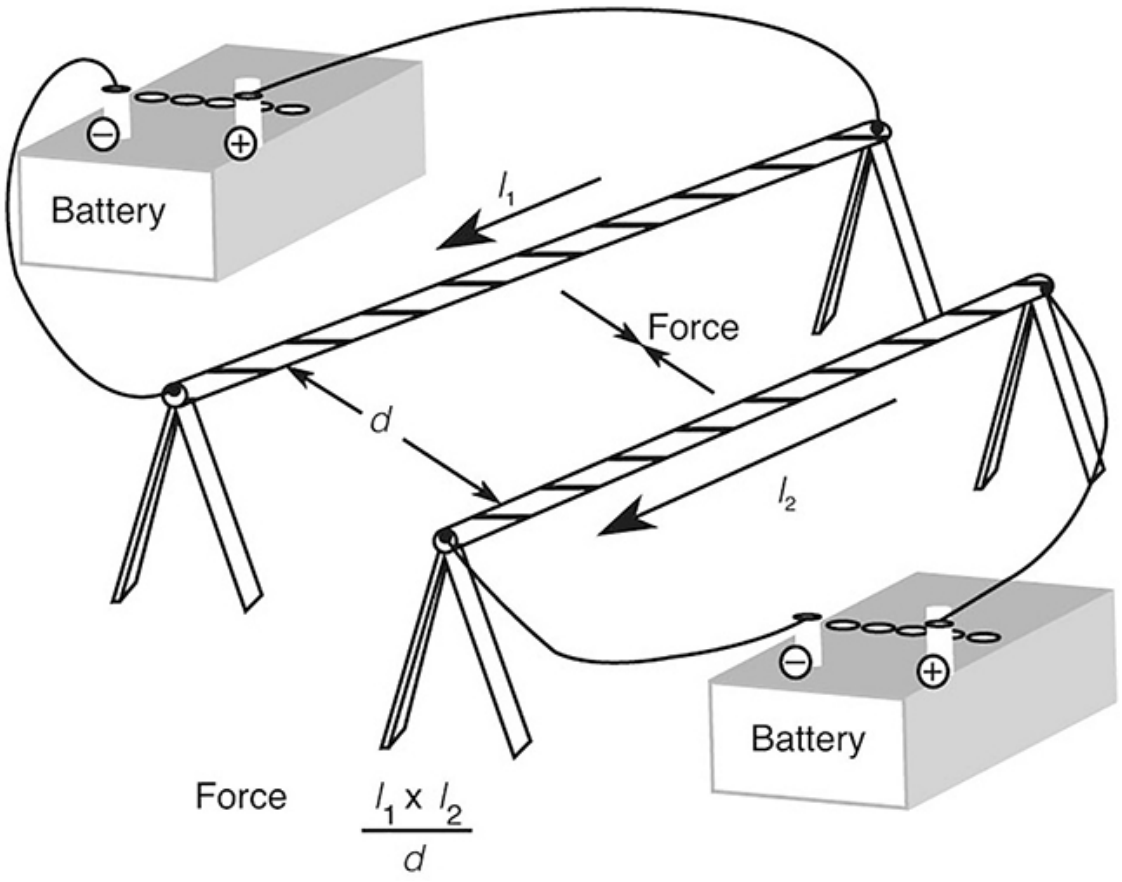


FIGURE 5.5 The force between current-carrying parallel wires.

Both electric and magnetic fields originate with electric charges. A motionless electric charge creates a static, or non-time-varying, electric field. A moving electric charge traveling at a constant velocity creates a static magnetic field at right angles to the direction of motion. An electric current is the result of many electric charges moving with a constant speed, which accounts for the magnetic field around the wire.

Rapidly varying the velocity of moving electric charges produces electromagnetic radiation and radio waves. The change in velocity is acceleration, so electromagnetic waves are associated with accelerating or decelerating charges. While there is more than one way to accelerate charged particles, for antennas this is caused by the application of rapidly varying voltages and currents to wires or other conducting structures.

Two conditions are necessary for generating radio waves with enough energy to be useful. The first is that linked time-varying electric and magnetic fields must exist in the same region of space. The second requirement is that the radiating structure must allow time-varying electromagnetic fields to break away from it and propagate into free space. As this is what makes an antenna effective, developing a familiarity with some of the geometries that allow for radiation of electromagnetic waves is one of the purposes of this book.

5.2.1 Voltage and the Electric Field

One of the most basic structures in which an electric field is important is the capacitor. A simple capacitor consists of two parallel metal plates separated by air or an insulating or dielectric material. The capacitance of the parallel plate capacitor is approximately

$$C = \frac{\epsilon A}{d} \quad (5.11)$$

where *C* is the capacitance in Farads, ϵ (epsilon) is the permittivity of the dielectric material between the plates, *A* is the area of one of the plates, and *d* is the spacing between plates.

In a vacuum, the permittivity is

$\epsilon_0 = 8.854 \times 10^{-12}\ \mathrm{F/m}$

(Farads per meter). The permittivity of a dielectric material is larger than the permittivity of free space. The permittivity of a dielectric is expressed as a multiple of

ϵ_0

and the relative permittivity,

ϵ_r

, such that

$$\epsilon = \epsilon_r \epsilon_0 \quad (5.12)$$

Values of the relative permittivity for common materials include air (1.0006), plywood (2.1), waxed paper (3.7), formica (6), and distilled water (81).

When a battery is connected across the plates of a capacitor (Fig. 5.6), negative charges flow from the negative terminal of the battery to the lower plate, where they repel the negative charges on the upper plate. This causes the negative charges from the upper plate to travel along the upper wire into the battery's positive terminal, leaving the upper plate with a shortage of electrons and a positive charge.

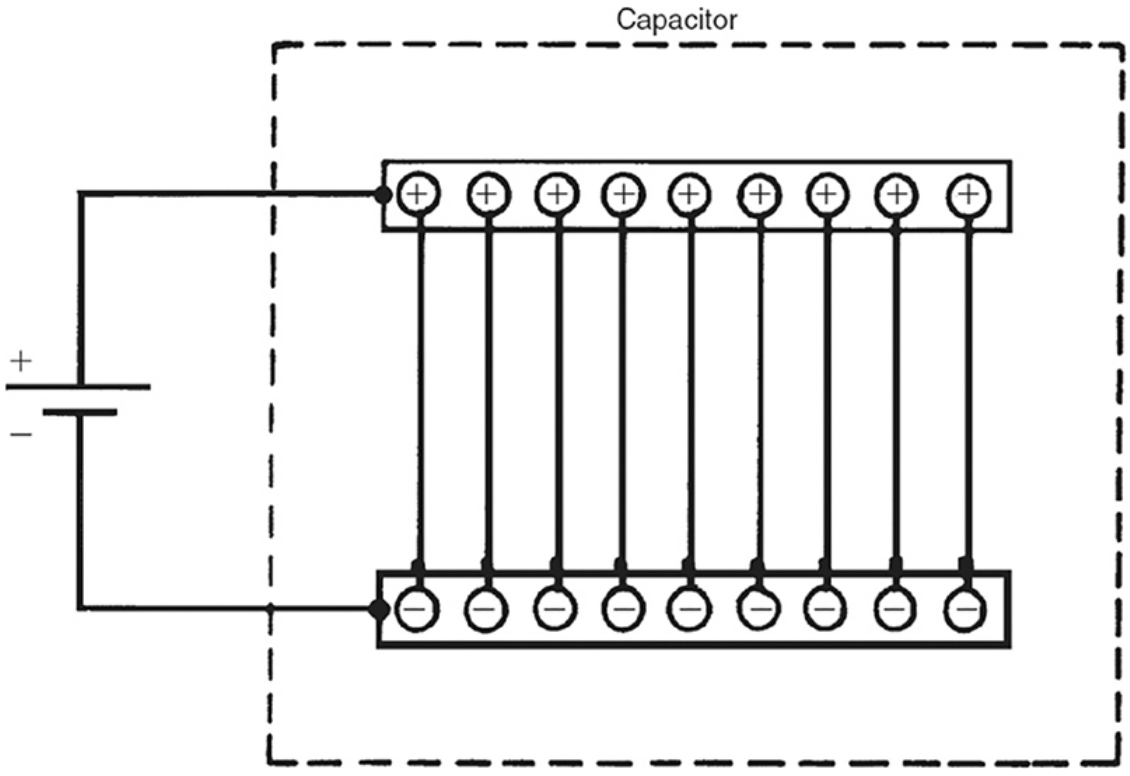


FIGURE 5.6 Charges on the plates of a capacitor.

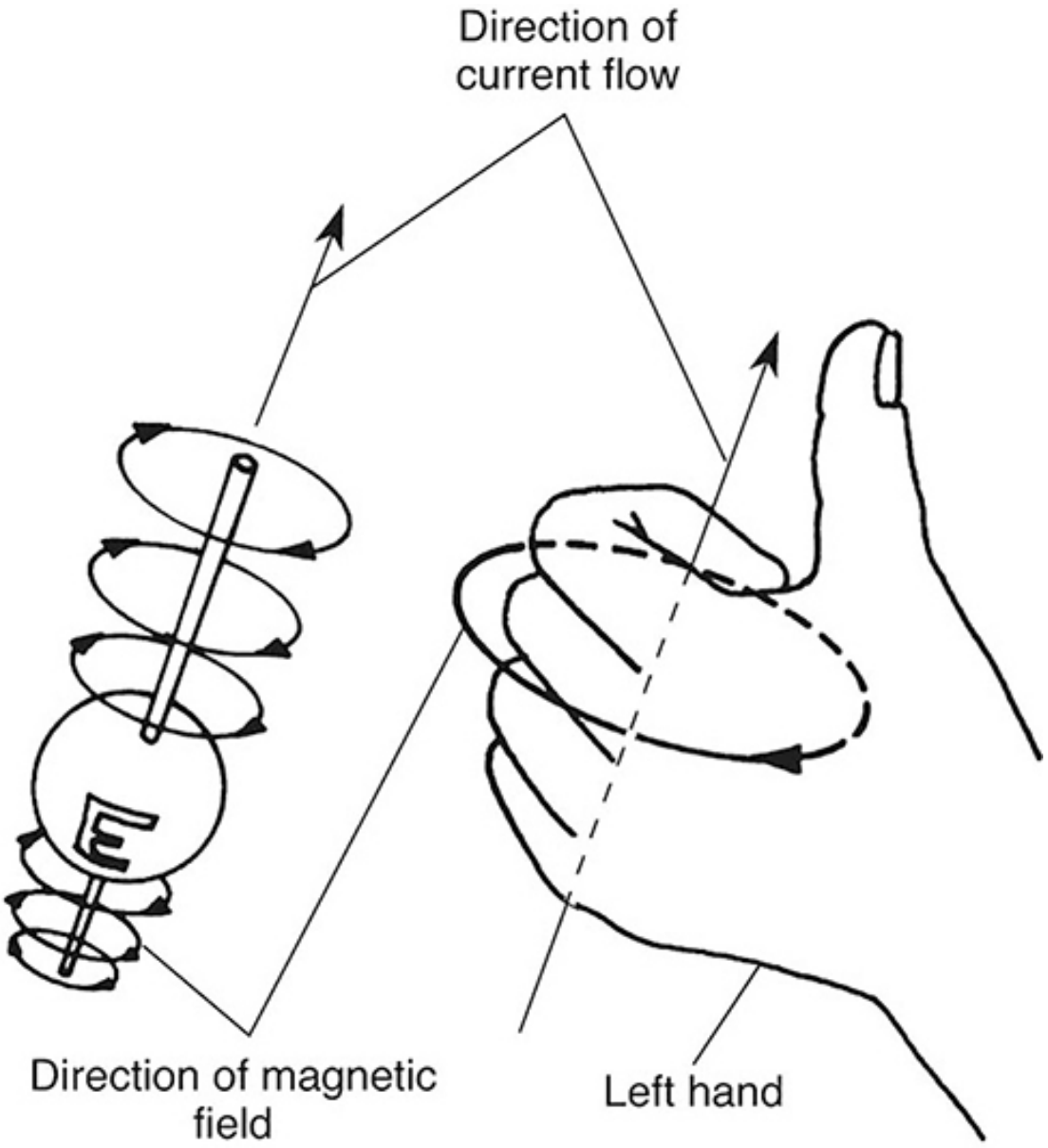
Because there is a voltage difference between the plates, an electric field exists in the space between the plates. Its direction is from the positive to the negative charges. The E-field is approximately constant between the plates (ignoring fringing around the edges) with a value

$$E = V/d$$

, where *d* is the distance between the plates. After the initial charging time, the voltage on the capacitor plates is constant and the E-field is static.

5.2.2 Current and the Magnetic Field

A charge moving at a constant velocity along a straight wire constitutes a steady current that produces a static magnetic field that is everywhere at right angles to the current. The field is circular surrounding the conductor. This is shown in Fig. 5.7.



F IGURE 5.7 Magnetic field around a half-wave dipole antenna.

The force on a moving electron from a magnetic field is proportional to the electron's velocity and at right angles to its direction of travel. A magnetic field causes a moving electron to change direction without changing its speed.

The magnetic field in a region of space can be represented by the magnetic field intensity *H* or the magnetic flux density *B*. For an

infinitely long, straight wire carrying a steady current *I*, the flux density at a point located a distance *r* from the wire is

$$B = \frac{\mu I}{2\pi r} \quad (5.13)$$

where μ is the permeability of the medium surrounding the current-carrying wire. In any region of permeability, μ, *B*, and *H* are related by

$$B = \mu H \quad (5.14)$$

For most common materials, the permeability is close to that of a vacuum,
$\mu_0 \simeq 1.26 \times 10^{-6}\ \mathrm{H/m}$
(henries per meter). The permeability is much larger for ferromagnetic materials. The permeability can be expressed relative to
μ_0
by

$$\mu = \mu_r \mu_0 \quad (5.15)$$

In ferromagnetic materials,
μ_r
ranges from a few hundred for cobalt and nickel to 5000 for iron and 100,000 or more for specialty magnetic materials such as mu-metal and permalloy. The high permeability intensifies the magnetic flux created by a given current. Transformers employ iron cores for efficient coupling between primary and secondary windings. RF transformers often use ferrite cores called toroids, as we shall see in Chap. 28.

5.2.3 Displacement Current and Maxwell's Equations

One of the earliest questions asked about capacitors was: When a capacitor is charging or discharging, how can current flow through the device when there is no connection between the plates of the capacitor? James Clerk Maxwell answered this question in 1861 when he proposed the existence of a displacement current (in contrast to the conduction current in a wire) in the space between the plates.

Because the displacement current flows between the plates of a capacitor even when it is in a vacuum, Maxwell concluded there must be some invisible medium permeating all of space that carries the displacement current. He called this medium the *æther*. It took scientists a half century to conclude there was no such thing. The displacement current is not carried by any medium. It is carried by space itself.

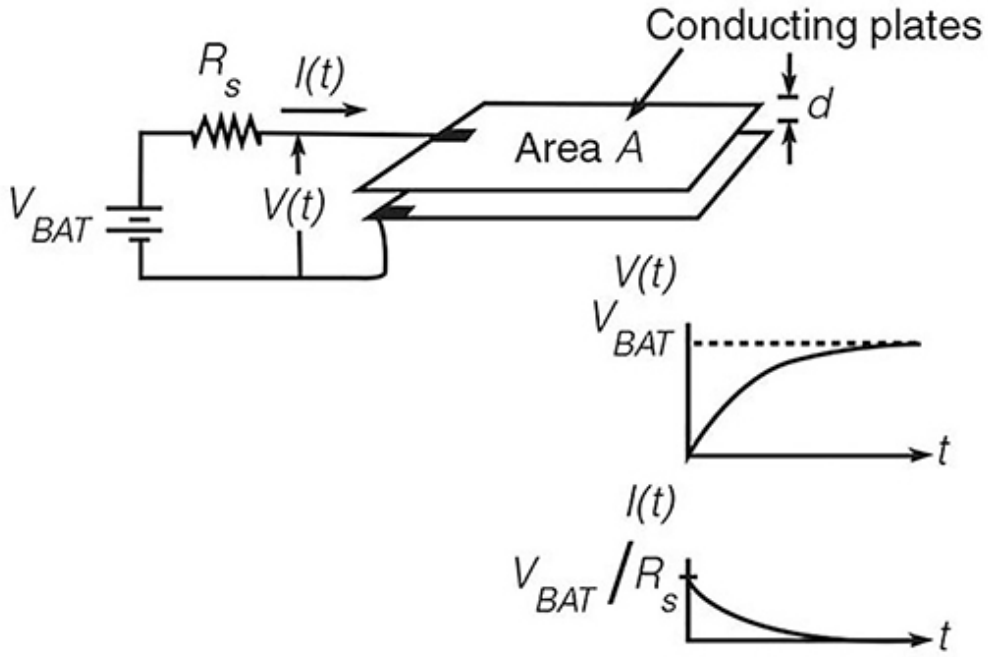
After discovering the displacement current, Maxwell showed that a changing electric field creates a magnetic field and a changing magnetic field creates an electric field. He added symmetric cross-coupling terms to then-existing equations that described the relationship between electric and magnetic fields. This allowed Maxwell to summarize the important characteristics of electromagnetism in a family of interrelated equations. Maxwell and other scientists and mathematicians found that the equations had wave solutions that traveled through space, much like sound waves and other mechanical vibrations.

Maxwell summarized all that he knew about electromagnetism in his *Treatise* of 1865, and it was there that he predicted, on the basis of his mathematical equations, the existence of electromagnetic waves, as well as the speed at which these waves must travel through space. Before Maxwell's time, people were aware of magnetic and electric forces, but no one suspected the connection between these forces and visible light. In one of the greatest discoveries in history, Maxwell found that the speed of the electromagnetic waves predicted by the equations was the same as the speed that had been measured for light. Maxwell showed that light itself is an electromagnetic wave.

A quarter century later, Heinrich Hertz demonstrated that he could create, transmit, and receive what we now call radio waves. Hertz thus took the first steps in determining the physical geometry of conductors (antennas, in other words) that would allow the generation and detection of the EM radiation predicted by Maxwell. Hertz generated a spark across a gap between two curved conductors and attempted to detect the fields from that spark with a similar set of conductors a few feet away. And so it was that Heinrich Hertz was quite likely the first to experience the magic of radio.

5.3 The Hertzian Dipole

We will use a capacitor to help understand how a basic antenna works. As a capacitor charges, current flows into the capacitor for a short time. We refer to this as a transient phenomenon, as opposed to the steady state voltage of the capacitor when fully charged. When the voltage across the capacitor is equal to the voltage of the source, the current stops. The charging time depends on the source voltage and resistance. The final value of the current is zero, just as the final voltage across the capacitor is the source voltage. As shown in Fig. 5.8, the voltage and current follow exponential curves.



$$V_{CAP} = V(t) = V_{BAT}\left(1 - e^{-\frac{t}{R_s C}}\right)$$

$$I_{CHG} = I(t) = \frac{V_{BAT}}{R_s} e^{-\frac{t}{R_s C}}$$

R_s = the sum of all resistances in the circuit, including the resistance of the interconnecting wire and the internal resistance of the battery itself

$$C = \frac{\epsilon A}{d}$$

F IGURE 5.8 Transient voltage and current when charging a capacitor.

Once the capacitor is charged, a steady E-field exists between the plates of the capacitor. Since the field is static field, it does not create a magnetic field. But during the short instant when current

flows, the field changes over time. This means that a changing magnetic field encircles each of the wires. With changing electric and magnetic fields around the circuit while the capacitor charges, during the transient charging period the circuit radiates a small portion of the energy supplied by the battery as electromagnetic waves.

Next, we will energize the capacitor not with a battery, but with a power amplifier or other source supplying a sinusoidal RF voltage represented by the circle marked E in Fig. 5.9. With a sinusoidal source, the current in the structure changes continuously and the structure has the potential to radiate electromagnetic waves continuously.

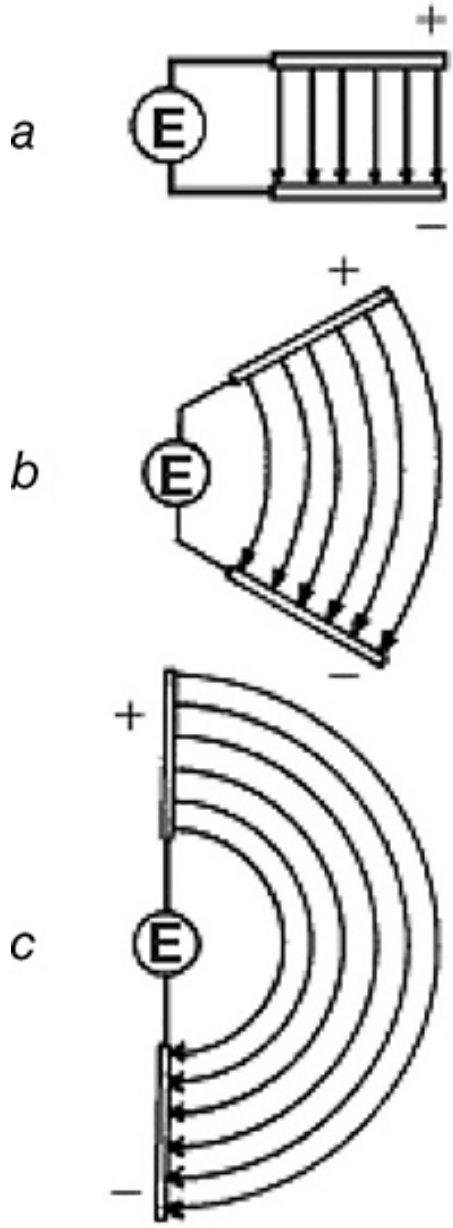


FIGURE 5.9 Electric field in a capacitor with the conducting plates arranged at various angles.

As the capacitor is a poor antenna, it radiates only a small amount of energy as electromagnetic waves. We can enhance the potential for radiation by modifying the geometry of the capacitor as shown in Fig. 5.9. If the plates are spread apart, the electric field between them curves to meet the plates at right angles (Fig. 5.9*b* and *c*). The straight field lines in Fig. 5.9*a* become arcs in Fig.

5.9*b* and approximately semicircles in Fig. 5.9*c*, where the plates are in a straight line. Since the field lines extend away from the conductors rather than remaining mostly confined between the plates as in the original capacitor, this new structure is a better radiator of electromagnetic waves.

In addition to altering the orientation of the plates, we can change their shape. Instead of flat rectangular metal sheets, we can make the two elements from cylindrical metal rods or wires. This is shown in Fig. 5.10. Electrons are pushed onto one rod or wire and pulled from the other in direct response to the sinusoidal source voltage applied to the two rods. The amplitudes of the static and quasi-static E- and H-fields in the vicinity of the two rods are changing at the same sinusoidal rate as the applied voltage.

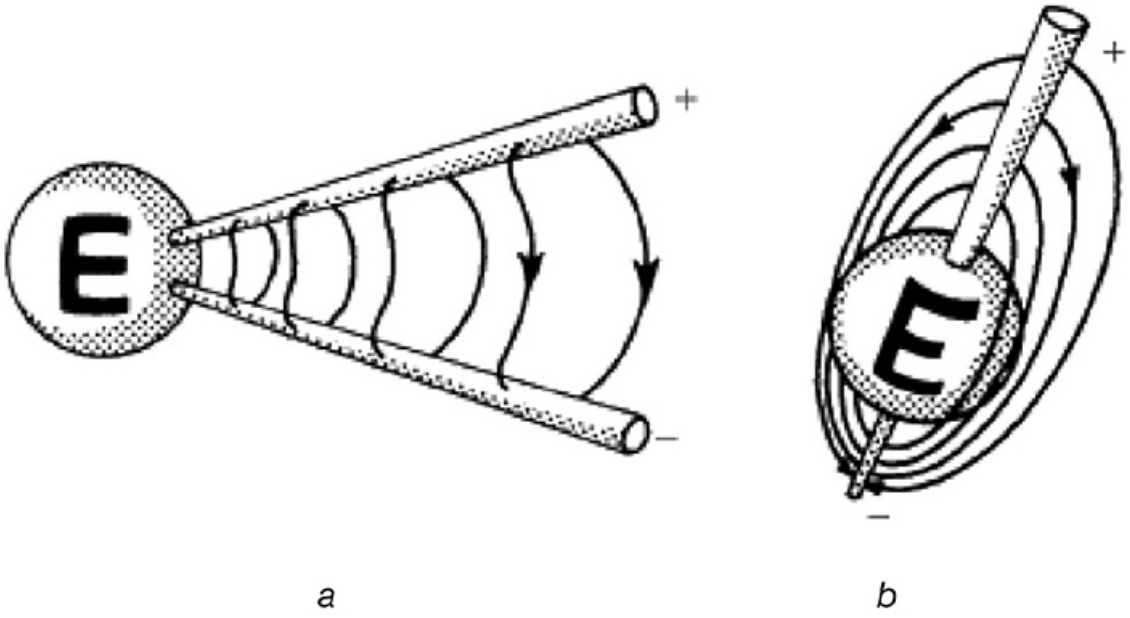


FIGURE 5.10 Electric field between wires arranged at two different angles. By changing the shape and geometry of the conductors, a parallel plate capacitor (the first sketch in Fig. 5.9) can be evolved into a dipole antenna.

Since the velocities of the electrons are constantly changing as they shuttle back and forth between the rods and the sinusoidal source, they are in a state of continuous acceleration and

deceleration, thus setting up linked time-varying electric and magnetic fields in the space enclosing the rods. Maxwell's equations predict the existence of new fields resulting from the acceleration of electric charge in the region between the two rods. Accelerating charges along the rods create a magnetic field. Since the magnetic field changes in time, they create electric fields, leading to a propagating wave with linked electric and magnetic fields that moves away from the antenna at the speed of light. This is the radiation mechanism of a dipole antenna.

Suppose we now attempt to observe the effects of these accelerating charges at a point in space far from the region around the rods. We would detect a sinusoidally varying E-field that is at right angles to a straight line drawn between the antenna and our observation point. Similarly, we detect a magnetic field that is at right angles to the line between the antenna and us, and at right angles to the E-field. Both the E-field and the H-field are at right angles to the direction of travel from the antenna to the receiving point.

These types of electromagnetic fields are not static or quasi-static fields, and they were totally unknown prior to Maxwell's great discovery. They represent a transverse electromagnetic wave originating from the region of the rods and traveling through space at the speed of light. The E- and H-fields associated with the accelerating electrons detach from the rods and move away from the antenna as the electrons move along the conductors in the antenna.

The electromagnetic field at a distant receiving point has the same frequency as the source voltage applied to the antenna, except for a delay that is proportional to the distance from the rods to the observation point. That delay is caused by the finite speed of light. The E- and H-fields are components of an electromagnetic wave propagating through space, and that wave is a real entity. The wave carries energy and can do useful work wherever it travels.

The received E-field at a point far from a very short (relative to wavelength) pair of rods oriented vertically and driven by a sine

wave at frequency
f_0
is described in spherical coordinates by

$$E_\theta = \frac{f_0 \mu I_0 h}{2r} \sin\theta \sin\left[2\pi f_0 \left(t - \frac{r}{c}\right)\right] \quad (5.16)$$

and the corresponding H-field is

$$H_\phi = \frac{f_0 \mu I_0 h}{\eta r} \sin\theta \sin\left[2\pi f_0 \left(t - \frac{r}{c}\right)\right] \quad (5.17)$$

To better understand these formulas, let's break them down piece by piece, looking at each term:

- f_0 is the frequency of the source that excites the antenna and I_0 is the peak amplitude of the sinusoidal current from the source flowing through the antenna. Obviously, the stronger the source, the stronger the field radiated by the antenna, so it's natural that the formula for the radiated field is proportional to I_0.
- μ is the permeability of the medium in which the antenna radiates. This is usually free space, so $\mu = \mu_0$ as in Eq. (5.15).
- In the magnetic field, $\eta = \sqrt{\mu/\epsilon}$ is the intrinsic impedance of the medium the antenna radiates in. This quantity has units of ohms and in free space, $\eta = \eta_0 \simeq 376.7\ \Omega$.
- *h* is the total height, or length, of the antenna formed by the two rods. These formulas are valid when *h* is small with respect to *λ*. A good value might be *λ*/20 or less. For longer antennas, more accurate formulas for the field radiated by a dipole antenna are available.
- *r* is the distance from the center of the two rods to a distant point where we are observing the E- and H-fields.

As the power radiated by the antenna spreads over an ever-increasing spherical area, the associated fields decrease with distance from the antenna according to $1/r$.

- θ (pronounced "thay-ta") is the angle from the axis of the antenna. The sin θ term tells us that the E-field drops to zero off the ends of the antenna and is maximum in the direction that is broadside to the antenna, where $\sin\theta = 1$.
- $\sin[2\pi f_0(t - r/c)]$ identifies the time-varying behavior of the E-field as a function of the drive signal frequency and the distance of the receiving point from the rods. Over time the sine function oscillates through multiple cycles from -1 to $+1$.
- The subscript θ for the E-field means the only component of the E-field that has any amplitude far from the two rods is an E-field lying in the same direction as the axis of the rods. The radiated wave is best analyzed in spherical coordinates represented by (r, θ, ϕ). An E-field comprised of solely E_θ is consistent with our earlier statement that the E- and H-fields are always at right angles to the direction of wave propagation.
- Similarly, the subscript ϕ for the H-field tells us that the only detectable component of the magnetic field at a distant point is in the ϕ direction as defined in spherical coordinates; that is, it is a circular field whose axis is the same as that of the rods that form the source antenna.

The wave represented by Eqs. (5.16) and (5.17) is a spherical wave. Power radiated by the antenna moves away from the source and propagates outward in space, forming an ever-expanding sphere. As the distance from the antenna becomes large, a small region on a large sphere appears flat, so at any distant receiving site, the wave appears to be a plane wave.

The ratio of the electric and magnetic field is

$$\frac{E_\theta}{H_\phi} = \eta = \sqrt{\frac{\mu}{\epsilon}} \quad (5.18)$$

This is the impedance of the medium in which the wave propagates. Since the electric field intensity has units of *V/m* and the magnetic field intensity has units of *A/m*, the ratio has units of *A/V* or ohms. The impedance of free space is

$$\eta_0 = \sqrt{\frac{\mu_0}{\epsilon_0}} = \sqrt{\frac{1.26 \times 10^{-6} \mathrm{H/m}}{8.854 \times 10^{-12}\ \mathrm{F/m}}} = 376.7\ \Omega \quad (5.19)$$

μ_0

and
ϵ_0
are the permeability and permittivity of a vacuum—free space, in other words. In older texts, the approximation
$\eta_0 \simeq 20\pi$
was often used. Interestingly, the vacuum of free space has at least two characteristics (
ϵ_0
and
μ_0
) and an impedance (measured in ohms). So the "nothing" of free space isn't exactly nothing, is it?

The antenna we have just described is known as a Hertzian dipole. The formulas are valid when the antenna is short compared to the wavelength of the signal applied to it—exactly the situation with Heinrich Hertz's laboratory setup of 1887. The shape of the antenna's E-field pattern is described by the sinθ term; it is donut-shaped in a three-dimensional view and a figure eight when viewed in two dimensions with the axis of the dipole lying in the plane of observation. Compared to an isotropic radiator, which is fictitious but which—if it did exist—would radiate equally in all directions, the directivity of this antenna is 1.5, or 1.75 dBi (decibels relative to

isotropic). From the standpoint of its pattern, this would not be a bad antenna, but the input impedance is problematic.

The input impedance of a Hertzian dipole exhibits both resistance and capacitative reactance:

$$Z_{\text{in}} = R_{\text{rad}} + X_C \quad (5.20)$$

where
X_C
is the quasi-static or low-frequency capacitance of the dipole's physical structure. For a very short dipole (
$h << \lambda$
),

$$R_{\text{rad}} \simeq 800\left(\frac{h}{\lambda}\right)^2 \quad (5.21)$$

For a dipole with
$h = \lambda/20$
, which is roughly the longest dipole this model is valid for,
$R_{\text{rad}} = 2\ \Omega$
. This is the radiation resistance of a lossless Hertzian dipole. A real antenna will also have ohmic resistance due to losses in the antenna. When a small radiation resistance is combined with ohmic losses, the power dissipated in the antenna is large in comparison to the power radiated by the antenna, and the radiation efficiency of the antenna is small.

The capacitive reactance term also complicates the problem of delivering power to the antenna in an efficient manner. In theory, a value of series inductance can be selected that will cancel out the capacitive reactance at
f_0
, but the
$I^2 R$

losses in a practical inductor will dissipate most of the RF power delivered to the system.

As a result of these effects, the Hertzian dipole is an impractical antenna, with typical radiation efficiencies of well under 1 percent. However, it is useful as a learning tool and serves as the basic mathematical building block for analyzing real-life antennas.

5.4 The Half-Wave Dipole

To have a reasonably large input resistance (so the radiation efficiency is close to 100%) and a small input reactance (so the antenna is resonant at the desired operating frequency), we need to make the dipole longer. In this case, the Hertzian dipole model is not accurate. We need to develop a model for longer dipole antennas that are comparable in size to the electromagnetic wavelength. The most common case is a dipole antenna that is a half wavelength long, or a half-wave dipole.

By making simplifying assumptions, we can obtain formulas for the radiation pattern, input resistance, and other antenna parameters for the half-wave dipole. With the Hertzian dipole, the current on the antenna arms is approximated as constant, so that the magnitude of the current is the same everywhere on the antenna. In reality, the current drops to zero at the ends of the dipole arms.

With a dipole antenna, a current introduced at the feedpoint travels along the wire to the ends of the dipole arms. At the ends, a reflection occurs and a wave starts back toward the center where the feedpoint is located. A standing wave is created on the dipole arms. This standing wave represents a combination of the outbound and reflected waves. The amplitude and phase of this standing wave can be used as a model for the current on the dipole when finding the field radiated by the antenna.

The most common models use either a straight-line decrease of current from the feed-point to each end that roughly approximates the standing wave on the dipole (triangular current model) or a

more accurate sinusoidal distribution that is zero at the dipole ends assuming a source frequency
f_0
(sinusoidal current model). The actual current on the dipole is a bit different from sinusoidal and can be found even more accurately if needed using a numerical model.

With the sinusoidal current model, the resulting E-field and H-field expressions for longer dipoles are similar to Eqs. (5.16) and (5.17), except that the sin θ term of those equations is now replaced with

$$f(\theta) = \frac{\cos\left(\frac{\pi h}{\lambda}\cos\theta\right) - \cos\frac{\pi h}{\lambda}}{\sin\theta} \quad (5.22)$$

For the special case of the half-wave dipole, Eq. (5.22) reduces to

$$f(\theta) = \frac{\cos\left(\frac{\pi}{2}\cos\theta\right)}{\sin\theta} \quad (5.23)$$

This is the radiation pattern of a half-wave dipole.

Broadside to the axis of a half-wave dipole (
$\theta = 90^\circ$
) the radiation pattern is equal to its maximum value of unity. In the endfire direction along the dipole axis (
$\theta = 0$
) the radiation pattern is zero or has a null. Figure 5.11 is a graphical representation of the geometrical variation given by Eq. (5.23)—the well-known donut pattern of a half-wavelength dipole. The half-wave dipole has 0.4 dB more gain than the Hertzian dipole and the lobe is slightly narrower than that of the Hertzian dipole.

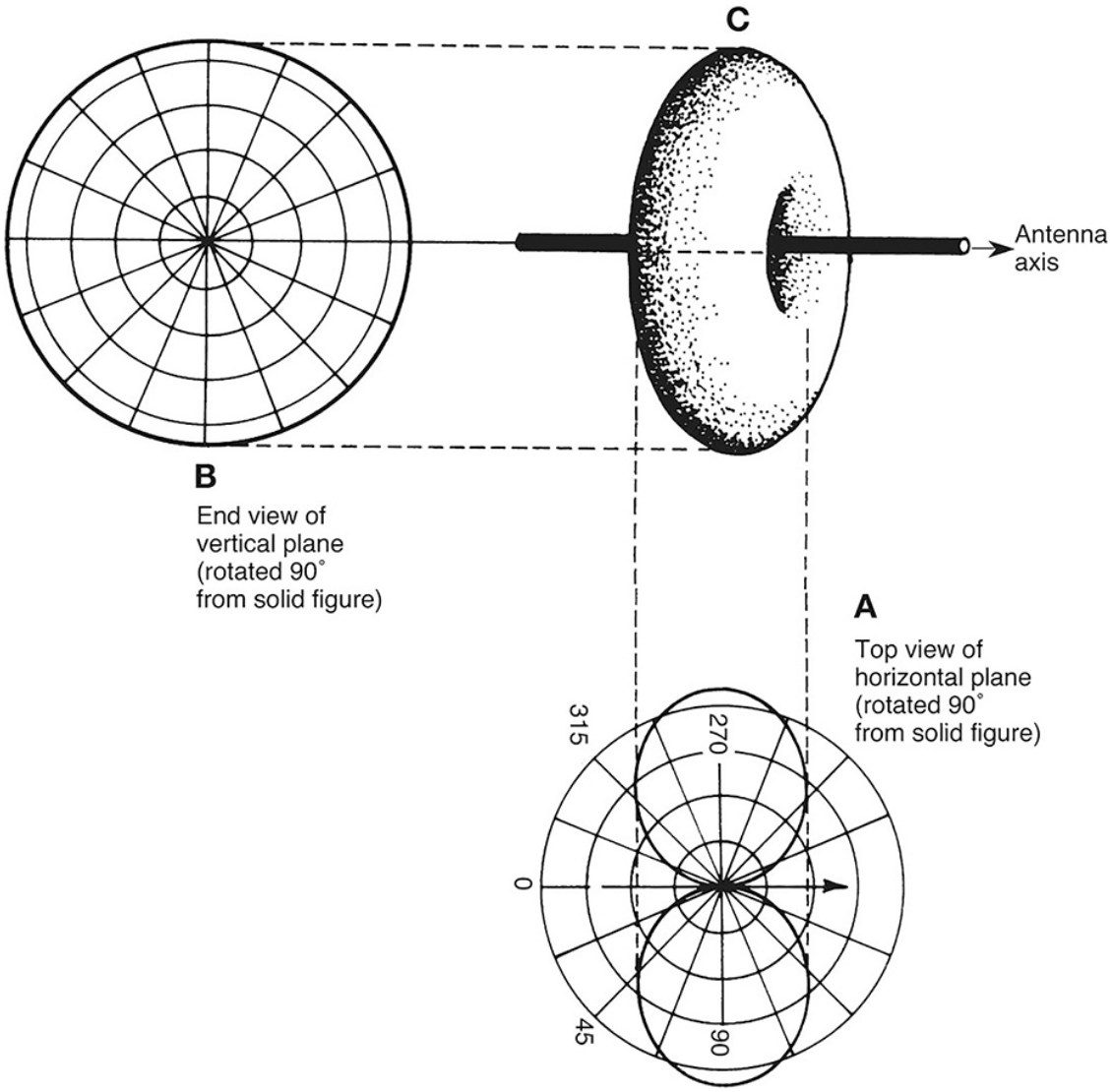


FIGURE 5.11 Free-space radiation pattern of $\lambda/2$ dipole.

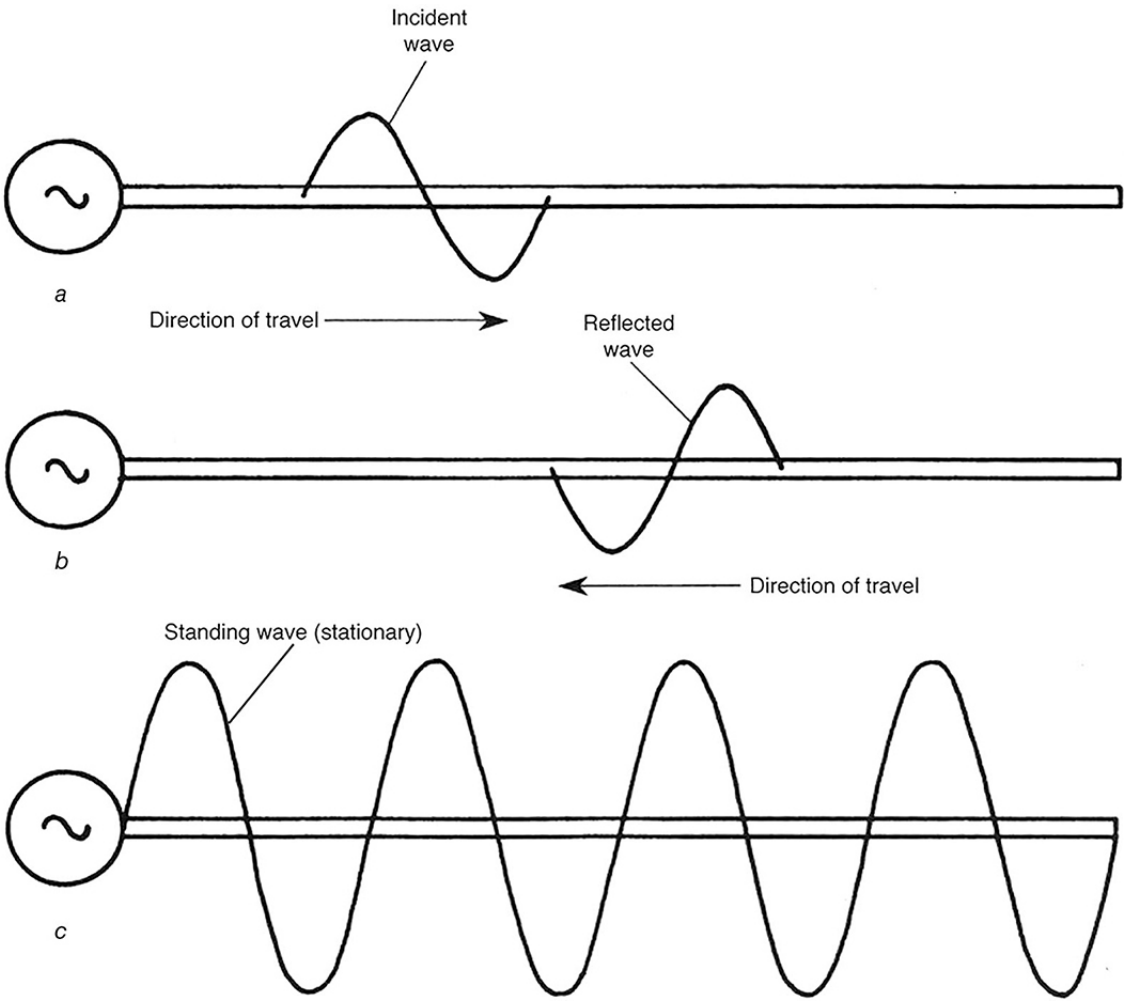
The big news is that the half-wave dipole is a more effective antenna compared to a short dipole. The radiation resistance is 73 Ω and the input impedance has only a small reactive component. This higher value of R_{rad} makes the antenna easy to match with commonly available transmission lines. More important, it is easier to deliver most of

the transmitter output power to the antenna rather than seeing it dissipate in the resistance of conductors and connectors.

5.4.1 Standing Waves

We have motivated the sinusoidal current model for the dipole antenna using standing waves. Standing waves are useful in understanding many other types of antennas that will be covered in this book. When a single-frequency RF signal is introduced on a conductor, an alternating current and voltage consisting of sinusoidal waves propagates along the wire. This is called a traveling wave or propagating wave. As the wave propagates along the conductor, the resistance of the conductor gradually diminishes its amplitude, but the wave continues to propagate until it reflects from the end of the conductor or some other junction. The wave traveling from the transmitter toward the end is called the incident or forward wave. The returning wave is called the reflected or reverse wave. The reflected wave adds to the forward wave and produces a standing wave on the conductor.

To visualize a standing wave, assume that the transmitter is turned on just long enough to allow one cycle of RF (solid line) to drive the line (Fig. 5.12*a*). Because it is an alternating current, it has a changing magnetic field associated with it. At the end of the conductor, the current path is broken, thus causing the magnetic field to collapse. This creates an induced voltage at the end of the conductor that causes a new current to flow back toward the source, as in Fig. 5.12*b*. Because electric and magnetic fields are linked by the physical relationships described by Maxwell's equations, propagating waves on a line always consist of linked current and voltage. An observer standing at a point midway along one side of the dipole would see a single cycle of RF current and voltage in the forward direction followed later by another cycle coming back.



F IGURE 5.12 Standing wave formation on an antenna. (*a*) Incident (forward) traveling wave. (*b*) Reflected traveling wave. (*c*) Resultant standing wave.

If now we let the transmitter send a continuous flow of sinusoidal energy of frequency

f_0

to the dipole, a continuous flow of forward and reflected voltages and currents will result. The two waves flow in opposite directions. The voltage and current observed at a point on the line are the sum of the voltages and currents associated with the forward and reverse waves.

To our observer standing at the same point on the dipole, all that is evident at that point is a single voltage or current that is the

vector sum of both voltages or both currents. When the two voltages are in phase with each other they reinforce, and the resultant seen by the observer is at its maximum value. When they are out of phase they cancel and the sum is at a minimum. The locations of the voltage and current maxima and minima depend on the frequency of the driving signal, the length of the line, and the junction or load at the end of the line.

For a conductor of any finite length, such as our dipole antenna, the points at which the maxima and minima of the resultant voltage or current occur (Fig. 5.12*c*) are stationary. For a continuous driving signal of fixed

f_0

the maxima and minima stand still even though both the incident and the reflected waves are constantly traveling out and back along the wire. This is why the wave pattern resulting from the sum of a forward and reverse wave is referred to as a standing wave.

Now we will return to the dipole antenna and apply the standing wave concept more carefully. With the feeding source in the center of the dipole, the two arms of the antenna can be viewed as separate transmission lines, each with a standing wave pattern. The standing current waveform is always zero at both ends of the dipole (because of the boundary condition there—it's an open circuit). There are voltage maxima at the ends of the wires or arms. At the ends of the dipole, the amplitude of the standing voltage waveform is always maximum, but of opposing polarities. At the source location in the middle of the dipole, there is a current maximum. For a half-wave dipole, the standing wave pattern on each arm of the dipole is approximately a quarter cycle of a sine wave.

If a radiator is longer than

$\lambda/2$

, the standing wave of current reverses direction every half-wavelength along the conductor. Starting with the boundary condition that the current at either end of a center-fed dipole must be zero, as we move toward the feedpoint, when we reach the next current node

$\lambda/2$
from the end of the dipole, the current reverses polarity for the next half-wavelength along the radiator (or as much of a half-wavelength as exists before reaching the feedpoint). Remembering that the current reverses in each adjacent
$\lambda/2$
section of a conductor will be key to understanding the patterns of loops, folded dipoles, collinear arrays, and other antenna types discussed in later chapters.

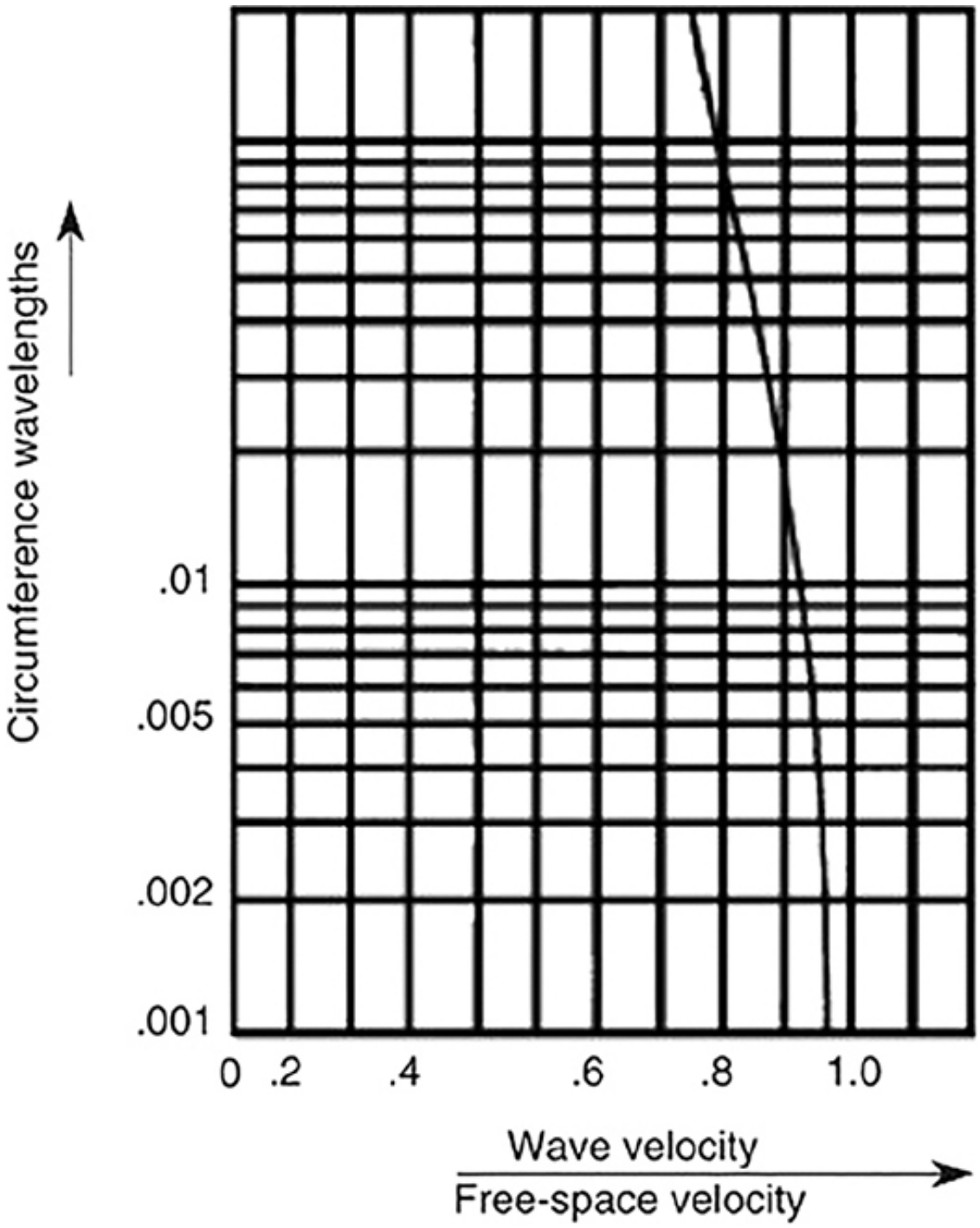
5.4.2 Velocity of Propagation and Antenna Length

In free space, electromagnetic waves travel at a constant velocity of 300,000 km (or approximately 186,000 mi) per second. This can be computed from physics constants using

$$c = \frac{1}{\sqrt{\mu_0 \epsilon_0}} \quad (5.24)$$

where
$\mu_0 \simeq 1.26 \times 10^{-6}\ \mathrm{H/m}$
is the permeability and
$\epsilon_0 \simeq 8.854 \times 10^{-12}\,\mathrm{F/m}$
is the permittivity of the vacuum. Because of the influence of the materials that make up the antenna structure, electromagnetic waves on an antenna usually propagate at a slower velocity. For this reason, a dipole antenna is resonant when it is slightly shorter than half of the wavelength of the radiated wave in free space.

The length of an antenna at resonance depends on several factors. A thin wire antenna has less effect on wave velocity than an antenna with a large cross section. As the circumference of the antenna increases, the wave velocity becomes progressively lower relative to its free-space velocity. The effect of antenna circumference on wave velocity is illustrated in the graph of Fig. 5.13.



F IGURE 5.13 Effect of antenna circumference on wave velocity.

Other factors can also lower wave velocity on the antenna. Stray capacitance increases the dielectric constant. In many installations

stray capacitance is dominated by the insulators used to give physical support to the antenna, as well as nearby objects made of metallic or high dielectric materials. The change in velocity resulting from stray capacitance is sometimes called an end effect, as if the conductor is slightly longer than its physical length in its electrical behavior. End effects for a dipole of wire that is thin compared to its length is counteracted by making the physical length about 5 percent shorter than half of the free space wavelength, as expressed in the formula

$$L = 0.95\left(\frac{492}{f}\right) = \frac{468}{f} \quad (5.25)$$

where *L* is the physical length in feet and *f* is the operating frequency in megahertz. This formula is reasonably accurate for determining the physical length of a resonant dipole antenna.

The capacitive end effect impacts the standing voltage and current waveforms along the antenna. When the standing waves are measured, the standing wave pattern at the nodes is not identically zero. Some current is necessary to charge the stray capacitance— especially that of the end insulators. This lowers the apparent impedance of the dipole at its ends from a theoretically infinite value (i.e., the open circuit at the dipole ends) to a few thousand ohms in practical dipoles, as suggested in Fig. 5.14.

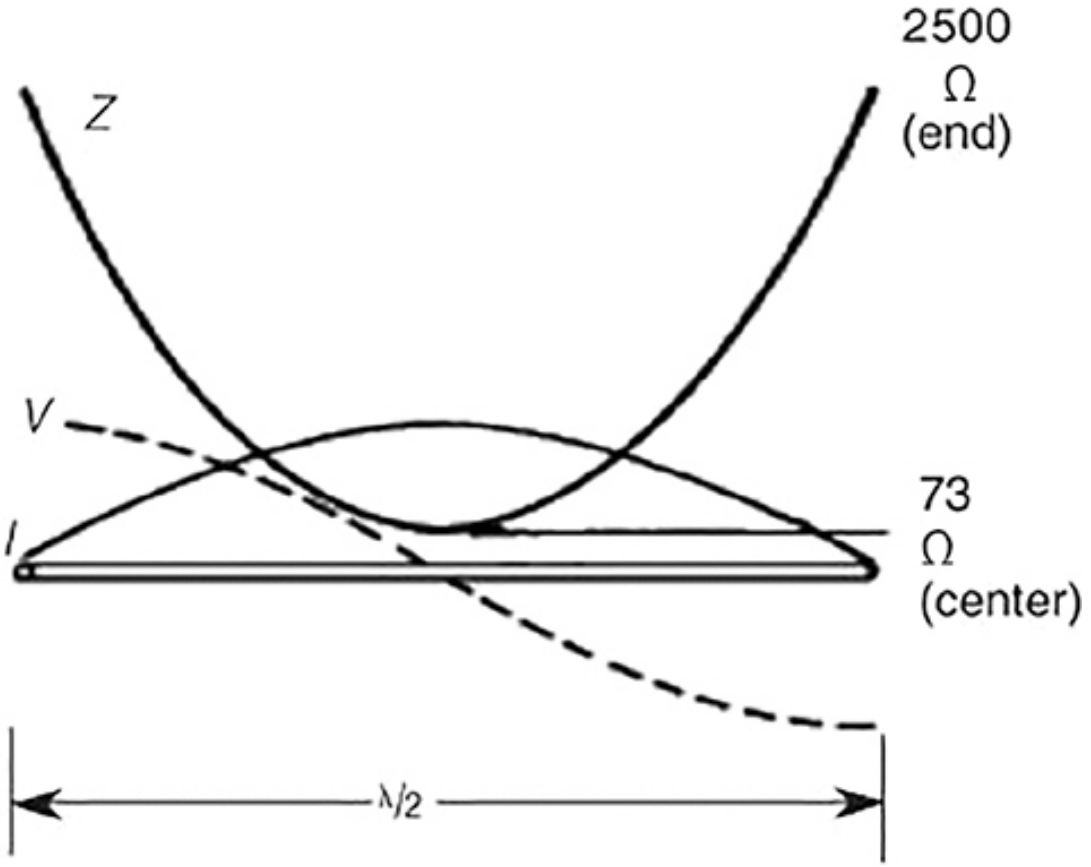


FIGURE 5.14 Impedance along a half-wave dipole antenna.

The half-wave dipole is the shortest resonant length, but dipole antennas that are an integer multiple of

$\lambda/2$

can also be resonant. Such antennas are said to be resonant at harmonic frequencies of

f_0

, the fundamental or design frequency of the

$\lambda/2$

dipole. If an antenna is four half-wavelengths at the transmitter frequency, it is being operated at the fourth harmonic of its lowest resonant frequency.

5.4.3 Half-Wave Dipole Feedpoint Impedance

The impedance of any antenna is of the form given in Eq. (5.8), with the real part representing power dissipated in losses and radiated by the antenna and the imaginary part is the reactance of the antenna.

The impedance of a dipole depends on the length of the antenna, the operating frequency, and the details of the antenna feed geometry. A half-wave dipole is usually fed at the center, but in some cases may be fed at other points along the antenna. Because the voltage and current vary along the length of the half-wave dipole, and the impedance at any point on the dipole is equal to the voltage at that point divided by the net current passing through that point, the impedance varies along the length of the antenna. If V is divided by I at each point of the voltage and current curves of Fig. 5.14, the result is the impedance curve Z. The resistance (i.e., the real part of the complex impedance) is a minimum of about 73 Ω at the center point and rises to 2500 Ω or more at the ends.

If the antenna is cut to a length corresponding to an exact resonance the feedpoint impedance at the center will be purely resistive. If the antenna is longer or shorter than resonance, the impedance is partly reactive and has a nonzero imaginary part. When the antenna is shorter than its resonant length, the feedpoint impedance will exhibit capacitive reactance; conversely, when the antenna is somewhat longer than the resonant length, inductive reactance will be seen.

If the frequency is increased, the physical length of a given antenna stays the same but the electrical length becomes a greater fraction of a wavelength, with a corresponding increase in the amount of inductive reactance seen at the feedpoint. Conversely, if the frequency is lowered, the electrical length decreases, and the feedpoint impedance becomes more capacitive in nature.

Unless it is compensated for by a matching reactance in the antenna feed network, the mismatch caused by the reactive part of the feedpoint impedance reduces the amount of power accepted by the antenna. For lengths other than an odd multiple of

$\lambda/2$
, dipoles exhibit sufficient feedpoint reactance to make the use of a matching network in the circuit somewhere between the transmitter and the feedpoint mandatory. Circulating currents in any wires or inductive components in the matching network contribute to resistive losses, lowering the efficiency of the antenna even more.

5.5 Antenna Losses and Radiation Efficiency

The purpose of an antenna system is to convert as much source energy as possible to electromagnetic radiation. The power accepted by an antenna is either radiated or absorbed by losses in the antenna and converted to heat.

Power can be dissipated as losses in the antenna in conductors and in any dielectrics in or near the antenna. Current flowing along the antenna conductor dissipates energy in the form of heat. In the
$I^2 R_{\text{ohmic}}$ loss, R_{ohmic}
is the ohmic resistance. The ohmic resistance is not the same as the wire's DC resistance as measured on a conventional ohmmeter. It is typically higher due to the skin effect, which limits the flow of RF energy to the surface of the conductor at high frequencies. Many antennas exhibit an
$I^2 R_{\text{shunt}}$
loss because of the leakage, or shunt, resistance of dielectric elements such as insulators. This effect of this resistance is often lumped in with the ohmic resistance into the loss resistance
R_{loss}
.

Radiated power expended in the form of radiation can be thought of as an
$I^2 R_{\text{rad}}$
loss, where
R_{rad}
is the radiation resistance. The energy dissipated in the radiation resistance is the useful part of the total power accepted by the

antenna. Because the actual power from the source is split between R_{rad} and the ohmic and shunt resistance terms, the latter two resistances should be kept as low as possible. The radiation resistance of a center-fed half-wave dipole antenna in free space is approximately 73 Ω. In a well-designed installation, the radiation resistance should be large compared to the loss resistances, and most of the available energy will be radiated as useful signal. The half-wave antenna is an efficient radiator under most circumstances.

The impedance of antenna can be expressed in an expanded form as

@@@@@@@@@@@@@@@@

$$Z_{\text{in}} = R_{\text{loss}} + R_{\text{rad}} + jX \quad (5.26)$$

where

R_{loss} is the part of the antenna resistance due to ohmic and dielectric losses,

R_{rad} is the radiation resistance, and *X* is the reactance. The radiation efficiency of an antenna is

$$\eta_{\text{rad}} = \frac{P_{\text{rad}}}{P_{\text{in}}} \quad (5.27)$$

or the ratio of the power radiated by the antenna to the power accepted by the antenna at its feed port. Using the relationship between resistance and dissipated power, the radiation efficiency can be expressed as

$$\eta_{\text{rad}} = \frac{R_{\text{rad}}}{R_{\text{in}}} = \frac{R_{\text{rad}}}{R_{\text{rad}} + R_{\text{loss}}} \quad (5.28)$$

where again

$R_{\text{loss}} = R_{\text{ohmic}} + R_{\text{shunt}}$
is the sum of the ohmic and dielectric resistances of the antenna.

The electrically short dipole model we used at the beginning of our analysis has a radiation pattern that is comparable to the $\lambda/2$ dipole, but its radiation resistance is so low and its input capacitance so large that it is difficult to avoid dissipating most of the accepted power in the antenna structure or the matching network for such a short antenna. The half-wave dipole has a more reasonable radiation resistance and higher radiation efficiency than the short dipole. As a general rule, antennas need to be comparable to the wavelength in size to have a decent radiation resistance and radiation efficiency.

5.6 Near Field and Far Field

Antennas are defined in terms of gain and directivity, both of which are measured by examining the radiated field far from the antenna. In free space, the power density far from the antenna falls off according to the inverse square law. That is, the intensity falls off according to the square of the distance ($1/r^2$), as in

$$S = \frac{P_{\text{rad}}}{4\pi r^2} \quad (5.29)$$

where P_{rad} is the total power radiated by the antenna and *S* is the power density at some point a distance *r* from the antenna, with *r* large enough to be in the far field of the antenna. In the far field of a transmitting antenna, the time-average power density in W/m^2 is related to the radiated electric field by

$$S = \frac{|E|^2}{2\eta} \quad (5.30)$$

where *E* is the electric field intensity expressed as a phasor and
$\eta \simeq 376.7\Omega$
is the characteristic impedance of space.

The near field of the antenna normally has a higher power density than the far field because of its proximity to the antenna radiator element, but it diminishes very rapidly with increasing distance according to a
$1/r^4$
function. The distance to the edge of the near field is a function of both the wavelength of the radiated signals and the antenna dimensions:

$$r_{\min} = \frac{2d^2}{\lambda} \quad (5.31)$$

where
$r_{\min}$
is the distance where the fields transition from near to far field behavior, *d* is the largest antenna dimension, and *λ* is the wavelength at the operating frequency. Energy transfer in the near field of an antenna is highly reactive, consisting predominantly of large circulating amounts of electric and magnetic energy close to the radiator being exchanged with the radiating antenna throughout each cycle of the oscillating fields.

Example 5.2 An antenna with a length of 6 cm radiates a 12-cm wavelength signal. Calculate the near-field distance.

Solution

$$\begin{aligned} r_{\min} &= \frac{2d^2}{\lambda} \\ &= \frac{(2)\ (6\ \text{cm})^2}{12\ \text{cm}} \\ &= 6\ \text{cm} \end{aligned}$$

Closer to the antenna than this distance, the near fields around the antenna are fairly complicated. Beyond $r_{\min}$, we are in the far field of the antenna, and the behavior of the radiated fields simplifies and the power density of the radiated fields decays according to the inverse square law of Eq. (5.29).

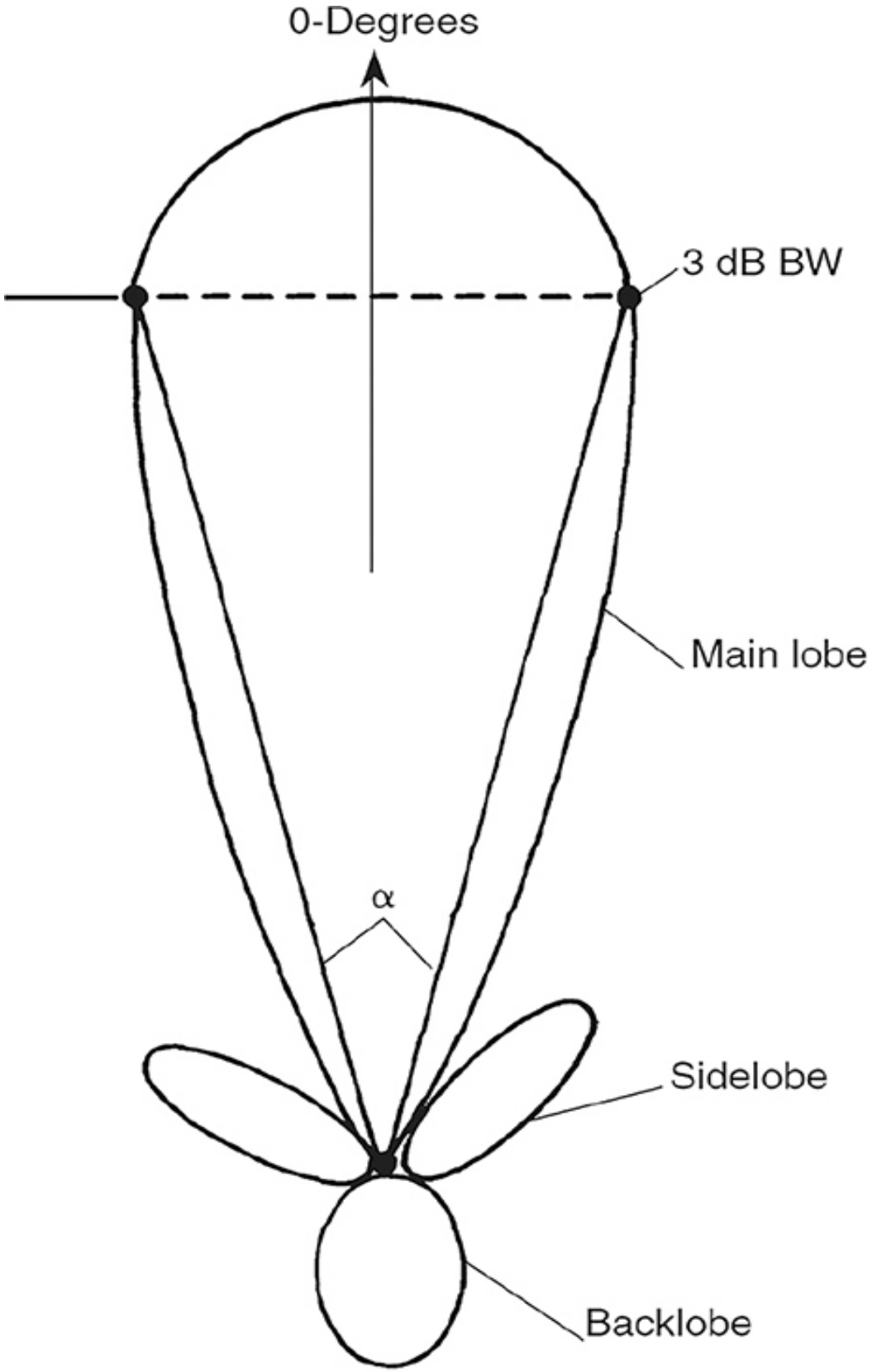
5.7 Antenna Pattern

The angular dependence of the far field radiated by an antenna is the antenna pattern or radiation pattern. The radiation pattern is defined by

$$f(\theta, \phi) = \frac{S(r, \theta, \phi)}{S_{\max}} \quad (5.32)$$

where $S(r, \theta, \phi)$ is the radiated power density at some point in the far field represented by the distance r from the antenna, the angle θ down from vertical, and the azimuth angle ϕ. $S_{\max}$ is the maximum power density over all angles θ and ϕ on the sphere of radius r around the antenna.

The antenna pattern is a function of angle relative to the antenna. We often plot radiation patterns as 1D cuts with one of the two angles fixed. Two methods for plotting antenna patterns are shown in Figs. 5.15 and 5.16. Figure 5.15 is a polar plot viewed from above, oriented such that the main lobe is centered on zero degrees. The plot of Fig. 5.16 is a rectangular method for displaying the same information.



a = 3-dB beamwidth

FIGURE 5.15 Directional antenna radiation pattern in polar form (top view).

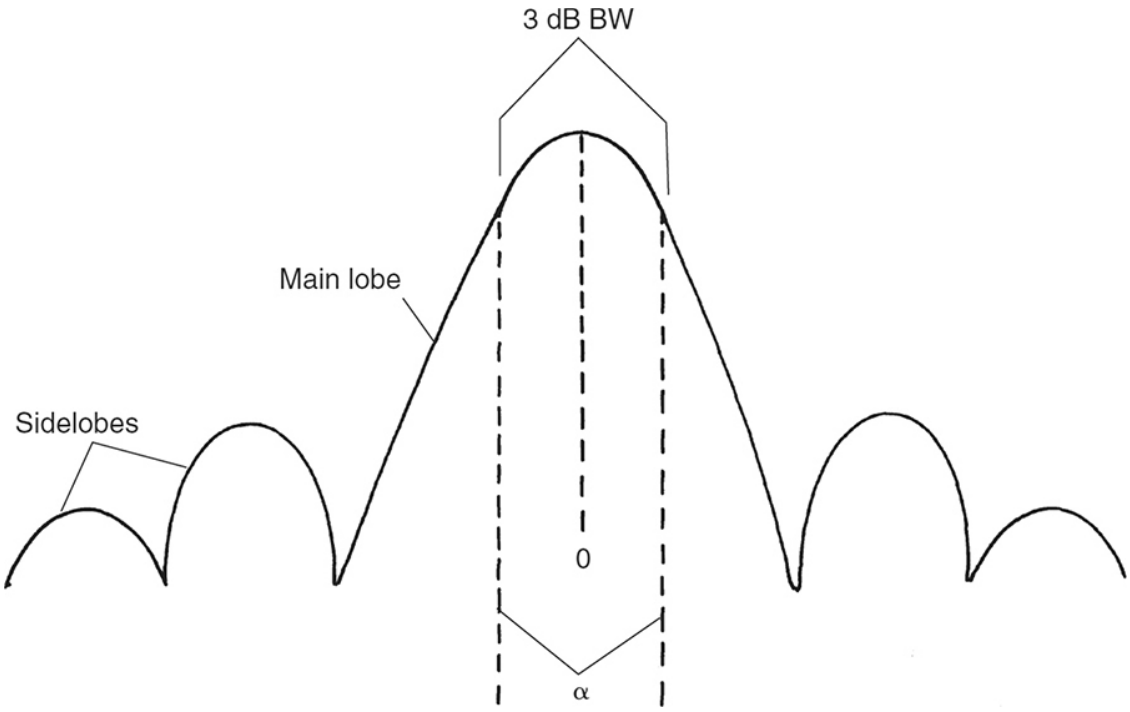


FIGURE 5.16 Radiation pattern plotted in a rectangular format.

A common method for specifying antenna directivity is the beamwidth. The angular displacement between points on the main lobe (see Figs. 5.15 and 5.16) where the power density drops to one-half (−3 dB) of its maximum at the peak of the lobe is the 3 dB beamwidth or half-power beamwidth (HPBW), shown as *a* in Fig. 5.15. On occasion, the null-to-null beamwidth of an antenna may be given, but normally it is understood that beamwidth is measured between the half-power points.

In addition to a main lobe, most antennas have additional minor lobes, such as the sidelobes and backlobe shown in Fig. 5.15. These can be inherent to the antenna itself or caused by structures and objects near the antenna. Sidelobes are generally undesirable. They can be minimized by careful design but often are impossible to entirely eliminate, particularly for electrically large antennas with

complex radiation patterns. For a given power density required at a distant receiver site, the transmitter must supply additional power to make up for power radiated into the sidelobes. In radar systems, for example, large sidelobes translate to errors in detecting target bearings. If a sidelobe is strong enough to detect a target, then the radar display will show the off-axis target as though it were in the main lobe of the antenna, resulting in an azimuth error.

5.8 Gain and Directivity

The directivity of an antenna is a measure of its ability to concentrate RF energy in a subset of directions rather than in all directions equally. Directivity is defined to be the ratio of the power density radiated by an antenna to the power density of an isotropic radiator with the same total radiated power, so that

$$D = \frac{S}{P_{\text{rad}}/(4\pi r^2)} \quad (5.33)$$

where S is the radiated power density in a given direction in W/m^2, r is the distance from the antenna, and P_{rad} is the power radiated by the antenna in Watts.

Directivity and gain for antennas are closely related. Gain is similar to directivity, but it includes the radiation efficiency of the antenna. Gain and directivity are related by

$$G = \eta_{\text{rad}} D \quad (5.34)$$

where η_{rad} is the radiation efficiency of the antenna. A highly directive antenna will have a large directivity, but if ohmic losses are large and the radiation efficiency is poor, the gain is smaller and the antenna may not be very useful for applications requiring high gain.

Gain can also be defined by a formula similar to Eq. (5.33) but with

P_{rad}

replaced by the power

P_{in}

accepted by the antenna from the transmitter power amplifier. By further replacing

P_{in}

with the available power from the transmitter, we can define the realized gain of the antenna. Of these antenna parameters, gain is the most common, but when the antenna loss is small, unknown, or unimportant for the application, directivity is used. Gain is a function of angle around the antenna, and technically Eq. (5.34) is the gain pattern. When just the term gain is used for an antenna, we often mean the peak value of the gain pattern or the peak gain.

Directivity and gain are associated with the radiation pattern of an antenna in three dimensions. We often present plots of gain or directivity in a single plane, such as a view from directly above the antenna that shows the directivity in the horizontal plane. But the radiated signal also varies with elevation angle above the ground. There is an inverse relationship between beamwidth and gain. Wide beamwidth means low gain, and narrow main beamwidth normally means high gain.

For some applications, low-gain antennas are desirable. For a handheld mobile device, the antenna should have a nearly isotropic radiation pattern so that the device can be held in any orientation and still maintain a good link to a base station. For long-range communication links, high-gain antennas are almost always used. Directional antennas primarily increase the level of the signal transmitted to or received from a given desired direction, but they also can reduce unwanted noise and mutual interference from neighboring cochannel stations.

5.9 Effective Isotropic Radiated Power

Antenna and transmitter systems are often rated in terms of effective isotropic radiated power (EIRP). EIRP is the product of the transmitter power and the antenna gain:

$$\text{EIRP} = GP_{\text{in}} \quad (5.35)$$

where
P_{in}
is the transmitter RF power delivered to antenna terminals. If an antenna has a gain of 3 dB, the EIRP is twice the transmitter output power. A 100 W output transmitter connected to a 3-dB antenna will produce a power density at a distant receiver equal to a 200 W transmitter feeding an isotropic radiator.

5.10 Reciprocity

A component or system that passes energy from one port to another equally well in either direction is reciprocal. Within limits, most physical systems are reciprocal, unless it includes an energized material such as a biased ferrite or semiconductor. The lower atmosphere is reciprocal, for example, but the ionosphere, with charged particles moving in the earth's magnetic field, is nonreciprocal.

Antennas are typically reciprocal. This means that they can be used as both transmitters or receivers. The radiation pattern of an antenna measured as a transmitter is nearly always identical to its receiving pattern. Reciprocity can break down in certain cases. High transmit power can damage an underdesigned antenna. If the antenna includes active components, such as amplifiers, then it is no longer reciprocal and only works for its designed operating mode. But as long as the antenna as a system is reciprocal, which is almost always the case, once we measure or compute the gain of an antenna, which is defined in its transmit mode of operation, we know its receiving properties as well.

5.11 Receiving Antennas and Effective Area

As we saw in Sec. 2.7, the receiving properties of an antenna are characterized by the effective area or effective aperture. The effective area for an antenna with a given gain is

$$A_{\mathrm{e}} = \frac{\lambda^2}{4\pi} G \quad (5.36)$$

The effective area can be larger than the physical extent of the antenna for a small dipole type antenna. Or, it can be close to or slightly smaller than the actual physical area for an aperture type antenna.

With the effective area, we can find the power received by the antenna for a load that is matched to the antenna impedance. This is given by Eq. (2.7):

$$P_L = S_{\mathrm{inc}} A_e \quad (5.37)$$

where
S_{inc}
is the power density of the radio wave incident on the receiving antenna in
$\mathrm{W/m^2}$
. If the load on the receiving antenna is not matched to the antenna impedance, then the power is smaller than this value by an impedance mismatch factor that can be calculated from the antenna and load impedances.

5.12 Friis Transmission Equation and Link Budgets

In Sec. 2.7, we gave a formula for the power at a receiver for a link between two isotropic radiators. With a simple modification, we can extend this formula for a link with antennas of any type. This is the Friis transmission equation,

$$P_L = P_t G_t G_r \left(\frac{\lambda}{4\pi r} \right)^2 \quad (5.38)$$

where
P_t
is the power accepted by the transmitting antenna,
G_t
is the gain of the transmitting antenna,
G_r
is the receive antenna gain, and *r* is the separation between the antennas. The factor in parentheses is the free space path loss. This formula only holds in free space and must be modified for propagation over terrain or other complex propagation environments.

The Friis transmission equation is the basis for a link budget that includes all of the factors that increase or decrease the signal strength as it propagates from the transmitter to the receiver. To facilitate this, the equation is often used in dB form:

$$P_L\,(\mathrm{dBm}) = P_t\,(\mathrm{dBm}) + G_t\,(\mathrm{dB}) + G_r\,(\mathrm{dB}) - L\,(\mathrm{dB}) \quad (5.39)$$

where *L* is the free space path loss in dB given by Eq. (2.9) and the powers are given in dB relative to a convenient reference such as 1 mW (dBm). This allows additional effects on the communication link, such as antenna mispointing, precipitation, or atmospheric attenuation to be included in a table format from which the signal power at the receiver can be conveniently obtained.

5.13 Signal to Noise Ratio

The Friis equation deals only with the signal of interest in a communication link. If the path loss is high, the resulting low signal strength at a receiver can be overcome by adding an amplifier to the system. But the amplifier also raises the noise floor and adds additional noise of its own. What matters more than just the signal

level itself is the ratio of the signal power to the noise power or the signal to noise ratio (SNR) at the receiver,

$$\mathrm{SNR} = \frac{P_s}{P_n} \quad (5.40)$$

where
P_s
is the received signal power and
P_n
is the noise power at the receiver. In a communication system, the amplifiers and other electronic components in the receiver actually lower the SNR slightly. If the SNR is insufficient for good link quality, the transmit power or antenna gains must be increased or sources of noise eliminated or reduced. At the receiver, once noise is added and the SNR at the receiver output determined, the SNR can only be improved with sophisticated noise cancellation techniques that are generally impractical for a typical radio link or a low SNR can be dealt with by coding techniques that can recover information from a signal of interest from a noisy communication channel.

The signal power can be computed from the transmitter power, transmit antenna gain, the path loss over the distance between the antennas, the receive antenna gain, and any impedance mismatches in the system. The noise power consists of galactic radiation (particularly below around 10 MHz), human-caused radio noise sources, interference from other radio services, thermal noise from ground and buildings and other warm objects around the receive antenna, and noise contributed by electronics in the receiver. Maximizing signal power and minimizing noise is the goal of most of the design considerations for antennas in this book.

5.13.1 SNR and Radiation Efficiency

The SNR in a radio link is influenced by the radiation efficiency of the antennas. At the transmit side, poor radiation efficiency means

that power is absorbed by the antenna. This translates directly to lower signal at the receiver and a decrease in SNR.

On receive, how SNR is affected by the efficiency of an antenna is more complicated. For some applications, SNR is strongly affected by antenna radiation efficiency, whereas for others efficiency is less important. Low radiation efficiency means that losses in the antenna structure due to metal components and dielectric materials are high. Losses in the receiving antenna add thermal noise, which can decrease the SNR depending on the level of noise from other sources in the system, such as the external environment around the receiving antenna.

For low-frequency receiving antennas operating below around 10 MHz, noise from the environment and galactic emissions are high enough that making the antenna more efficient has only a modest impact on the SNR at the antenna output. At these frequencies, it is difficult to make the antenna large enough to be comparable to the wavelength anyway, so antenna efficiencies in some cases can be as low as a few percent without materially affecting SNR. The antennas for mobile devices used in noisy terrestrial environments at microwave frequencies also do not need to be particularly efficient and radiation efficiencies of around 70 percent are common.

Antennas used for satellite communications operate in an environment with low external noise, so the additional thermal noise added by losses in a receive antenna can have a significant impact on SNR. In applications where SNR is at a premium radiation efficiencies can be well into the high 90 percent range and in many cases close enough to 100 percent that it is difficult to measure the small antenna loss. This is the case for the typical feed horn and parabolic reflector antenna used for a satellite ground terminal.

5.14 Electrically Small Antennas and Bandwidth Limitations

One of the less intuitive aspects of antennas is the relationship between antenna size and bandwidth. For many applications, it is desirable for an antenna to be as small as possible. This is the case for portable radios and handheld wireless devices which may contain a dozen or more antennas for different services. We have already seen that due to low feedpoint input resistance, electrically small antennas (i.e., antennas that are small compared to the wavelength at the operating frequency) can have low radiation efficiency. But low efficiency is not the only issue with small antennas.

A small antenna tends to be highly reactive. A small dipole, as we have seen, is essentially a capacitor and has a very low radiation resistance. The input impedance is dominated by a large, negative imaginary part. In order for a transmitter power amplifier to drive substantial power into the antenna, a matching network is required. In the case of the capacitative small dipole, the matching network is a tuning inductor in series with the antenna feed port.

The antenna and matching network consists of the capacitative antenna, its small input resistance, and the inductor. This is an RLC circuit as discussed in Sec. 5.1. With a small antenna input resistance in series with the antenna reactance, the qualify factor (Q) of the antenna and matching network is large. Consequently, the bandwidth of the matched antenna is small, making it usable only for narrowband communications and low data rates. If the antenna is lossy, that decreases the Q of the antenna and matching circuit, increasing the bandwidth, but the loss dissipates some of the power introduced to the antenna and the radiated power may be insufficient. In the technical literature, there are mathematical formulas that give lower bounds, often referred to as Wheeler-Chu bounds, for the quality factor and achievable bandwidth for a small antenna based on its electrical size and radiation efficiency.

CHAPTER 6
Transmission Lines and Impedance Matching

Transmission lines and waveguides are conduits for guiding radio frequency (RF) energy between elements of a radio system. In a typical station with a transmitter and a power amplifier, one transmission line carries RF power to an amplifier input and a second line delivers transmitter output energy from the amplifier to the antenna. A third line may carry the incoming RF energy from a transmit/receive switch or separate receiving antenna to the station receiver. Still other lines may switch filters in and out, or pass transmitter RF through signal monitors, power meters, and other station accessories.

The objective is to transfer as much energy as possible between the source and load and minimize the unintentional loss of energy in the transmission line or waveguide. The longer the distance between the source and the radiating unit(s), the greater the potential for loss of energy in the distribution system.

The geometry of a transmission line is designed to minimize the tendency of the structure to act as an antenna and radiate energy rather than guiding it to the load. In contrast, the geometry of an antenna has been designed to maximize its tendency to radiate. Some antennas are simply poor transmission lines. A classic example is the leaky wave antenna.

Many transmission lines have two or more conductors, but waveguides use an enclosed structure with only one contiguous conductor to guide RF energy. Transmission lines with two or more conductors can operate at low frequencies down to DC (zero frequency). Single conductor waveguides only guide energy above a given frequency that depends on the size and shape of the structure.

If we know the geometry of the conductors in a transmission line and the dielectric properties of the material that surrounds the conductors, we can use software tools or equations in this chapter to calculate the parameters that govern its operation in a system. These include the phase velocity and characteristic impedance, or Z_0, of the line. To reduce reflections at junctions in an RF system, components that are joined together normally have the same characteristic impedance. This is called impedance matching.

Standard values for transmission line characteristic impedances have evolved over the years. Microwave systems, test instruments, and many radio installations utilize a system impedance of 50 Ω. Cable television systems often use 75 Ω. On occasion, 300-, 450-, and 600-Ω lines are used.

6.1 Types of Transmission Lines

A simple though poorly performing transmission line is a single metal wire between the transmitter and some point on a horizontal antenna. The return path for this type of line is the earth or other ground path between the antenna and the transmitter. This transmission line exhibits many undesirable characteristics and it has been supplanted by two-conductor lines with controlled geometries.

6.1.1 Parallel Conductor Lines

Another basic configuration is the parallel conductor line shown in Figs. 6.1 through 6.5. Two conductors, both of diameter *d*, are separated by a dielectric (or air) at a spacing *S*. Parallel lines have been used at VLF, MF, and HF for over a century. Even antennas for the lower VHF bands are often fed with parallel lines. At higher frequencies, better transmission lines are available.

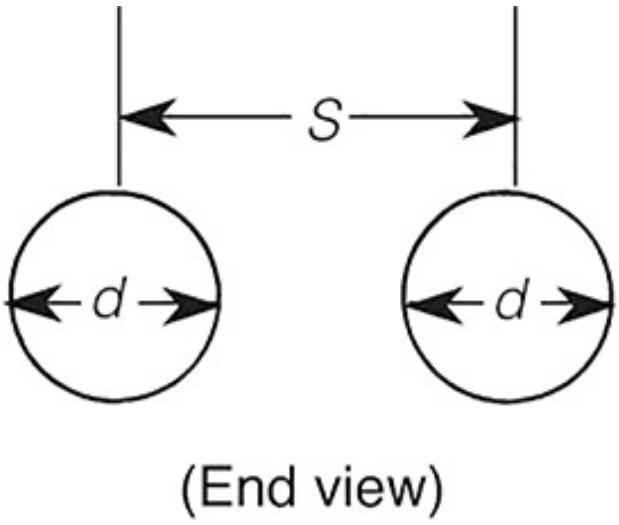


FIGURE 6.1 Parallel line transmission line (end view).

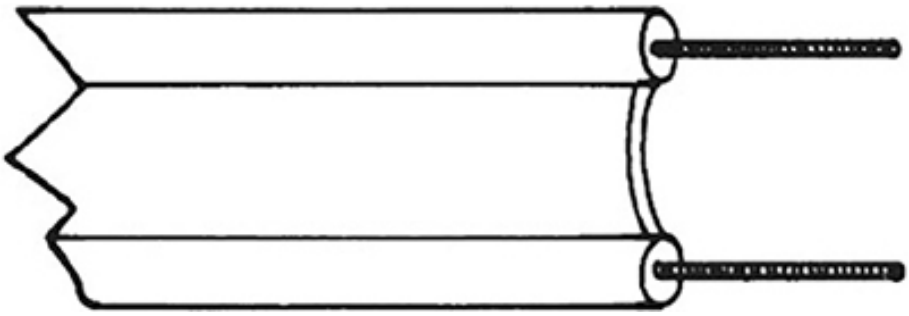


FIGURE 6.2 Parallel open-wire line construction details.

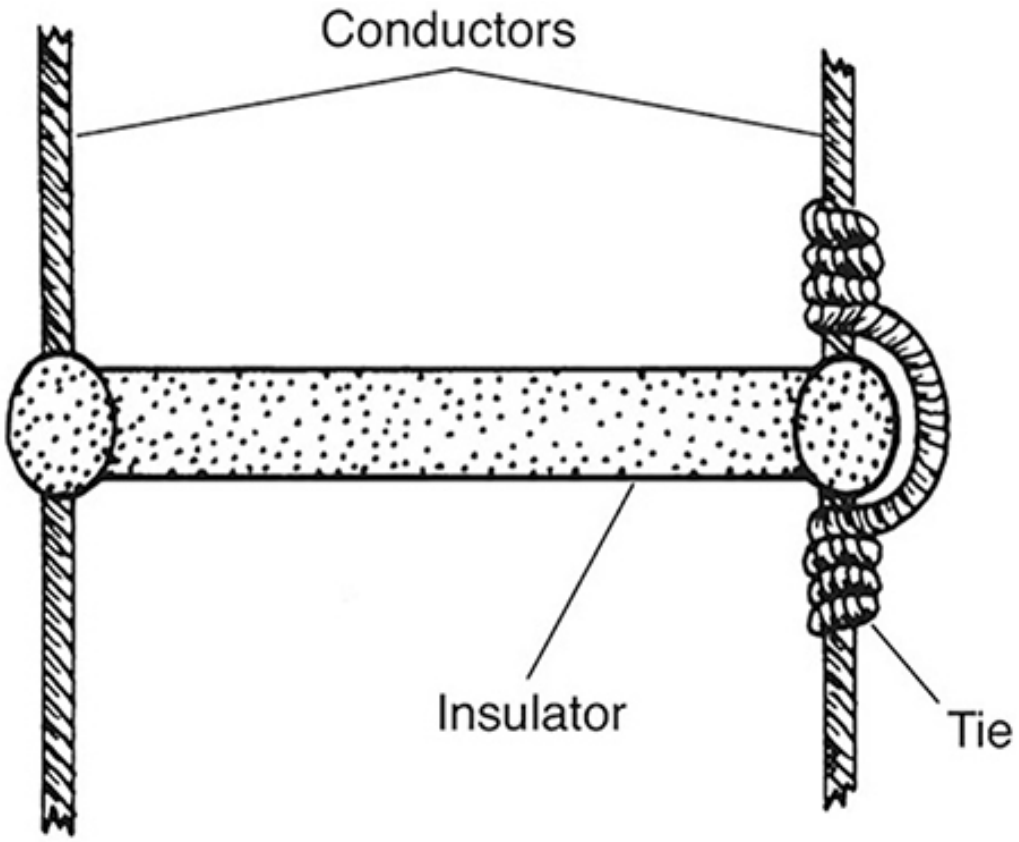


FIGURE 6.3 Twin-lead transmission line.

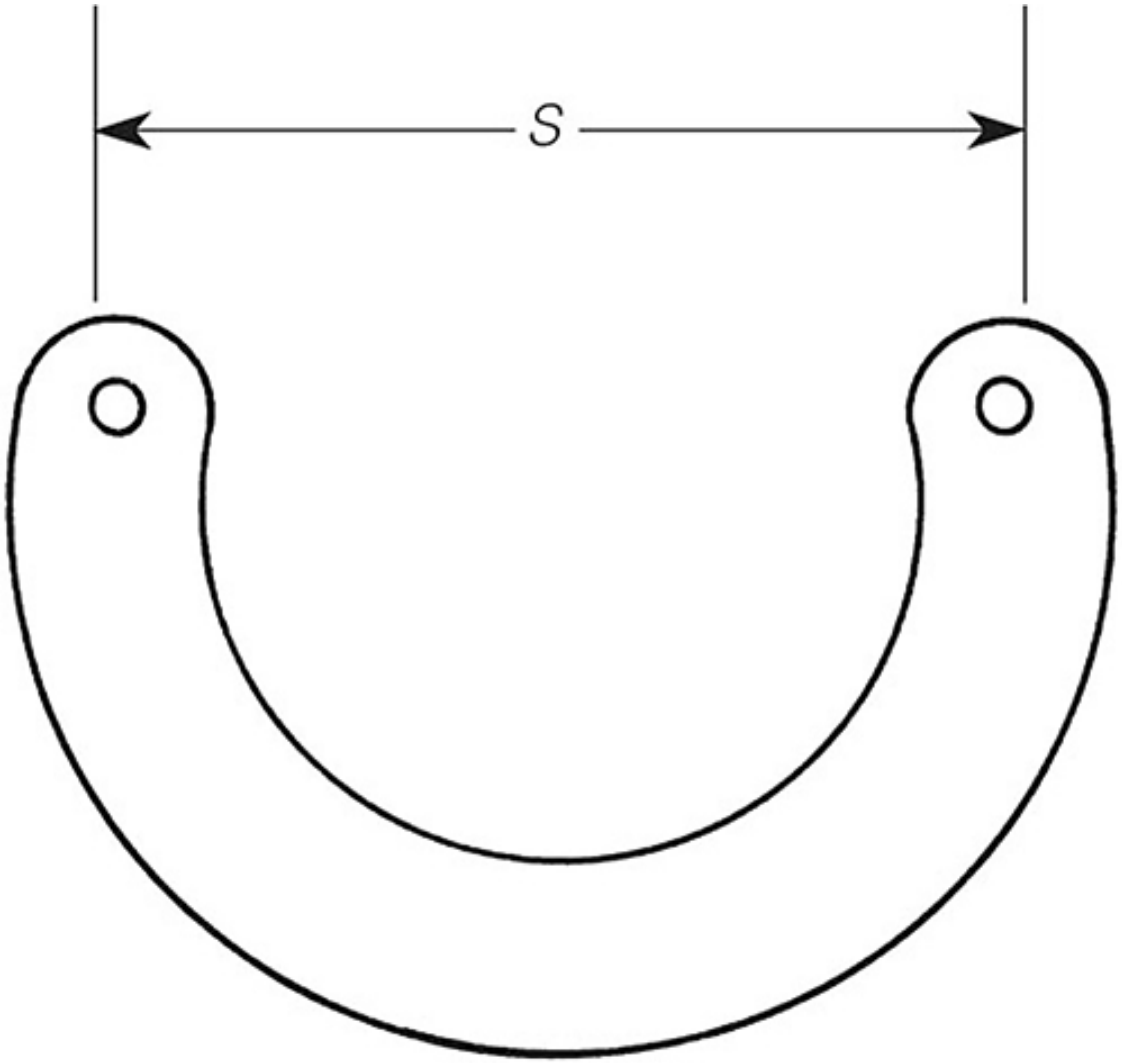


FIGURE 6.4 Shielded twin-lead transmission line.

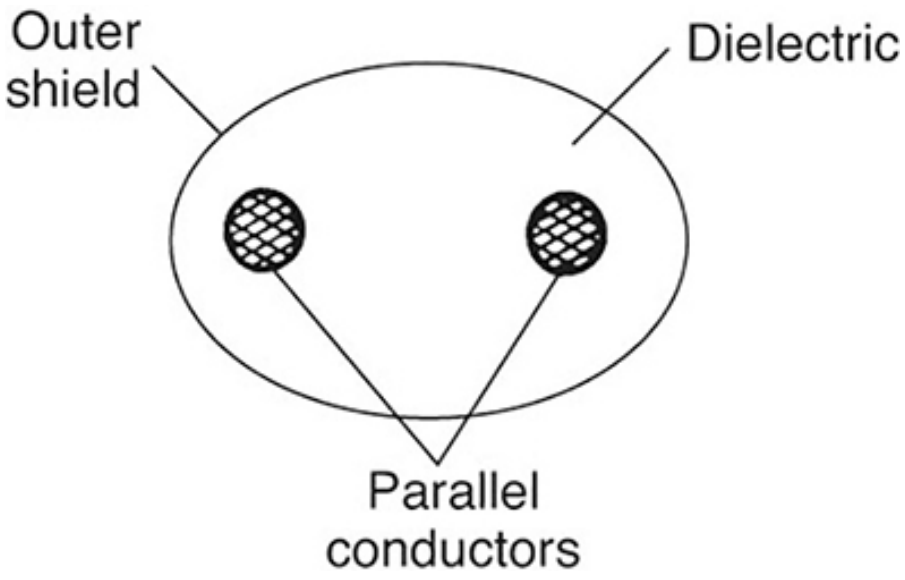


FIGURE 6.5 Horseshoe parallel line spreader.

Early parallel lines were open-wire line (OWL), as shown in Fig. 6.2. Here, the wire conductors are separated by an air dielectric; spacing between the two wires is maintained by rigid ceramic or plastic insulators.

Figure 6.3 shows a type of parallel line called twin-lead. This was used for analog television antennas. It consists of two parallel conductors encased in a flexible plastic dielectric. Twin-lead is designed for a characteristic impedance of 300 Ω to simplify impedance matching to a folded dipole broadcast antenna.

Figure 6.4 shows a form of parallel line called shielded twin-lead. This type of line uses the same structure as TV-type twin-lead, but adds a shield layer surrounding it. The shield may be braided, a thin layer of aluminum foil, or both. This reduces susceptibility to electrical noise and interference.

Another type of spacer is shown in Fig. 6.5. Generally formed from plastic or ceramic, its U shape reduces losses, especially in wet weather, by providing increased leakage path length between the two conductors relative to their spacing.

Multiwire parallel line is occasionally used. Four-wire and six-wire transmission lines can help when the distance between transmitter and antenna is extremely long (i.e., a substantial portion of a mile).

6.1.2 Coaxial Lines

Another common type of transmission line is coaxial cable (Figs. 6.6 through 6.12), often abbreviated as coax (pronounced "co-ax"). This form of line consists of two cylindrical conductors sharing a common axis (hence "coaxial"), and separated by a dielectric (Fig. 6.6). The outer conducting cylinder is hollow and the inner one is usually solid. The dielectric may be polyethylene or polyethylene foam or at higher frequencies teflon and other specialized low-loss dielectrics. The dielectric normally fills the space between the two conductors over the entire length of the cable.

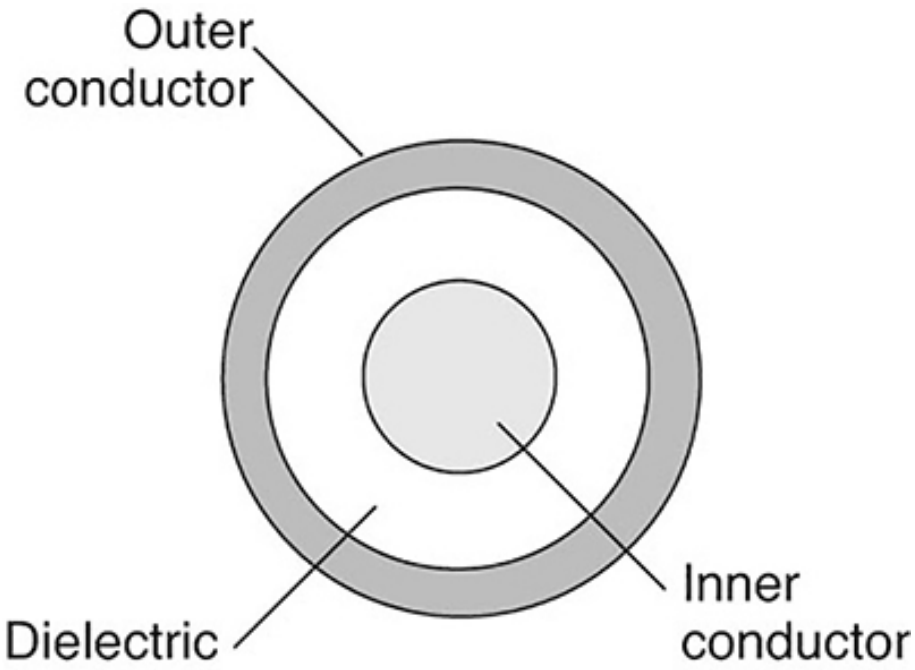


FIGURE 6.6 Coaxial cable (end view).

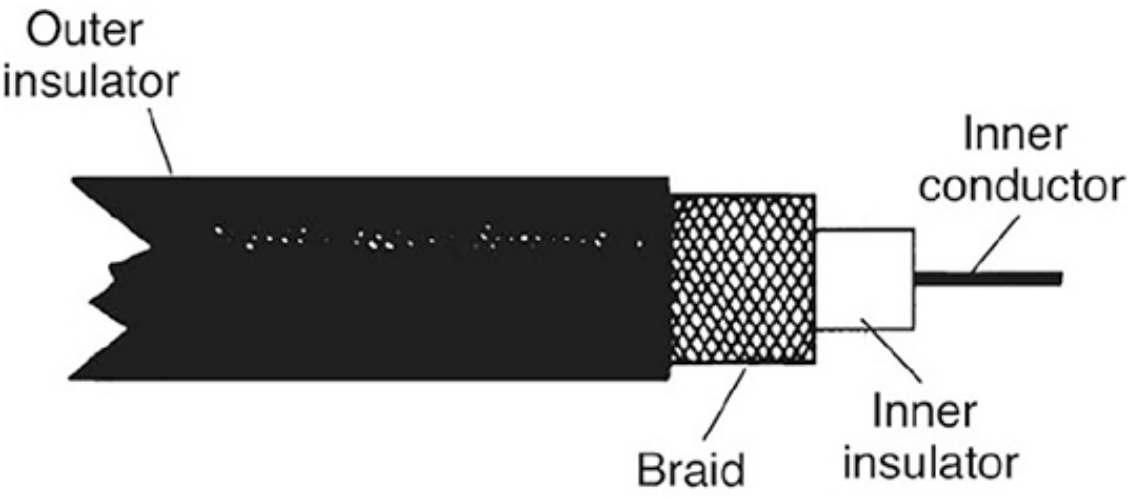


FIGURE 6.7 Coaxial cable (side view).

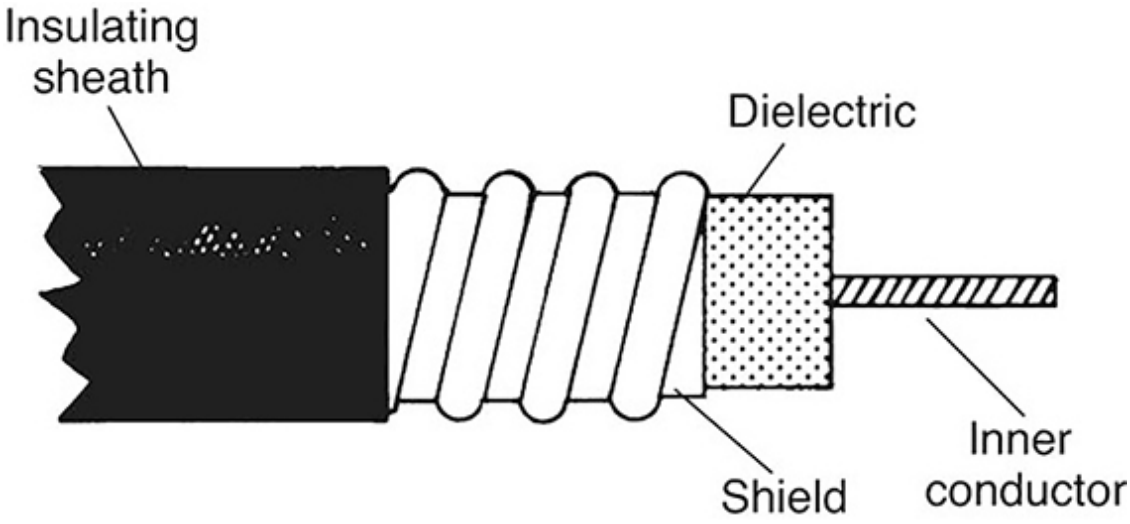


FIGURE 6.8 Helical coaxial cable.

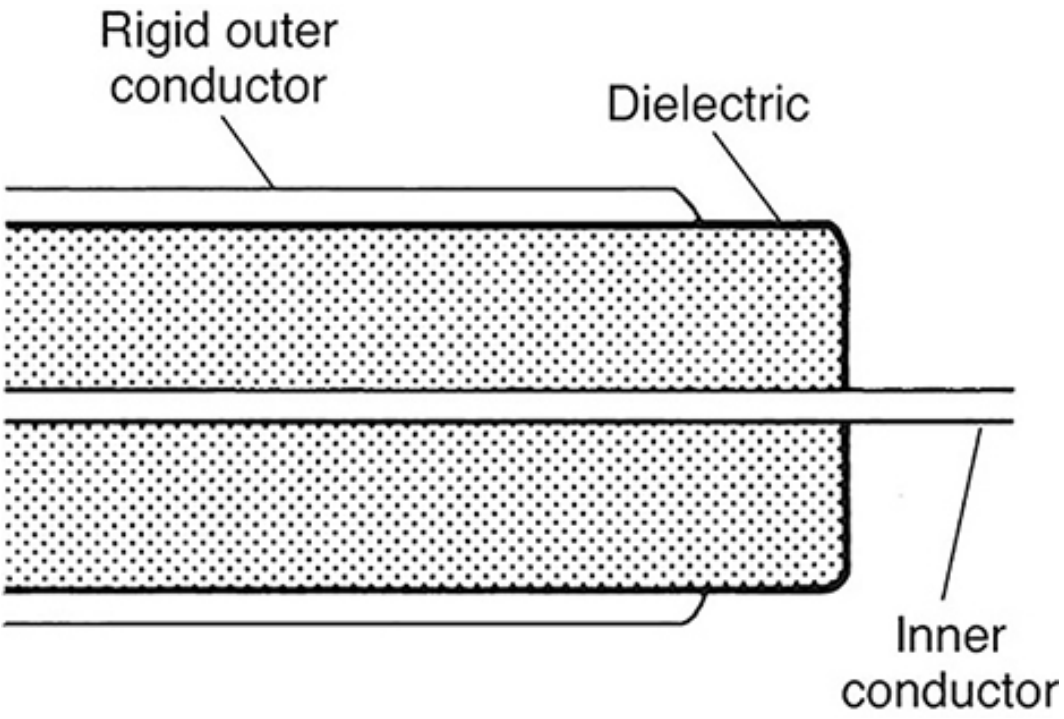


FIGURE 6.9 Rigid coaxial line (hardline).

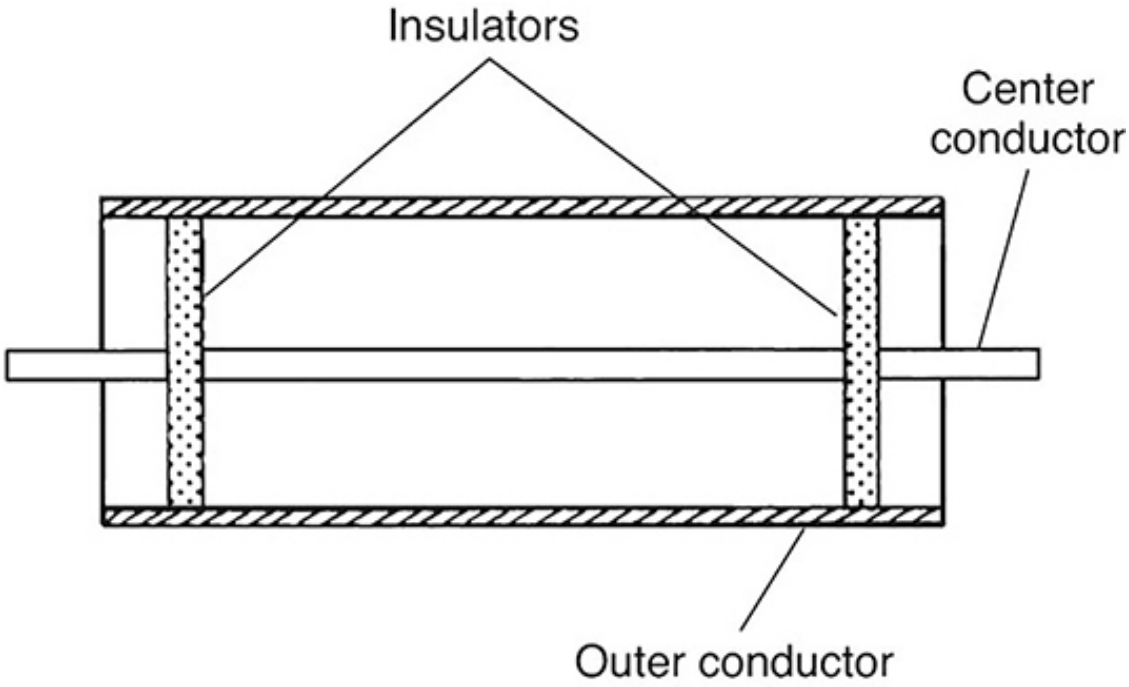


FIGURE 6.10 Gas-filled hollow coaxial line.

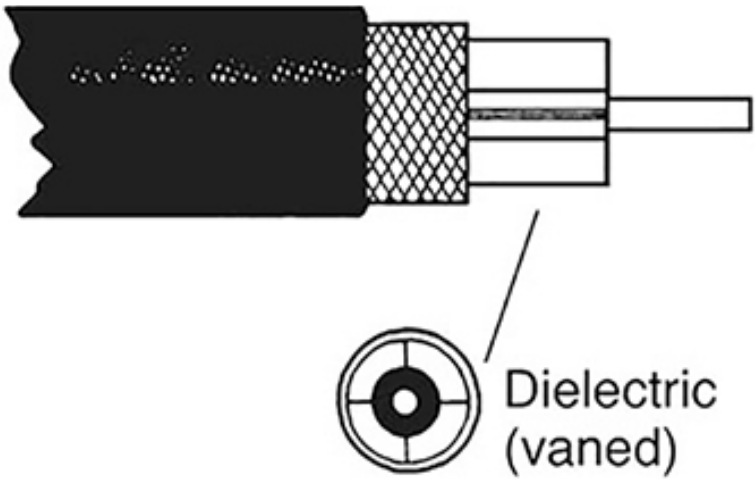


FIGURE 6.11 Articulated coaxial line.

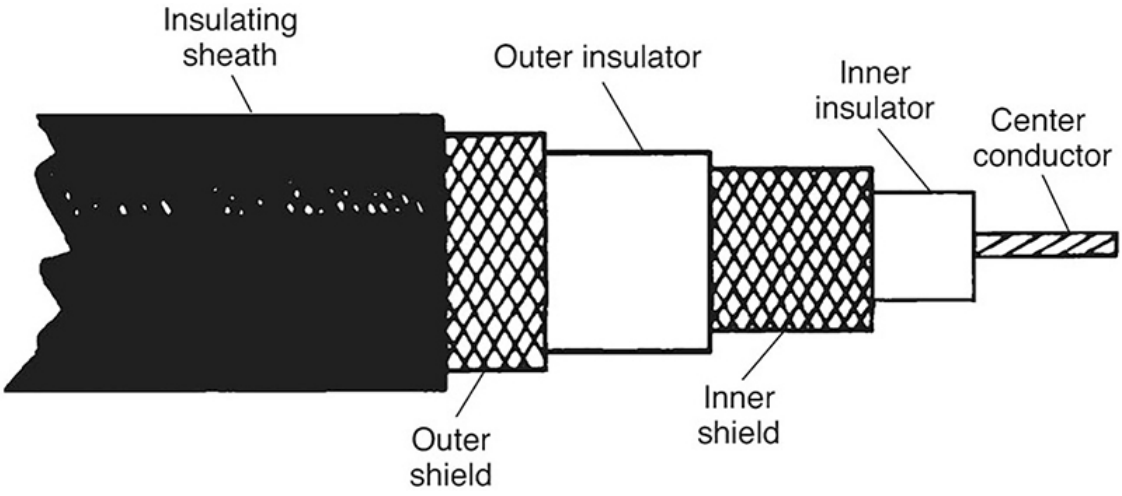


FIGURE 6.12 Double-shielded coaxial line.

Flexible coaxial cable, characterized by the RG-8, RG-58, RG-59, and RG-213 families, is the most common form. The outer conductor is made of braid or foil (Fig. 6.7). Another form of flexible or semiflexible coaxial line is helical line (Fig. 6.8), in which the outer conductor is spiral wound. Hardline (Fig. 6.9) is coaxial cable that uses a thin-wall aluminum tube as the outer conductor. Some hardline used at microwave frequencies has a rigid outer conductor and a solid dielectric. Semi-rigid coax provides mechanical stiffness but can be formed into bends and loops.

For extremely low loss requirements circular discs are spaced along the inside of the line to keep the inner conductor centered with respect to the outer conductor and the line is often filled with dry air or dry nitrogen and kept slightly pressurized (Fig. 6.10).

Some flexible microwave coaxial cable uses a solid "air-articulated" dielectric (Fig. 6.11), in which the inner insulator is not continuous around the center conductor but is ridged. Reduced dielectric losses increase the usefulness of the cable at higher frequencies. Double-shielded coaxial cable (Fig. 6.12) provides an extra measure of protection against radiation from the line, as well as from electromagnetic interference (EMI) from outside sources getting into the system.

6.1.3 Microstrip, Stripline, and Coplanar Waveguide

Microstrip line (Fig. 6.13) is a form of transmission line used at high UHF and microwave frequencies. Microstrip consists of a critically sized conductor over a ground-plane conductor, and separated from it by a dielectric. Striplines are sandwiched between two ground planes and are separated from each by the dielectric board. Coplanar waveguide is fabricated by milling or etching two gaps between a central conductor and the surrounding ground plane. These types of transmission lines are fabricated on layered structures and are used in PCBs and integrated circuits.

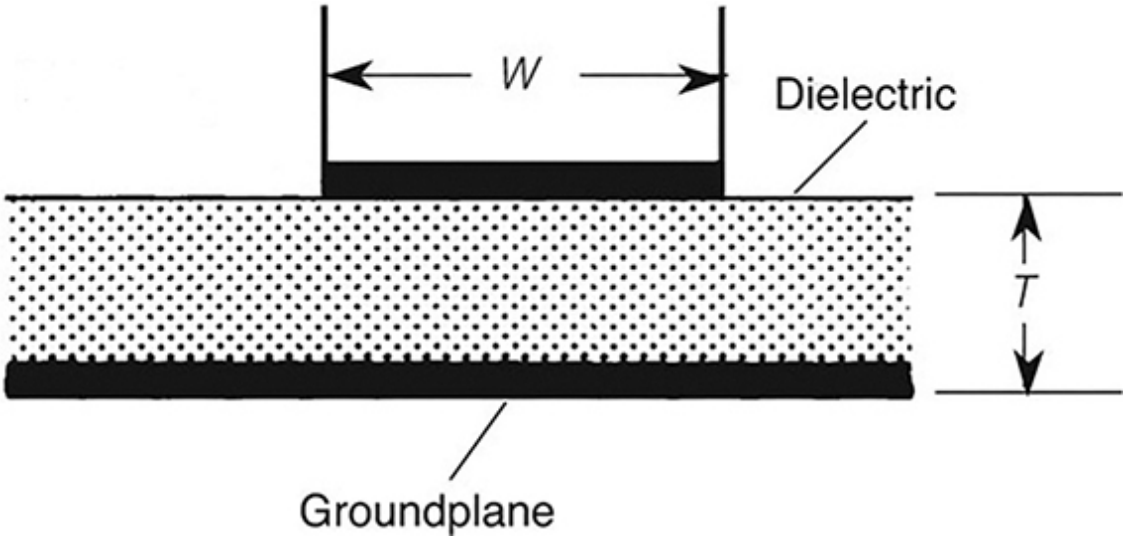


FIGURE 6.13 Microstrip transmission line used in printed circuit boards.

6.2 Lumped Element Model

If we take a short length of two-conductor transmission line of any kind and measure its electrical characteristics at some operating frequency with simple test equipment, we will discover four parameters:

- Each of the two conductors has a small series inductance between its two ends. All wires—even straight ones—have some inductance. We define L' as the sum of these inductances.
- Between the conductors we measure a small shunt capacitance, C'. This is a measure of the capacitive coupling between two closely spaced conductors including the dielectric material between them.
- Each of the two conductors exhibits a very small series resistance between its two ends, measured at the operating frequency. R' is the sum of the resistive losses in the two conductors.

- At the same end of the short length of transmission line we measure a high shunt resistance or a low shunt conductance, G', where G' is the reciprocal of the shunt resistance) across the two conductors. This is due to the small conductivity of the dielectric material filling the space between the two conductors. The shunt conductance may be increased by unintended effects like moisture in the line.

In defining these four parameters, we have primed each of them to indicate that the value we are measuring is for a unit length of transmission line. The length of transmission line used to obtain these parameters must be much shorter than the wavelength of the highest frequency we intend to send through the line.

If we now use these four parameters to create a lumped element model of this short length of transmission line, we obtain the equivalent circuit of Fig. 6.14. The transmission line can be modeled by chaining together many of these short lumped element circuits.

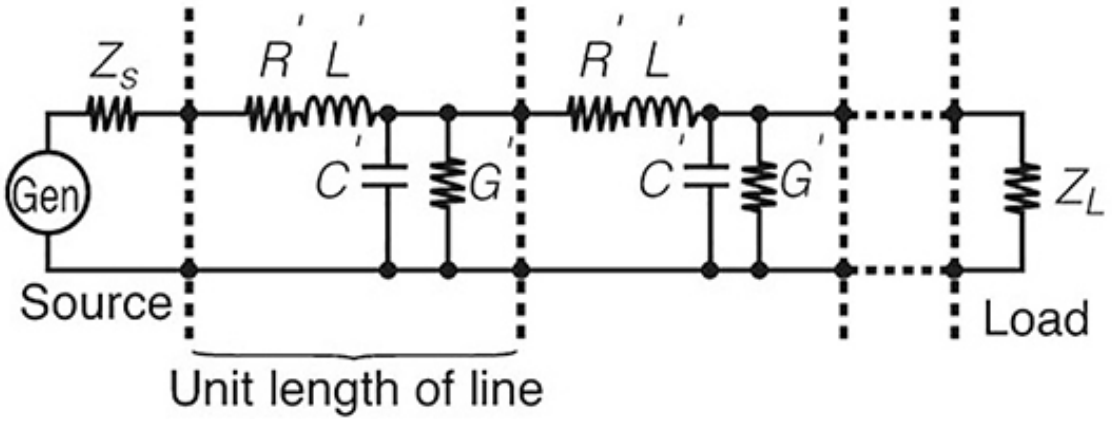


FIGURE 6.14 Transmission line lumped element model.

6.3 Characteristic Impedance

Using network analysis, we can use the lumped element model for a transmission line to derive differential equations for the current and voltage on the line. The equations can be solved to find formulas

for the current and voltage waves on the line. In the solution, the key parameter is the ratio between the current and voltage wave solutions on the transmission line. This is called the characteristic impedance
Z_0
.

The characteristic impedance is a function of the four lumped element parameters:

$$Z_0 = \sqrt{\frac{R' + j\omega L'}{G' + j\omega C'}} \quad (6.1)$$

where
R'
is the series resistance per unit length of two conductors,
G'
is the shunt conductance between two conductors per unit length,
L'
is the series inductance per unit length of two conductors,
C'
is the shunt capacitance per unit length between conductors, *j* is the imaginary unit
$\sqrt{-1}$
, and ω is the angular frequency in radians per second
$(\omega = 2\pi f)$
. The lumped element parameters are in turn a function of the physical geometry of the line, including conductor size, shape, and spacing, and the dielectric constant of the insulating material between the conductors, so the characteristic impedance is determined by the shape and composition of the transmission line.

In general,
Z_0
is complex and has both a resistive and a reactance component. The imaginary part is usually small and represents loss in the transmission line. The real part is more important for basic

transmission line analysis and gives the ratio between current and voltage waves on the line.

For certain values of R', G', L', C', and the operating frequency, the transmission is lossless and Z_0 is purely resistive or has a small enough imaginary part that we can approximate the line as lossless. Three such cases are:

- $R' = G' = 0$. In an ideal cable, the series resistance is zero and the shunt resistance is infinite, so (6.1) reduces to the following simplified form for a lossless cable:

$$Z_0 = \sqrt{\frac{L'}{C'}} \quad (6.2)$$

- $\omega L' >> R'$ and $\omega C' >> G'$. Although not zero, the series resistance and shunt conductance are negligible with respect to the series inductance and shunt capacitance at the operating frequency. Because the effective resistance of a wire at RF is modified by a frequency-dependent phenomenon called skin effect, R' is a function of frequency. It is not appropriate, for instance, to blindly use the DC or ohmic value of *R* at RF.
- $R'/L' = G'/C'$ and both R' and G' are small, but not necessarily negligible. If this condition is met, the line is distortionless or dispersionless even though it may not be lossless.

Example 6.1 A nearly lossless transmission line (R' and G' are very small) has an inductance of 3.75 nH per unit length and a capacitance of 1.5 pF per unit length. Find the characteristic impedance.

Solution

$$Z_0 = \sqrt{\frac{L'}{C'}} = \sqrt{\frac{3.75 \times 10^{-9}\ \mathrm{H}}{1.5 \times 10^{-12}\ \mathrm{F}}} = 50\ \Omega$$

The characteristic impedances of various kinds of two-conductor transmission lines are:
Parallel line:

$$Z_0 = \frac{276}{\sqrt{\epsilon}} \log_{10}\left(\frac{2S}{d}\right) \quad (6.3)$$

where ϵ is the dielectric constant of the insulator between the conductors, S is the center-to-center conductor spacing, and d is the conductor diameter.
Coaxial line:

$$Z_0 = \frac{138}{\sqrt{\epsilon}} \log_{10}\left(\frac{D}{d}\right) \quad (6.4)$$

where D is the inside diameter of the outer conductor and d is the inner conductor diameter.
Microstrip:

$$Z_0 = \frac{377}{\sqrt{\epsilon_t}} \frac{T}{W} \quad (6.5)$$

where
ϵ_t
is the relative dielectric constant of the printed circuit board (PCB) substrate, *T* is the substrate thickness, and *W* is the width of the conductor strip. The relative dielectric constant
ϵ_t
is a combination of the relative dielectric constant of the substrate material used in the PCB and the dielectric constant of the air above the PCB and becomes closer to the substrate relative dielectric constants for larger values of the ratio *W/T*.

Example 6.2 A microstrip transmission line is built on a 4-mm-thick PCB that has a relative dielectric constant of 5.5. Calculate the characteristic impedance if the width of the strip is 2 mm.

Solution

$$Z_0 = \frac{377}{\sqrt{\epsilon_t}} \frac{T}{W} = \frac{377}{\sqrt{5.5}} \left(\frac{4}{2} \right) = 321\ \Omega$$

We often need to design a transmission line to fit a specific system impedance (e.g., 50 Ω). We can make some choices of printed circuit material (hence, dielectric constant) and thickness, but even these are usually limited in practice by the availability of standardized boards. The microstrip width is the variable parameter. Equation (6.5) can be rearranged into the form

$$W = \frac{377\, T}{Z_0 \sqrt{\epsilon_t}} \quad (6.6)$$

A characteristic impedance of 50 Ω is accepted as standard for most RF and microwave systems. This value is a compromise between power-handling ability and loss. Peak power handling and low-loss operation may not occur at the same characteristic impedance. The maximum power-handling ability for coaxial cables occurs at 30 Ω

while the lowest loss occurs at 77 Ω. 50 Ω is a reasonable tradeoff. For cable TV distribution, the RF power levels are small and distribution lines are long, so a higher impedance (75 Ω) is used to take advantage of the reduced attenuation factor.

6.4 Phase Velocity

For many transmission lines, the velocity of a wave on the line is less than the free-space velocity (i.e., less than the speed of light in vacuum). The phase velocity on the line is related to the geometry of the conductors and the dielectric constant of the insulating material that separates the conductors. The velocity of wave propagation on a transmission line is

$$v = \frac{1}{\sqrt{L'C'}} \quad (6.7)$$

In the usual transmission line, μ is essentially the same as for free space, but ϵ will be somewhat larger than
ϵ_0
, with the result that v in the line will be less than the speed of propagation in free space. For lossless lines, the velocity of propagation on the line is

$$v = \frac{c_0}{\sqrt{\epsilon_r}} \quad (6.8)$$

The wavelength of the wave in the line is

$$\lambda = \frac{\lambda_0}{\epsilon_r} \quad (6.9)$$

The phase velocity can be specified as a number v (m/s) or as a decimal fraction
v_F
of

c_0
, the speed of light in a vacuum (approximately
$3 \times 10^8 \text{ m/s}$
). If the velocity factor of a transmission line is rated at 0.66, then the velocity of the wave is
$0.66c_0$
, or
$1.98 \times 10^8 \text{ m/s}$
.

For some transmission lines, the phase velocity depends on frequency. When that is the case, signals applied to one end of the line are said to suffer dispersion or frequency dispersion as they propagate along the line. A complex waveform consisting of multiple-frequency components will become distorted by the time it reaches the far end of the line, since the different frequencies will have experienced differing amounts of attenuation over their common path.

The phase velocity is used when designing transmission line transformers and other devices for which the length of the line is important. Transmission line length is often specified in terms of electrical length, which can be either an angular measurement (e.g., 180 degrees or *π* radians), or a value relative to the wavelength on the line (e.g., one half-wavelength, which is the same as 180 degrees). The phase velocity or velocity factor must be known before transmission lines can be selected and cut to length for specific situations.

Table 6.1 lists the velocity factor and characteristic impedance for common types of transmission lines. These are nominal values. For precise work the actual velocity factor for any given line can be measured, using techniques and instruments described in Chap. 30.

TABLE 6.1 Transmission Line Characteristics

Type of Line	Z_0 (ohms)	Velocity Factor

1/2-in TV parallel line (air dielectric)	300	0.95
1-in TV parallel line (air dielectric)	450	0.95
TV twin-lead	300	0.82
UHF TV twin-lead	300	0.80
Polyethylene coaxial cable	*	0.66
Polyethylene foam coaxial cable	*	0.79
Air-space polyethylene foam coaxial cable	*	0.86
Teflon coaxial cable	*	0.70
CATV hardline	75	0.8–0.9
* Various impedances depending upon cable type.		

6.5 Loss in Transmission Lines

Loss in a transmission line is modeled with the series resistance R' associated with the conductors and the shunt conductance G' due to the dielectric material. For an ideal transmission line, the conductor resistance is zero and the dielectric is a perfect insulator, so $R' = G' = 0$. When either R' or G' or both are large enough to matter, the transmission line must be modeled as lossy.

The behavior of propagating waves on a lossy transmission line can be characterized by a complex number, the propagation

constant:

$$\gamma = \sqrt{(R' + j\omega L')\,(G' + j\omega C')} \quad (6.10)$$

The real part of the propagation constant is the attenuation constant *α* and the imaginary part is the phase constant *β*:

$$\gamma = \alpha + j\beta \quad (6.11)$$

We can represent a sinusoidal (single-frequency) waveform on the line using

$$V(z,t) = V_0 e^{-\alpha z}\cos(\omega t - \beta z) \quad (6.12)$$

where *V*(*z*, *t*) is the amplitude of the wave as a function of its location along the line (the *z* axis) and the time of observation,
V_0
is the amplitude of the input waveform at the peak of its sine wave or cycle,
$e^{-\alpha z}$
is a decaying exponential term representing the effect of line loss,
$\omega t = 2\pi f t$
is the portion of the cosine argument that mathematically accounts for the sinusoidal or cyclic variation in the amplitude of the wave with the passage of time, and *βz* is the sinusoidal spatial variation in the amplitude of the wave along the line.

The real part *α* is associated with loss on the transmission line. *α* is called the attenuation constant of the transmission line, and it is a dimensionless number with units of nepers per meter (Np/m). If *α* is large, the line is highly lossy and a propagating wave decays quickly. For a lossless line, the attenuation constant is zero and a wave can propagate arbitrarily far along the line without decaying in amplitude. The attenuation constant for a given type of line is often expressed in dB/100 ft.

The wavenumber *β* is related to the wavelength on the line by
$\beta = 2\pi/\lambda$

. In terms of the velocity of propagation on the line, the wavenumber is

$$\beta = \frac{2\pi f}{v} \quad (6.13)$$

With this formula, together with the attenuation on the line, we can predict the variation in the amplitude of a wave of frequency *f* from the phase velocity or velocity factor of the transmission line.

6.6 Transmission Line Responses

When we employ a transmission line to deliver RF energy from a transmitter to an antenna, the signal is often narrowband and we can approximate the transmission line behavior by its response at a single frequency. This is the line's sinusoidal steady-state response.

In other cases, it is helpful to understand what happens when a transient propagates through a transmission line system. This is done using the line's step-function or transient response. Step-function analysis involves the application of a single voltage transition at the input of the line. The input voltage changes from zero (or an initial value) to a nonzero (or final) value instantaneously and is held there until all action along the entire transmission line has died out.

6.6.1 Step-Function Response of a Transmission Line

Before we get into the details of transmission line step response, let's look at a mechanical analogy that should be familiar to all of us. A taut rope (Fig. 6.15*a*) is tied to a rigid wall that does not absorb any of the energy in the pulse propagated down the rope. When the free end of the rope is given a vertical displacement (Fig. 6.15*b*) by yanking up on it once, a wave is propagated along the rope, toward the wall, at velocity *v* (Fig. 6.15*c*). When the pulse hits the wall (Fig. 6.15*d*), it is reflected (Fig. 6.15*e*) and propagates back along the rope toward the free end (Fig. 6.15*f*).

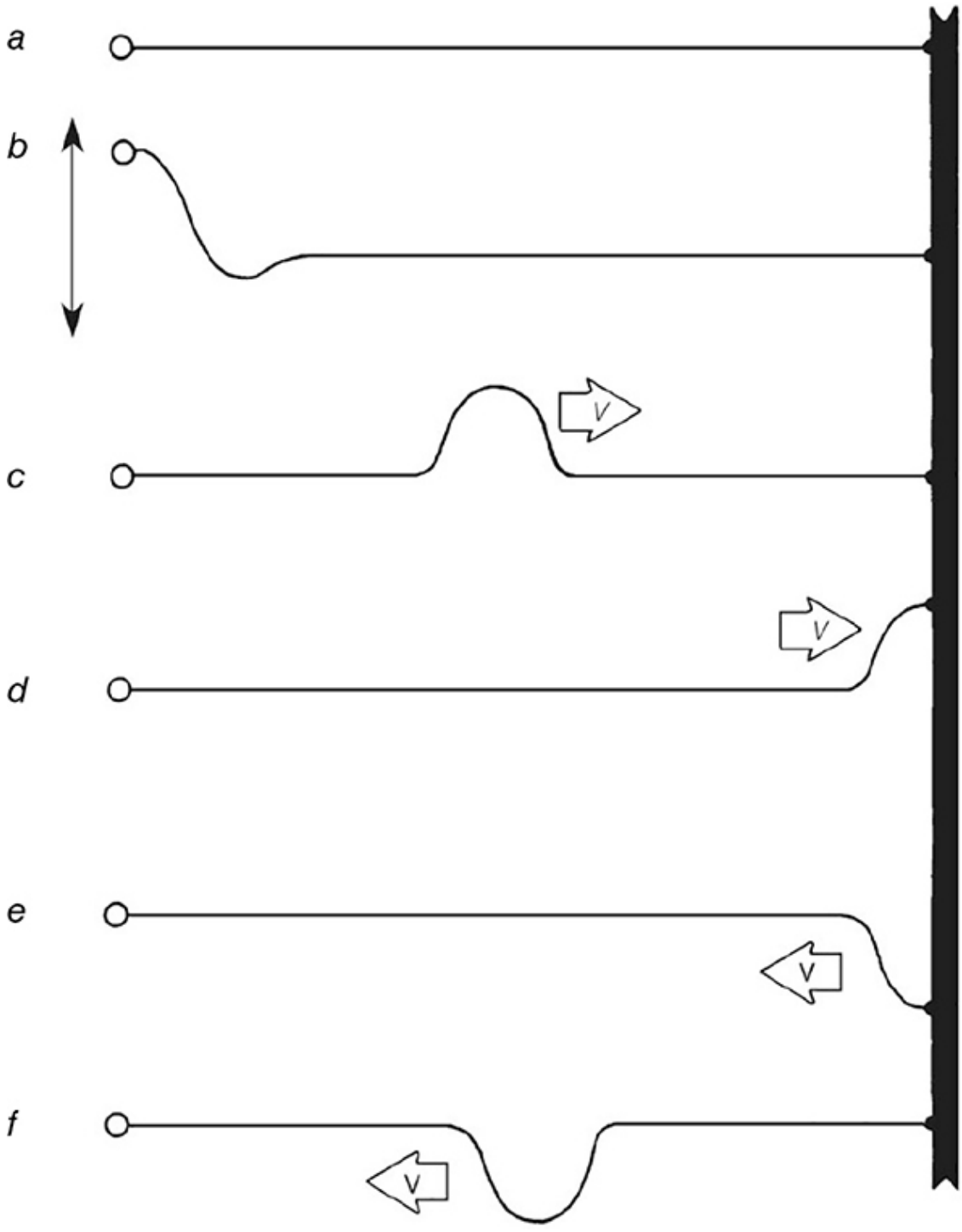


FIGURE 6.15 Rope analogy to transmission line.

If a second pulse is propagated down the line before the first pulse dies out, then there will be two pulses on the line at the same time (Fig. 6.16). When the two pulses are both present on the rope,

the resultant deflection of the rope will be the sum of the two pulses. If a continuous train of pulses is applied to the line at certain resonant repetition rates, the net deflection resulting from the combined incident and reflected pulse amplitudes at each point along the rope will create standing waves—one example of which is shown in Fig. 6.17—that do not propagate in either direction on the rope but, instead, go up and down without moving along the rope.

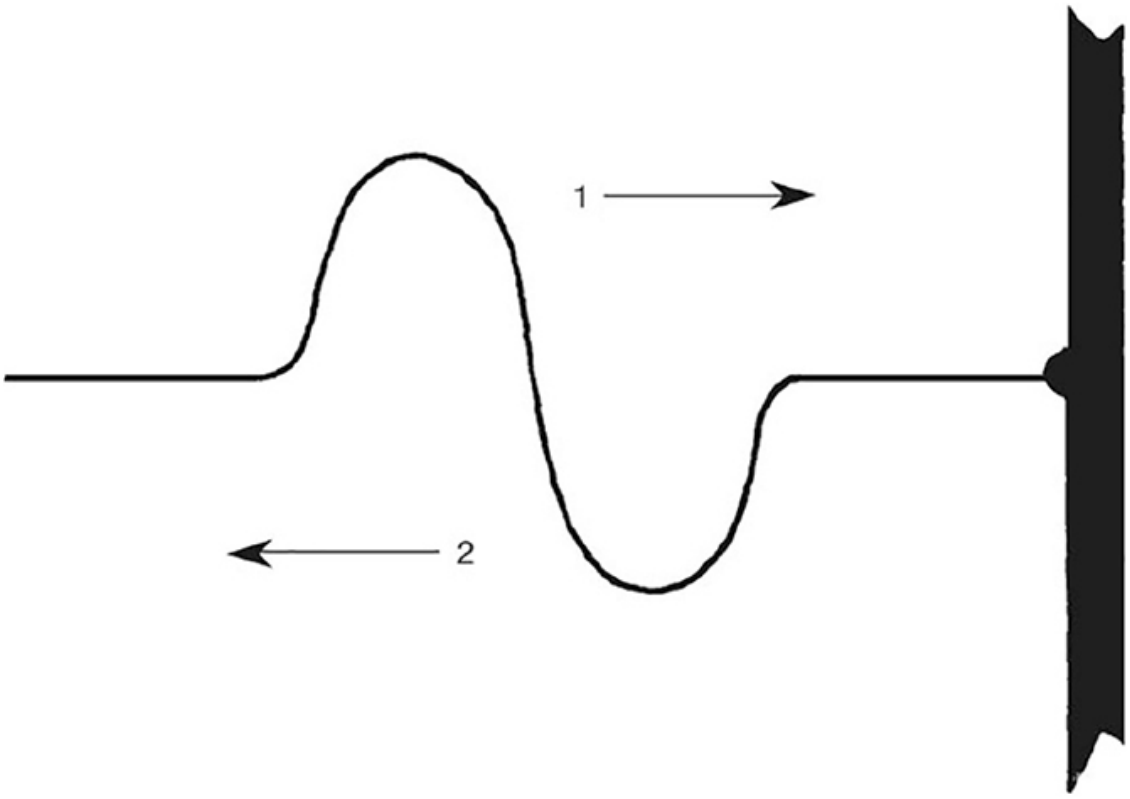


FIGURE 6.16 Interfering opposite waves.

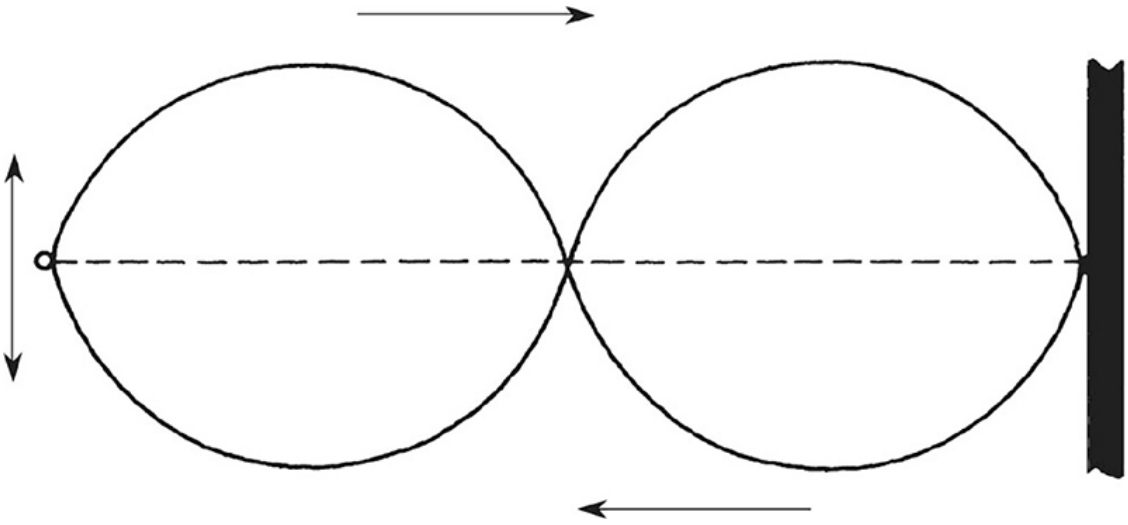


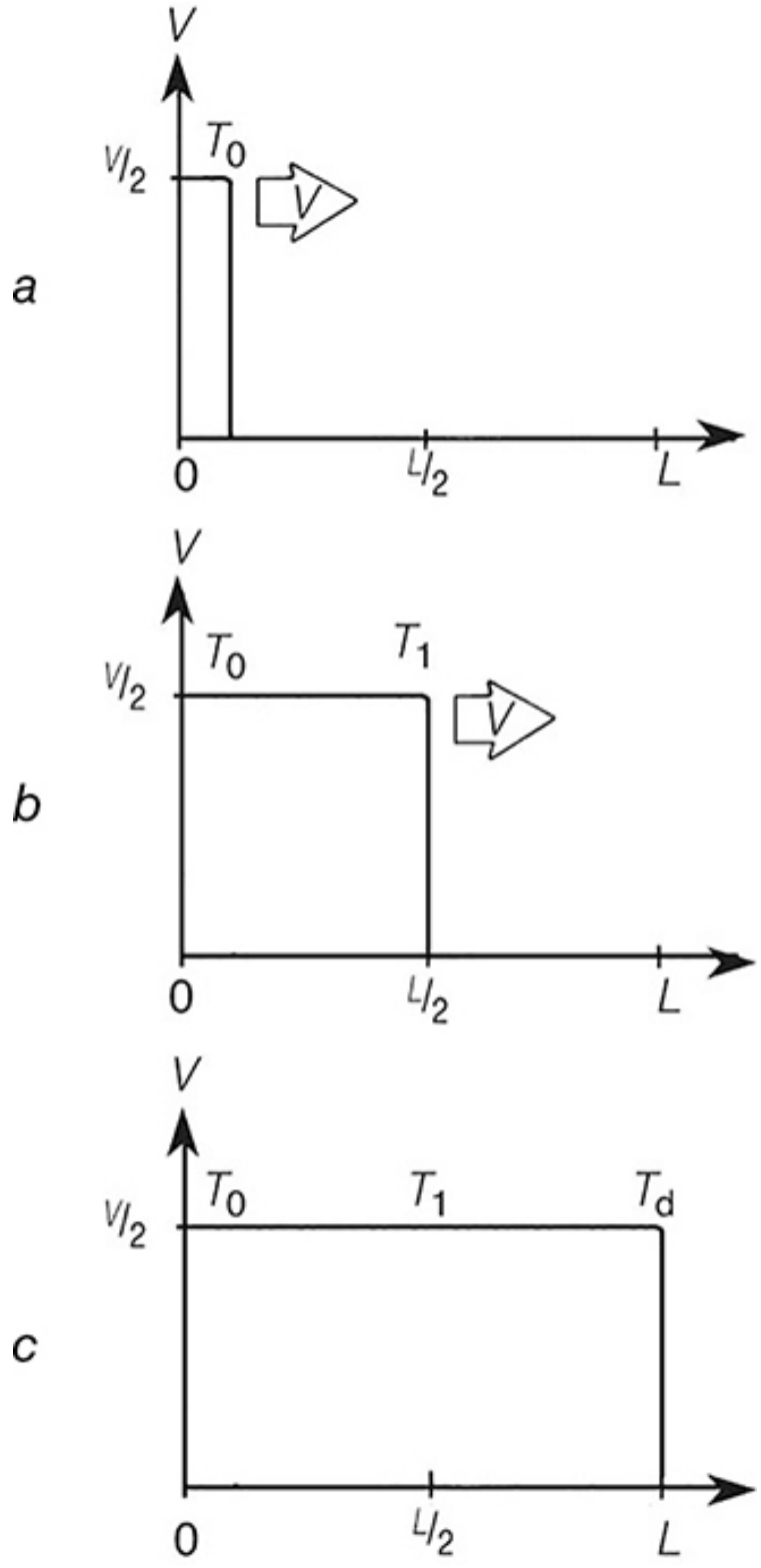
FIGURE 6.17 Standing waves.

Now let's look at a transmission line with characteristic impedance Z_0 connected to a load impedance Z_L, as shown in Fig. 6.14. The generator at the input of the line consists of a voltage source *V* with source impedance Z_S in series with a switch S_1. Assume for the present that both the source and the load impedances are pure resistances (i.e., *R*+*j*0) equal to Z_0, the characteristic impedance of the transmission line.

When the switch is closed at time T_0 (Fig. 6.18*a*), the voltage at the input of the line (V_{IN}) jumps to *V*/2. In Fig. 6.14, you may have noticed that the $L'C'$ circuit resembles a delay line. As might be expected, the voltage wavefront propagates along the line at a velocity given by

$$v = \frac{1}{\sqrt{L'C'}} \quad (6.14)$$

where
L'
is the line inductance in Henrys per meter and
C'
is the capacitance in Farads per meter.



F IGURE 6.18 Step-function propagation along transmission line at three points.

At time T_1 (Fig. 6.18*b*), the wavefront has propagated one-half the distance *L*, and by T_d it has propagated the entire length of the cable (Fig. 6.18*c*).

If the load is perfectly matched to the line (i.e., $Z_L = Z_0$), the load absorbs the wave and no component is reflected. But in a mismatched system (where $Z_L \neq Z_0$), a portion of the wave is reflected back up the line toward the source.

6.7 Reflection Coefficient

Consider a length of transmission line connecting a transmitter to an antenna. The transmitter launches a forward wave that travels down the line toward the antenna, which acts as a load on the end of the transmission line. If the impedance of the antenna or load is matched to the characteristic impedance of the line, the forward wave is fully absorbed by the load and becomes power radiated by the antenna. If there is a mismatch, part or all of the forward wave reflects and power is carried by a reverse wave back toward the generator.

Impedance mismatches can vary from a short circuit at the end of a transmission line to an open circuit (no load at all). In between are all the possible combinations of resistance and reactance. If the impedance at the end of a line is close to the line characteristic impedance, the reflection is small. The greater the difference in the line impedance and the load, the larger the reflection. With an open circuit, short circuit, or a purely reactive load, all the power in the forward wave is reflected from the load and returns to the input end of the transmission line.

At the end of a transmission line, the reflection coefficient is defined to be

$$\Gamma = \frac{V_{\mathrm{ref}}}{V_{\mathrm{fwd}}} \quad (6.15)$$

where
V_{fwd}
is the amplitude of the incident or forward wave on the transmission line and
V_{ref}
is the amplitude of the reflected wave. Since the amplitudes are complex, Γ is also a complex number that includes the magnitude and phase of the reflection from the load impedance.

For passive load impedances, the absolute value of the reflection coefficient varies from
-1
to
$+1$
, depending upon the nature of the reflection.
$\Gamma = 0$
indicates a perfect match with no reflection, while
-1
indicates a short-circuited load (
$Z_L = 0$
), and
$+1$
indicates an open circuit (
$Z_L = \infty$
).

Mathematically, reflections are related to the boundary conditions that hold at the junction of the transmission line and the load. For the case of a short circuit, the voltage at the load is always zero. When a forward wave arrives, a reflected wave is produced with equal and opposite amplitude, so that the total voltage at the short circuit load remains zero. This results in the reflection coefficient of

-1
mentioned above for the short circuit load.

Similarly, if we disconnect the load completely, so that the transmission line is looking into an open circuit. In this case,

$Z_L = \infty$
and
$I_L = 0$
. But if the current in the advancing wave is to be continuous, there must be an equal and opposite current flowing at the end of the transmission line to make the net load current zero. We call this new current
I_{ref}
, and it must be exactly equal in magnitude to
I_{fwd}
but of opposite polarity through the load. For the reverse voltage wave,
$V_{\text{ref}} = V_{\text{fwd}}$
with a positive sign, so
$\Gamma = 1$
also as mentioned above.

Because there are both forward and reflected waves on the line, the total voltage and current *V* and *I* are the sum of the forward and reflected voltages and currents. At the load end of the line, the ratio of the total current and voltage is equal to the load impedance:

$$Z_L = \frac{V}{I} = \frac{V_{\text{fwd}} + V_{\text{ref}}}{I_{\text{fwd}} + I_{\text{ref}}} \quad (6.16)$$

The forward and reverse wave voltage and current are related by the characteristic impedance according to

$$I_{\text{fwd}} = \frac{V_{\text{fwd}}}{Z_0} \quad (6.17\text{a})$$

$$I_{\text{ref}} = -\frac{V_{\text{ref}}}{Z_0} \quad (6.17\text{b})$$

The negative sign means that the direction of power flow of the reverse wave is opposite that of the forward wave. Using these expressions in Eq. (6.16) and solving for the ratio
$\Gamma = V_{\text{ref}}/V_{\text{fwd}}$
leads to the reflection coefficient formula

$$\Gamma = \frac{Z_L - Z_0}{Z_L + Z_0} \quad (6.18)$$

Example 6.3 A 50-Ω transmission line is connected to a 30 Ω resistive load. Calculate the reflection coefficient Γ.

Solution

$$\begin{aligned} \Gamma &= \frac{Z_L - Z_0}{Z_L + Z_0} \\ &= \frac{50\ \Omega - 30\ \Omega}{50\ \Omega + 30\ \Omega} \\ &= 0.25 \end{aligned}$$

Example 6.4 In Example 6.3, the incident (forward) voltage is 3 V rms. Calculate the reflected voltage.

Solution Since
$\Gamma = V_{\text{ref}}/V_{\text{fwd}}$
, the reflected voltage is

$$\begin{aligned} V_{\text{ref}} &= \Gamma V_{\text{fwd}} \\ &= (0.25)(3\ \text{V}) \\ &= 0.75\ \text{V} \end{aligned}$$

Using the relationship between load impedance and transmission line characteristic impedance, we can determine the phase of the reflected voltage. For resistive loads ($Z_L = R_L + j0$):

- If Z_L/Z_0 is less than 1.0, the reflected signal is 180 degrees out of phase with the incident voltage.
- If the ratio Z_L/Z_0 is greater than 1.0, the reflected signal is in phase with the incident voltage.

If $Z_L = Z_0$, there is no reflection.

The step-function (or pulse) response of the transmission line provides a way to analyze the line and the load using an oscilloscope. Figure 6.19 shows (in schematic form) the test setup for time domain reflectometry (TDR) measurements. As discussed in Chap. 30, these measurements can be done more easily using a vector network analyzer (VNA), but the TDR setup is quite satisfyingly visual.

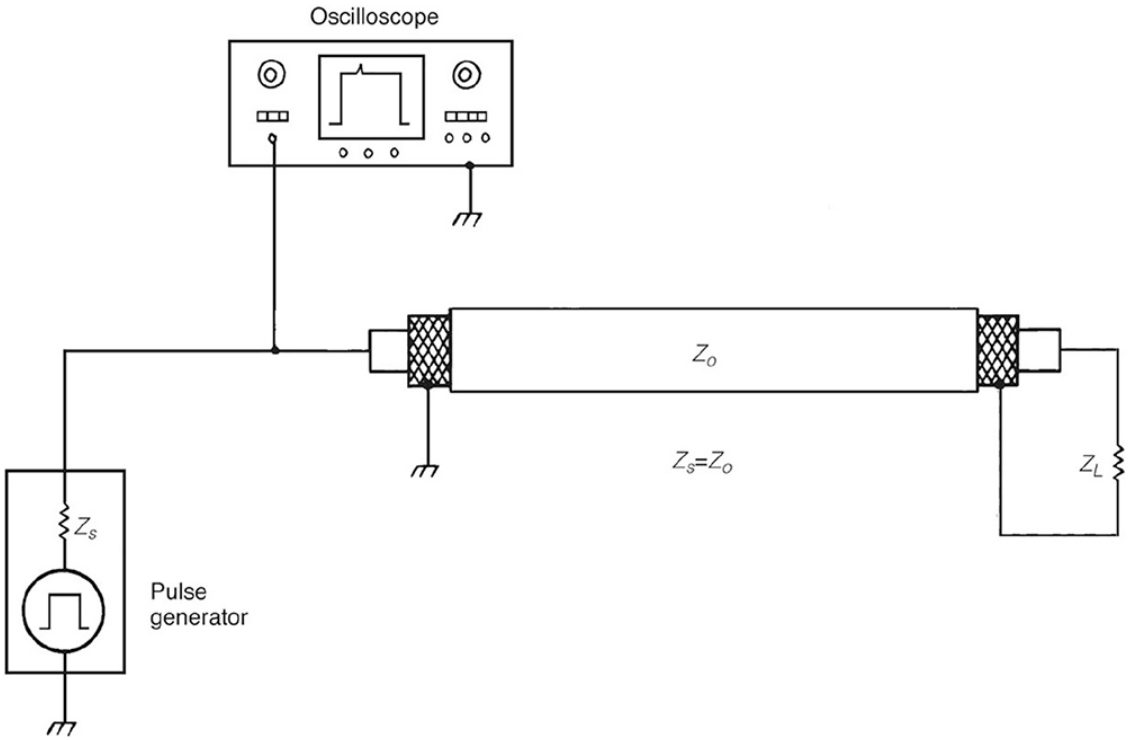


FIGURE 6.19 Time domain reflectometry setup.

An oscilloscope and a pulse or square-wave signal generator are connected in parallel across the input end of the transmission line. Figure 6.20 shows a test setup for HF lines. The small shielded box contains a square-wave oscillator circuit. Although a crystal oscillator can be used, an RC-timed circuit running at 1 MHz or less is sufficient. The pulse generator is connected in parallel with the cable under test and the input of the oscilloscope. A BNC tee connector and a double male BNC adapter are used to connect the line and pulse generator with the scope.

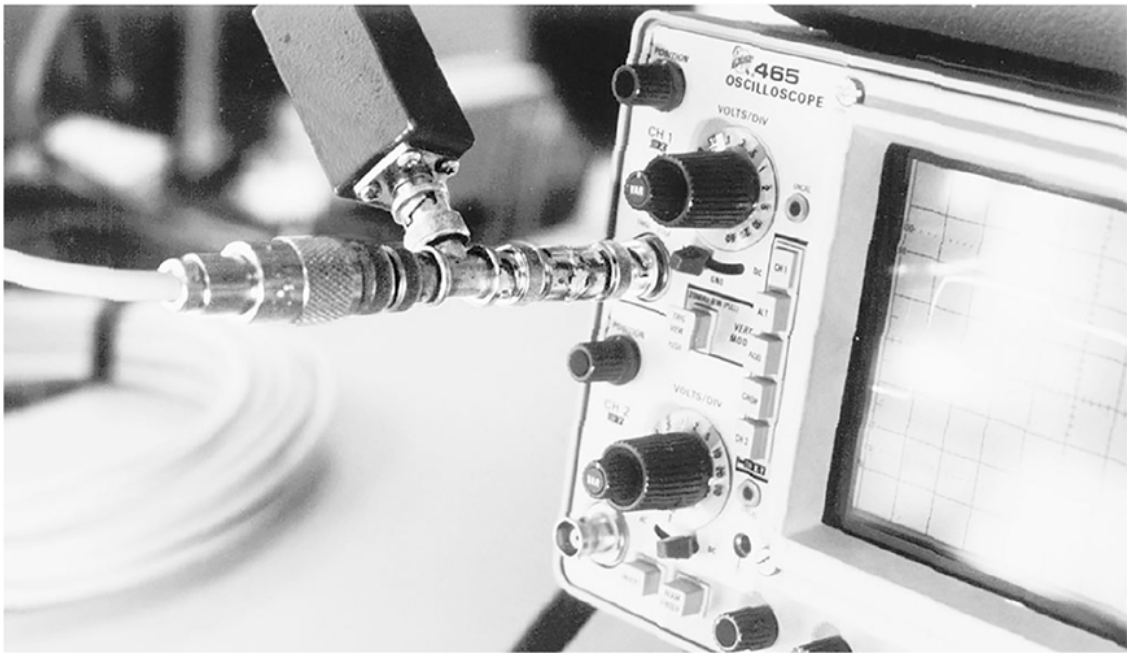


FIGURE 6.20 Test setup for time domain reflectometry.

If a periodic waveform is supplied by the generator, the oscilloscope will show the sum of the reflected and incident pulses. The duration of the pulse (i.e., pulse width) is adjusted so that the returning reflected pulse arrives approximately in the center of the incident pulse.

The images of Figs. 6.21 and 6.22 are TDR displays for several different values of

Z_L

. Approximately 30 m of coaxial cable with a velocity factor of 0.66 was used in a test setup similar to that shown in Fig. 6.19. The pulse width was approximately 0.9 μs (microseconds). The horizontal sweep time on the scope was adjusted to show only one pulse, which in this case represented half of a 550-kHz square wave (Fig. 6.22).

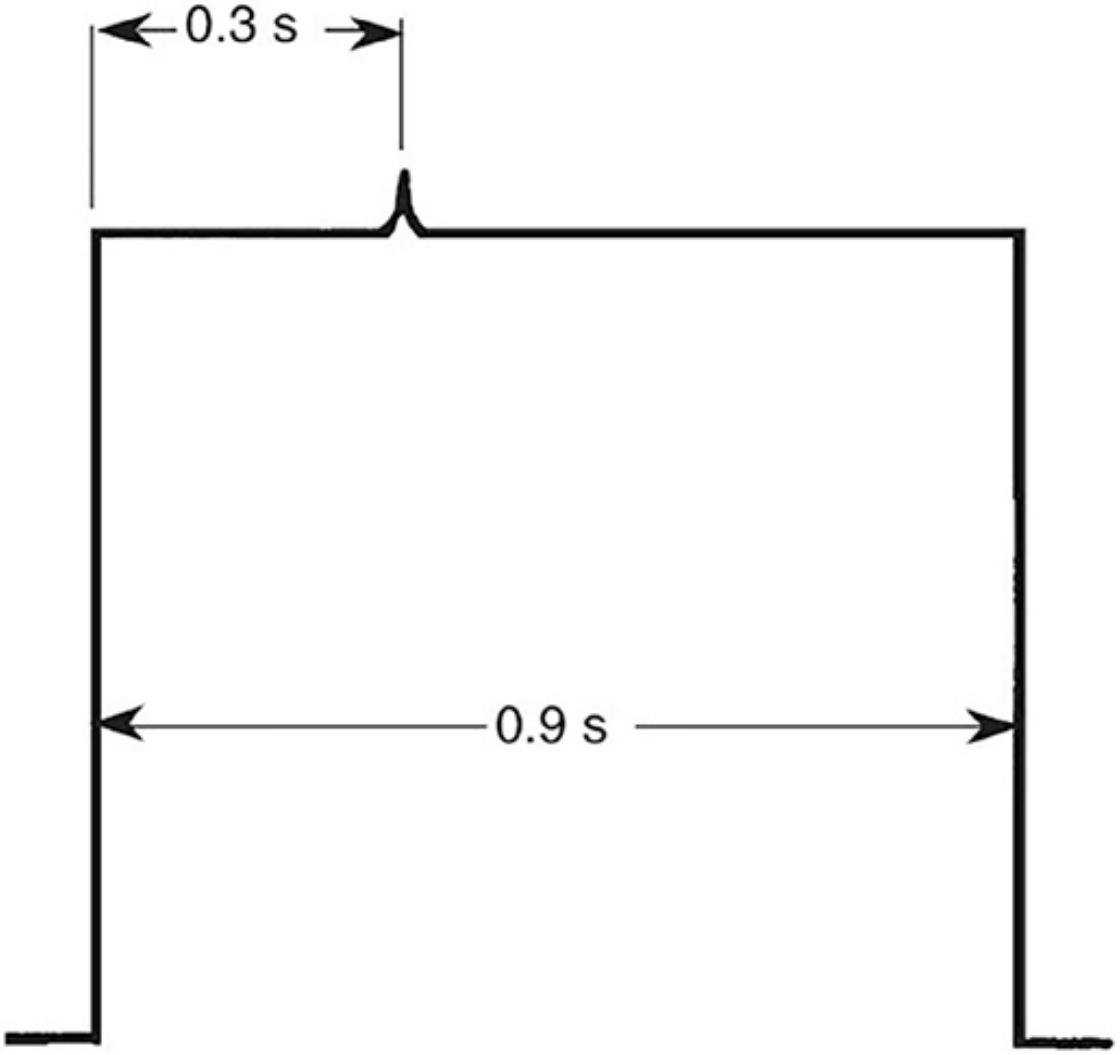


FIGURE 6.21 Idealized TDR pulse. The small peak is the reflected signal interfering with the forward pulse.

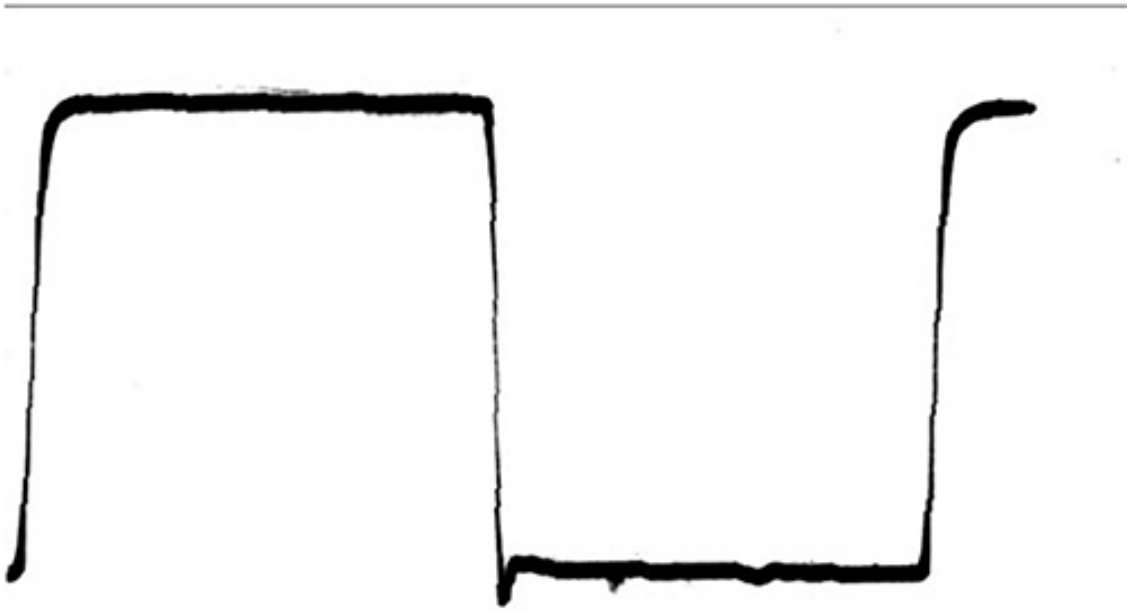


FIGURE 6.22 TDR pulse with $Z_L = Z_0$.

The trace in Fig. 6.22 corresponds to a load matched to the line ($Z_L = Z_0$). A slight discontinuity, representing a small reflected wave, can be seen on the high side of the pulse. Even though the load and line are nominally matched, any connector or other dimensional change may present a slight impedance discontinuity. Any discontinuity, damage along the line, sharp bend, or other anomaly can cause a slight impedance variation that is observable as a small reflection. This principle is used to find breaks in buried cables with TDR equipment.

In the example shown, the anomaly occurs at about one-third of the 0.9-μs duration (or 0.3 μs) after the onset of the pulse. This tells us that the reflected wave arrives back at the source 0.3 μs after the incident wave leaves. Because this time period represents a round-trip, you can conclude that the wave required 0.3 μs/2, or 0.15 μs, to propagate the length of the line. Since we know the velocity factor, we can find the approximate length of the cable:

$$\begin{aligned} l &= cv_F T \\ &= \left(3 \times 10^8 \mathrm{m/s}\right) \times 0.66 \times \left(1.5 \times 10^{-7} \mathrm{s}\right) \\ &= 29.7 \mathrm{\ m} \end{aligned}$$

which agrees within experimental accuracy to the 30 m actual length.

Figures 6.23 through 6.28 show the signal waveform at the input end of the transmission line resulting from a step-function input for various values of mismatched load impedance (i.e., Z_L not equal to Z_0). When the load impedance is less than the line impedance (in this case, $0.5Z_0$), the reflected wave is inverted as seen in Fig. 6.23 and subtracts from the incident wave along the top of the pulse. The reflection coefficient can be determined by examining the relative amplitudes of the two waves.

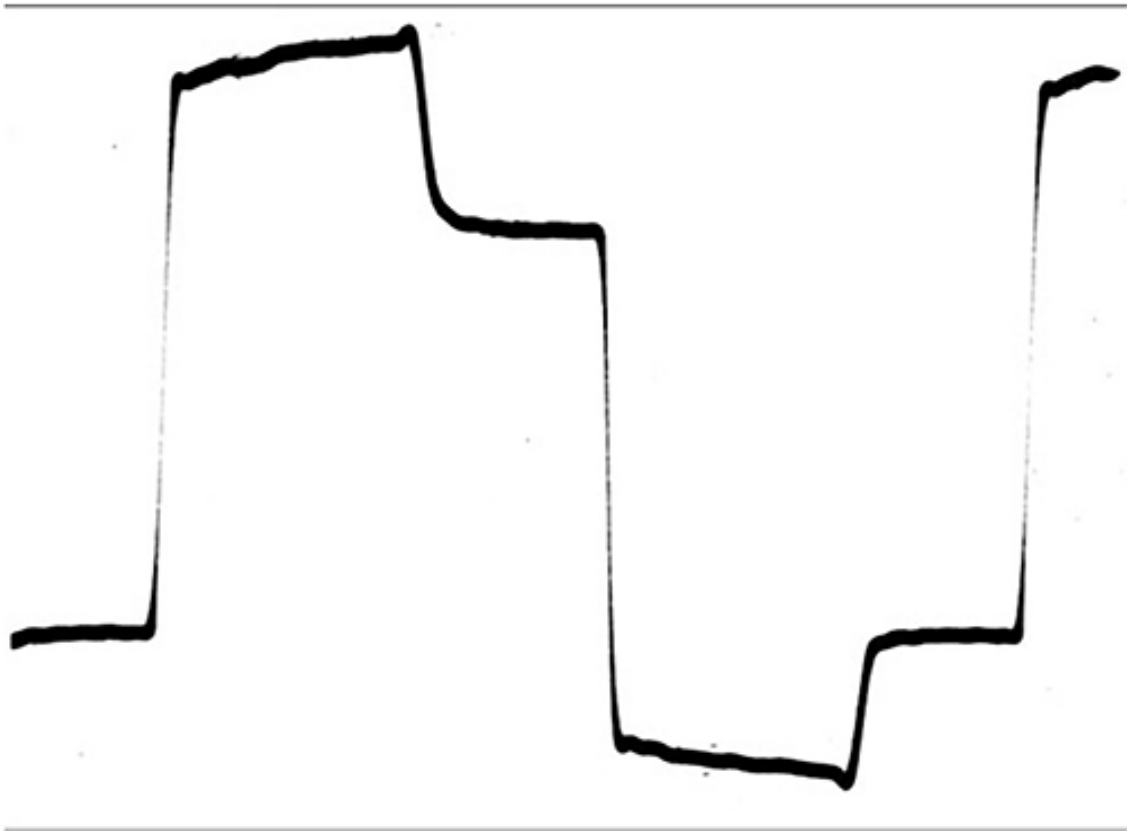


FIGURE 6.23
$Z_L < Z_0$
.

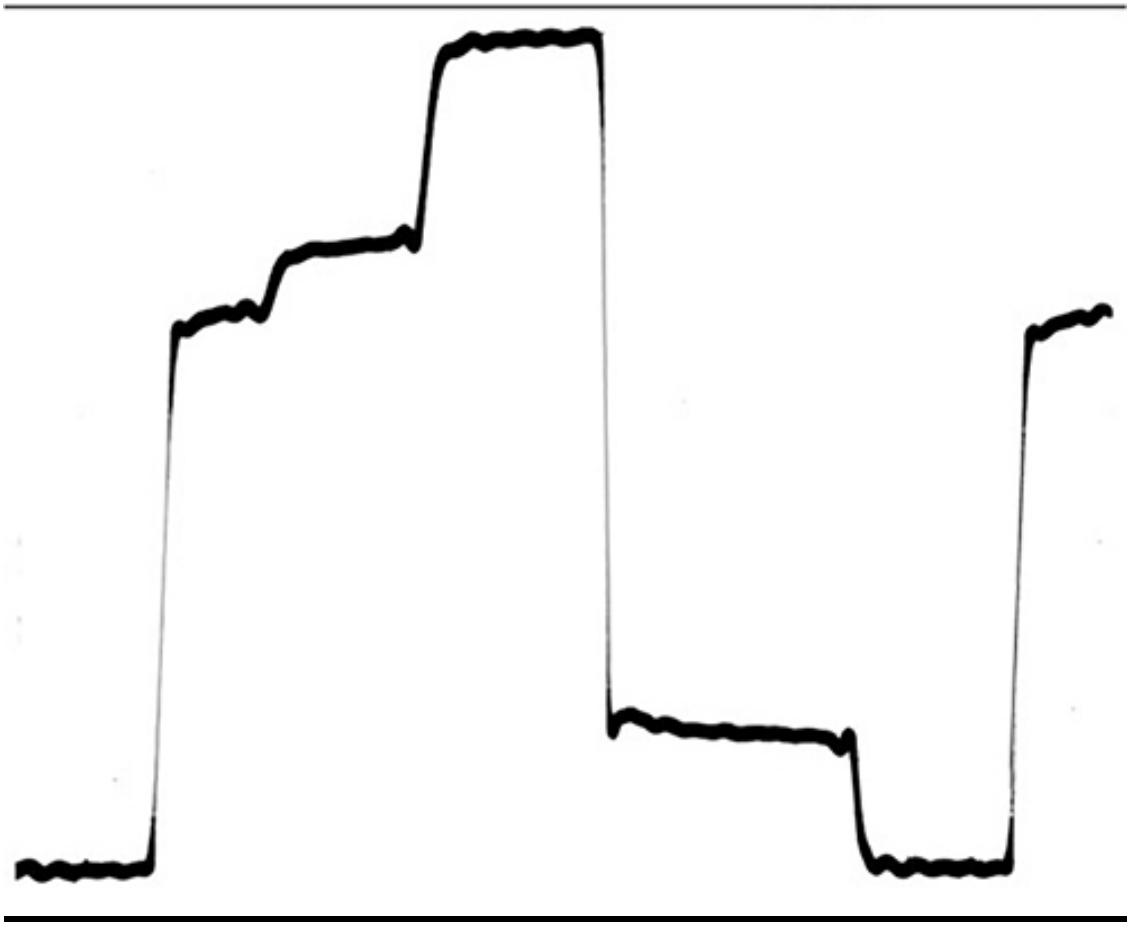


FIGURE 6.24
$Z_L > Z_0$
.

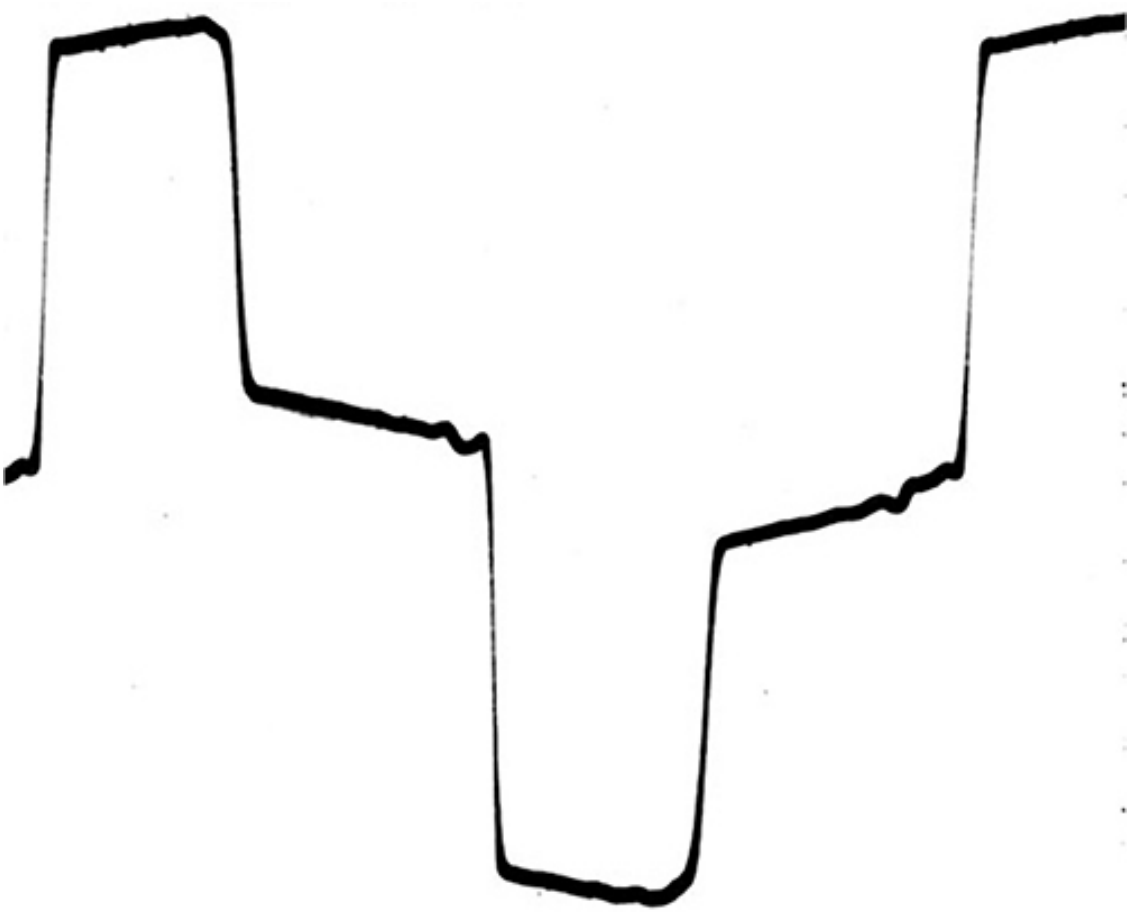


FIGURE 6.25
$Z_L = 0$
.

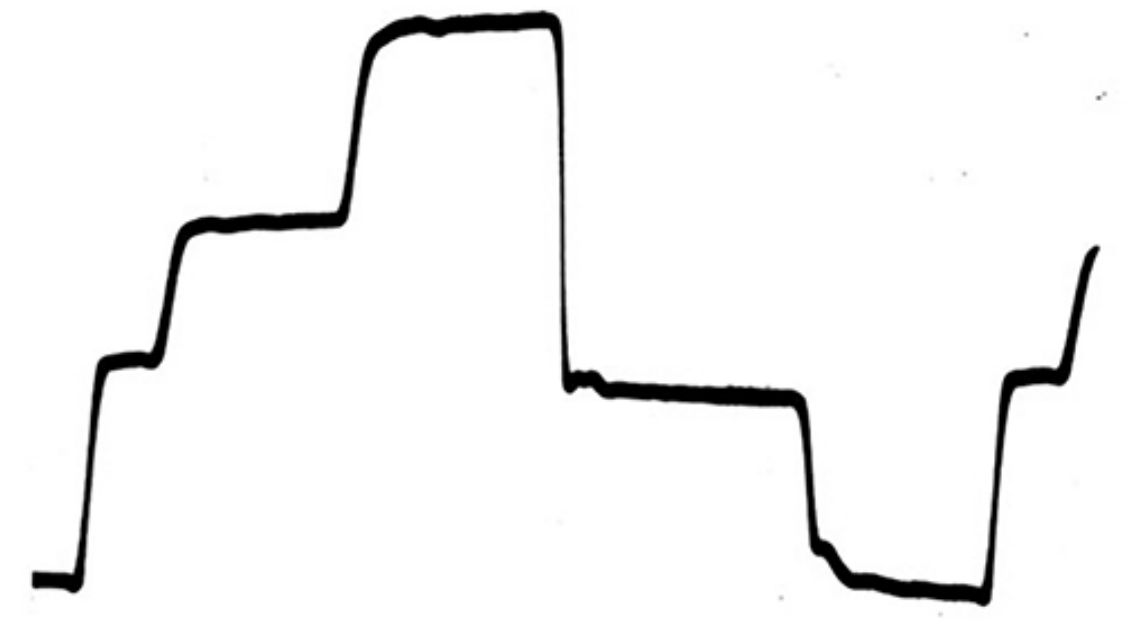


FIGURE 6.26
$Z_L = \infty$
.



FIGURE 6.27
$Z_L = 50 - jX_C$
.

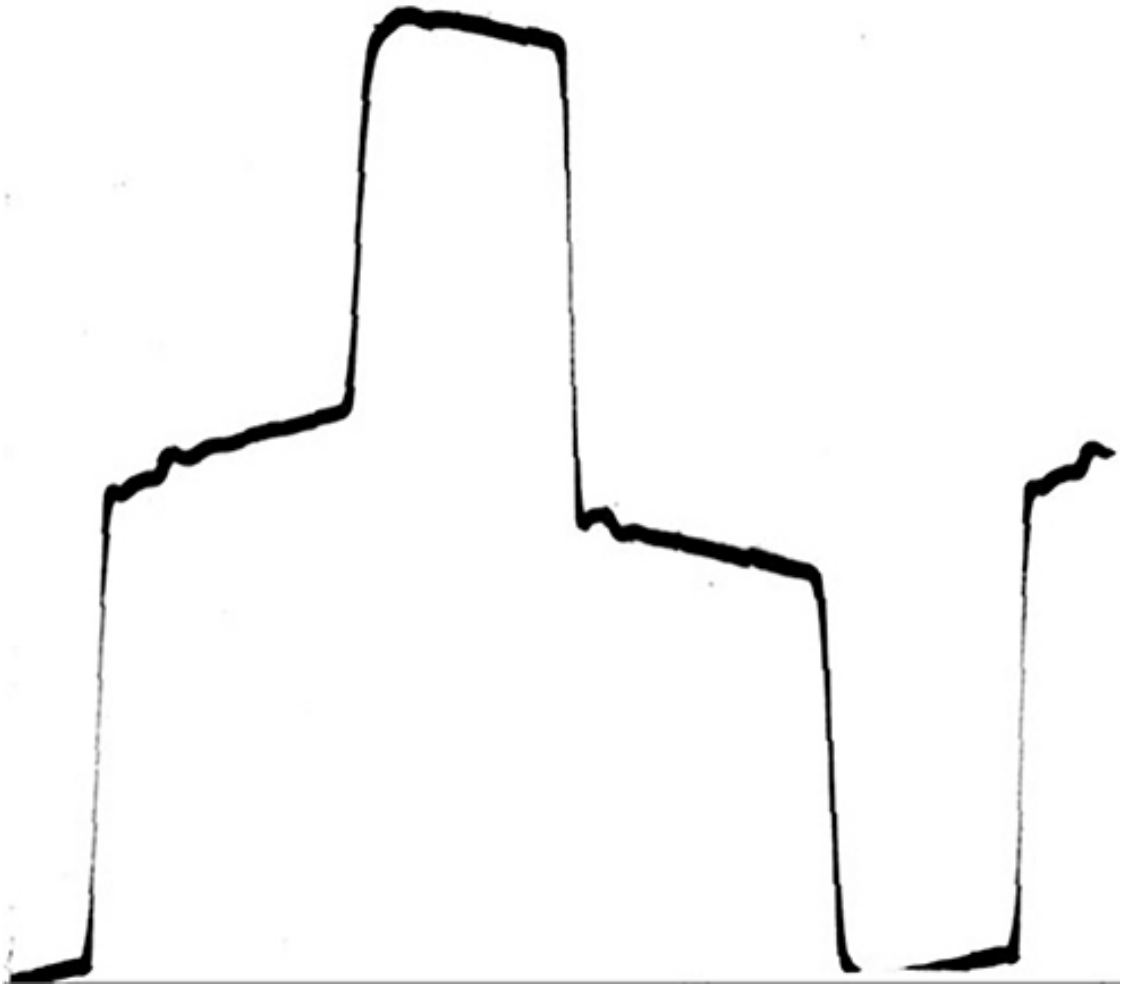


FIGURE 6.28
$Z_L = 50 + jX_L$
.

When
$Z_L = 2Z_0$
, the waveform of Fig. 6.24 results. In this case, the reflected wave is in phase with the incident wave, so it adds to the incident wave as shown. Waveforms for a short-circuited load and an open-circuited load are shown in Figs. 6.25 and 6.26, respectively. Finally, examples of waveforms resulting from adding capacitive and inductive reactance to a 50 Ω load are displayed in Figs. 6.27 and 6.28. The waveform in Fig. 6.27 resulted from a capacitance in

series with a 50 Ω (matched) resistance; the waveform in Fig. 6.28 resulted from a 50 Ω resistance in series with an inductance.

6.8 Sinusoidal Steady State

When a continuous wave (CW) RF signal is applied to a transmission line, the excitation is sinusoidal (Fig. 6.29). When the generator is turned on, transients reflect from the load and propagate back and forth on the transmission line. After a few reflections, the transients die out, and the line stabilizes at a steady-state condition.

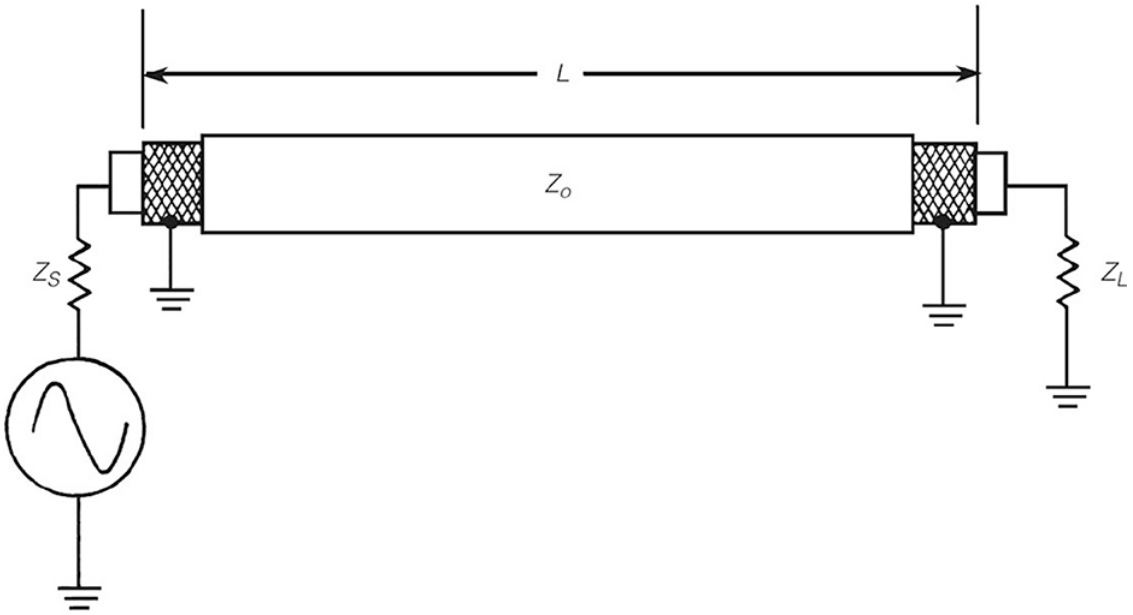


FIGURE 6.29 AC-excited transmission line.

As with the step function considered above, with a sinusoidal excitation on a transmission line that is not matched to its load, some of the energy is absorbed by the load and some is reflected back up the line toward the source. The interference of the forward and reflected waves creates standing waves on the transmission line. If the peak voltage or current is measured along the line, it will vary in a way that depends on the load as shown in Fig. 6.30.

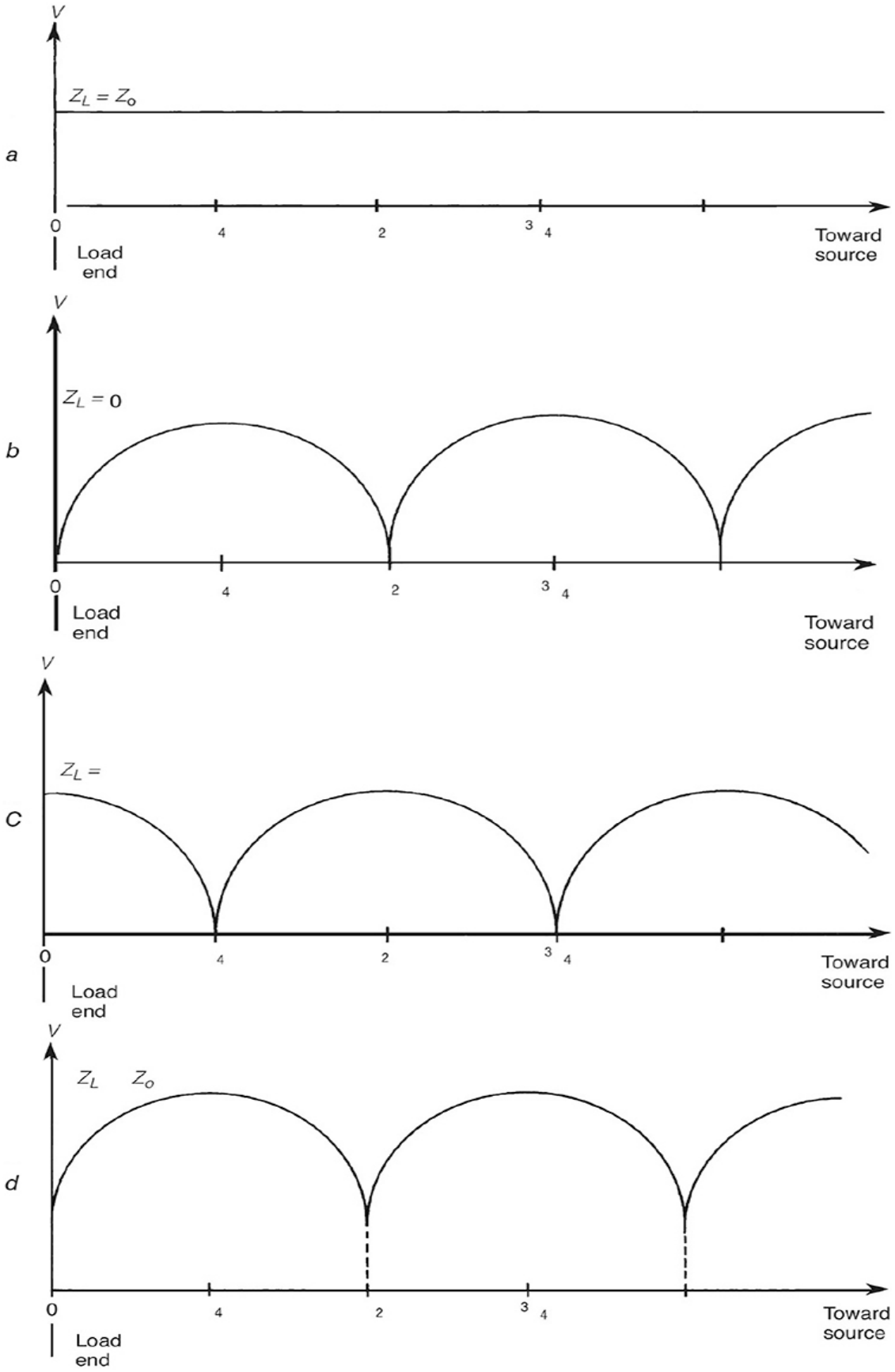

V
$Z_L = Z_o$
a
0
4
2
3 4
Load end
Toward source
V
$Z_L = 0$
b
0
4
2
3 4
Load end
Toward source
V
$Z_L =$
C
0
4
2
3 4
Load end
Toward source
V
Z_L Z_o
d
0
4
2
3 4
Load end
Toward source

FIGURE 6.30 Voltage versus electrical length. (*a*) Matched impedances. (*b*) $Z_L = 0$, (*c*) $Z_L = \text{infinite}$. (*d*) Z_L not equal to Z_0.

Figure 6.30*a* is a voltage-versus-length curve for a matched line—that is, where $Z_L = Z_0$. Since there is no reverse wave, there is no standing wave on the line and we only measure the peak voltage of the forward wave, which is constant along the line.

Figure 6.30*b* shows the voltage distribution over the length of the line when the load end of the line is shorted (i.e., $Z_L = 0$). At the load end the voltage is zero, which results from the boundary condition of a short circuit. The forward and reverse waves move in and out of phase along the line, resulting in a standing wave pattern that moves from zero to a peak value every half-wavelength along the line.

When the line is unterminated or open-circuited (i.e., $Z_L = \infty$), the pattern in Fig. 6.30*c* results. The standing wave pattern has the same shape as Fig. 6.30*b* for the shorted line, but the phase is shifted 90 degrees, or a quarter-wavelength, along the line.

If Z_L is not equal to Z_0

and is neither zero nor infinite, the standing wave pattern is similar to that depicted in Fig. 6.30*d*. In this case, the nodes exhibit some finite voltage, V_{MIN}, instead of zero.

6.8.1 Standing Wave Ratio

The degree of load mismatch can be measured using the standing wave ratio (SWR). The SWR can be measured using either the current or voltage along the line. If the SWR is measured with the voltage, it is sometimes referred to as the voltage standing wave ratio (VSWR).

SWR on a transmission line is important for several reasons:

- The higher the SWR, the higher the maximum voltages on a transmission line for a given power applied or delivered to the load. High SWR can cause unexpected destruction of the line and any electronics equipment attached to it.
- Solid-state transmitters, transceivers, and amplifiers are often designed to reduce RF output or shut down completely in the presence of high SWR.
- The reflected wave on a transmission line represents energy rejected by the load. For a specified power delivered to the antenna, a large SWR means higher required power ratings for components in the system.
- A transmission line operating with high SWR will exhibit greater loss per unit length than its specified loss for a matched load impedance.

VSWR can be specified several ways:

- From incident (or forward) voltage and reflected voltage:

$$\mathrm{VSWR} = \frac{|V_{\mathrm{fwd}}| + |V_{\mathrm{ref}}|}{|V_{\mathrm{fwd}}| + |V_{\mathrm{ref}}|} \quad (6.19)$$

- From transmission line voltage measurements (Fig. 6.30*d*):

$$\mathrm{VSWR} = \frac{V_{\max}}{V_{\min}} \quad (6.20)$$

- From load and line characteristic impedances:

$$(Z_L > Z_0)\ \mathrm{VSWR} = Z_L / Z_0 \quad (6.21)$$

$$(Z_L < Z_0)\ \mathrm{VSWR} = Z_0 / Z_L \quad (6.22)$$

- From the forward and reflected power:

$$\mathrm{VSWR} = \frac{1 + \sqrt{P_{\mathrm{ref}} / P_{\mathrm{fwd}}}}{1 - \sqrt{P_{\mathrm{ref}} / P_{\mathrm{fwd}}}} \quad (6.23)$$

- From the reflection coefficient:

$$\mathrm{VSWR} = \frac{1 + |\Gamma|}{1 - |\Gamma|} \quad (6.24)$$

It's also possible to determine the magnitude of the reflection coefficient from the VSWR:

$$|\Gamma| = \frac{\mathrm{VSWR} - 1}{\mathrm{VSWR} + 1} \quad (6.25)$$

VSWR is sometimes expressed as a ratio. For example, when Z_L is 100 Ω and Z_0 is 50 Ω, the VSWR is $Z_L / Z_0 = 100\ \Omega / 50\ \Omega = 2$, which can be written as $\mathrm{VSWR} = 2.0 : 1$

. VSWR can also be expressed in decibel form:

$$\text{VSWR (dB)} = 20 \log_{10}(\text{VSWR}) \quad (6.26)$$

Example 6.5 A transmission line is connected to a mismatched load. Calculate the VSWR and the VSWR decibel equivalent if the reflection coefficient Γ is 0.25.

Solution

$$\begin{aligned} \text{(a) VSWR} &= \frac{1 + |\Gamma|}{1 - |\Gamma|} \\ &= \frac{1 + 0.25}{1 - 0.25} \\ &= 1.67 : 1 \end{aligned}$$

$$\begin{aligned} \text{(b)VSWR(dB)} &= 20 \log_{10} (\text{VSWR}) \\ &= 20 \log_{10} 1.67 \\ &= 4.4 \text{ dB} \end{aligned}$$

6.8.2 Mismatch Losses

The power reflected from a mismatched load reduces performance in a radio link and can in some cases cause damage. Reflected power can damage the output stage of a transmitter. This is a common enough issue that many transmitters include circuitry to sense high VSWR and turn down output power proportionately. When making measurements, a poor VSWR on the transmission lines that interconnect devices under test, instruments, and signal sources can cause erroneous readings and invalid measurements.

Determination of VSWR losses must take into account mismatches at the generator and load ends of a transmission line. Figure 6.29 shows a transmission line of impedance Z_0 interconnecting a load impedance Z_L

to a source with output impedance Z_S. There is a potential for impedance mismatch at both ends of the line.

When one end of the line is matched, the mismatch loss (ML) is

$$\mathrm{ML} = -10 \log_{10} \left[1 - \left(\frac{\mathrm{SWR} - 1}{\mathrm{SWR} + 1} \right)^2 \right] \quad (6.27)$$

which from (6.25) is

$$\mathrm{ML} = -10 \log_{10} \left(1 - |\Gamma|^2 \right) \quad (6.28)$$

Example 6.6 A coaxial transmission line with a characteristic impedance of 50 Ω is connected to the 50 Ω output (Z_0) of a signal generator and a 20 Ω load impedance Z_L. Calculate the mismatch loss.

Solution

(a) We first find the VSWR:

$$\begin{aligned} \mathrm{VSWR} &= \mathrm{Z}_0 / \mathrm{Z}_L \\ &= (50\ \Omega)/(20\ \Omega) \\ &= 2.5\ :\ 1 \end{aligned}$$

(b) The mismatch loss is:

$$\begin{aligned} \mathrm{ML} &= -10\ \log\left[1-\left(\frac{\mathrm{SWR}-1}{\mathrm{SWR}+1}\right)^2\right] \\ &= -10\ \log\left[1-\left(\frac{1.5}{3.5}\right)^2\right] \\ &= 0.89 \end{aligned}$$

When both ends of the line are mismatched, a different equation is required:

$$\mathrm{ML} = 20\log_{10}[1 \pm |\Gamma_1\Gamma_2|] \quad (6.29)$$

where
Γ_1
is the reflection coefficient at source end of line and
Γ_2
is the reflection coefficient at the load end of the line. The solution to Eq. (6.29) has two values:
$1+|\Gamma_1\Gamma_2|$
and
$1-|\Gamma_1\Gamma_2|$
.

The equations reflect the mismatch loss for low-loss transmission lines. Loss on the transmission line can exacerbate the effect of standing waves on the line. Interference between forward and reflected waves produces increased current at antinodes—which increases ohmic losses—and increased voltage at antinodes—which increases dielectric losses. As dielectric loss increases with frequency, overall loss can become worse at higher frequencies. Line losses can be used with the reflection coefficient to determine total loss on a line with a given VSWR,

$$\mathrm{Loss} = 10\log_{10}\left(\frac{n^2-|\Gamma|^2}{n-n|\Gamma|^2}\right) \quad (6.30)$$

where
$n = 10^{A/10}$
and *A* is the total attenuation presented by line in dB with a matched line (
$Z_L = Z_0$
).

Example 6.7 A 50-Ω transmission line is terminated in a 30-Ω resistive impedance. The line is rated at a loss of 3 dB/100 ft at 1 GHz. Calculate (a) loss in 5 ft of line, (b) reflection coefficient, and (c) total loss in a 5-ft line mismatched per above.

Solution

$$\begin{aligned}(a) A &= \frac{3\text{dB}}{100\ f\text{t}} \times 5\ f\text{t} \\ &= 0.15\text{dB}\end{aligned}$$

$$\begin{aligned}\text{(b) } \Gamma &= \frac{Z_L - Z_0}{Z_L + Z_0} \\ &= \frac{50 - 30}{50 + 30} \\ &= 0.25\end{aligned}$$

$$\begin{aligned}\text{(c) } n &= 10^{(A/10)} \\ &= 10^{(0.15/10)} \\ &= 1.04\end{aligned}$$

$$\begin{aligned}\text{Loss} &= 10\ \log\left(\frac{n^2 - \Gamma^2}{n - n\Gamma^2}\right) \\ &= 10\ \log_{10}\left[\frac{(1.04)^2 - (0.25)^2}{1.04 - (1.04)(0.25)^2}\right] \\ &= 0.2\ \text{dB}\end{aligned}$$

Compare the matched line loss (
$A = 0.15\ \text{dB}$
) with the total loss (

$\text{Loss} = 0.2 \text{ dB}$
), which includes mismatch loss and line loss. The difference is only 0.05 dB. If the VSWR or the total line length were considerably larger, the loss would rise.

6.9 Impedance Matching in Antenna Systems

Although antennas can normally be used for both transmit and receive and are reciprocal devices, impedance matching considerations can be different for transmit and receive antennas. For transmitters, proper matching of the transmitter to the feedline and the feedline to the antenna is essential for maximizing our radiated signal. In the MF and lower HF bands, reception is usually limited by atmospheric noise, so impedance matching for receive antennas has little effect on the signal to noise ratio.

In simple designs, the transmitter, feedline, and antenna may have the same impedance. Most transmitters, transceivers, and amplifiers are designed for maximum power transfer to a 50-Ω resistive load. But antenna impedances vary considerably for many reasons and can be different in practice from designed values. The $\lambda/2$ dipole seldom appears to be exactly 75 Ω or purely resistive. Wire arrays, Yagis, and loops can have feedpoint impedances substantially higher or lower or more reactive than the ideal $\lambda/2$ dipole. In these cases, a matching network at the junction of the feedline and the antenna is likely to be needed.

We often want to match an antenna with some impedance to a given transmission line characteristic impedance, such as 50 Ω. In some cases, we want to match for maximum power transfer to a load. For maximum power transfer to a load with complex impedance, we match the system so that the impedance driving the load is complex conjugate of its impedance. This is a conjugate match.

Two common matching techniques used for antenna systems are:

- Distributed matching systems (stubs, transmission line sections, etc.)
- Lumped element tuners (antenna tuning units, or ATUs)

The second category is covered in Chap. 28.

6.9.1 Distributed Matching Networks

Matching circuits that depend on the propagation of the signal along a carefully chosen length of conductor for their matching capabilities are called distributed networks. The examples described in the following sections are typically employed either as a series element placed between the antenna and feedline or as a shunt element directly across the antenna feedpoint or a short distance along the feedline from it.

Since the dimensions of most distributed matching networks are a function of the operating frequency or wavelength, they only work over a limited bandwidth centered on the design frequency. Most of the networks described here are useful for single-band antennas and arrays.

6.9.2 Coaxial Cable Baluns

Figure 6.31 shows a balun that transforms impedances at a 4:1 ratio, with $R_2 = 4R_1$. The length of the balun (or U-shaped) section of coaxial cable is

$$L = \frac{492 v_F}{f} \quad (6.31)$$

where *L* is the physical length of the U-shaped section of cable in feet, v_F is the velocity factor of the cable (typically 0.66–0.85), and *f* is the operating frequency in MHz.

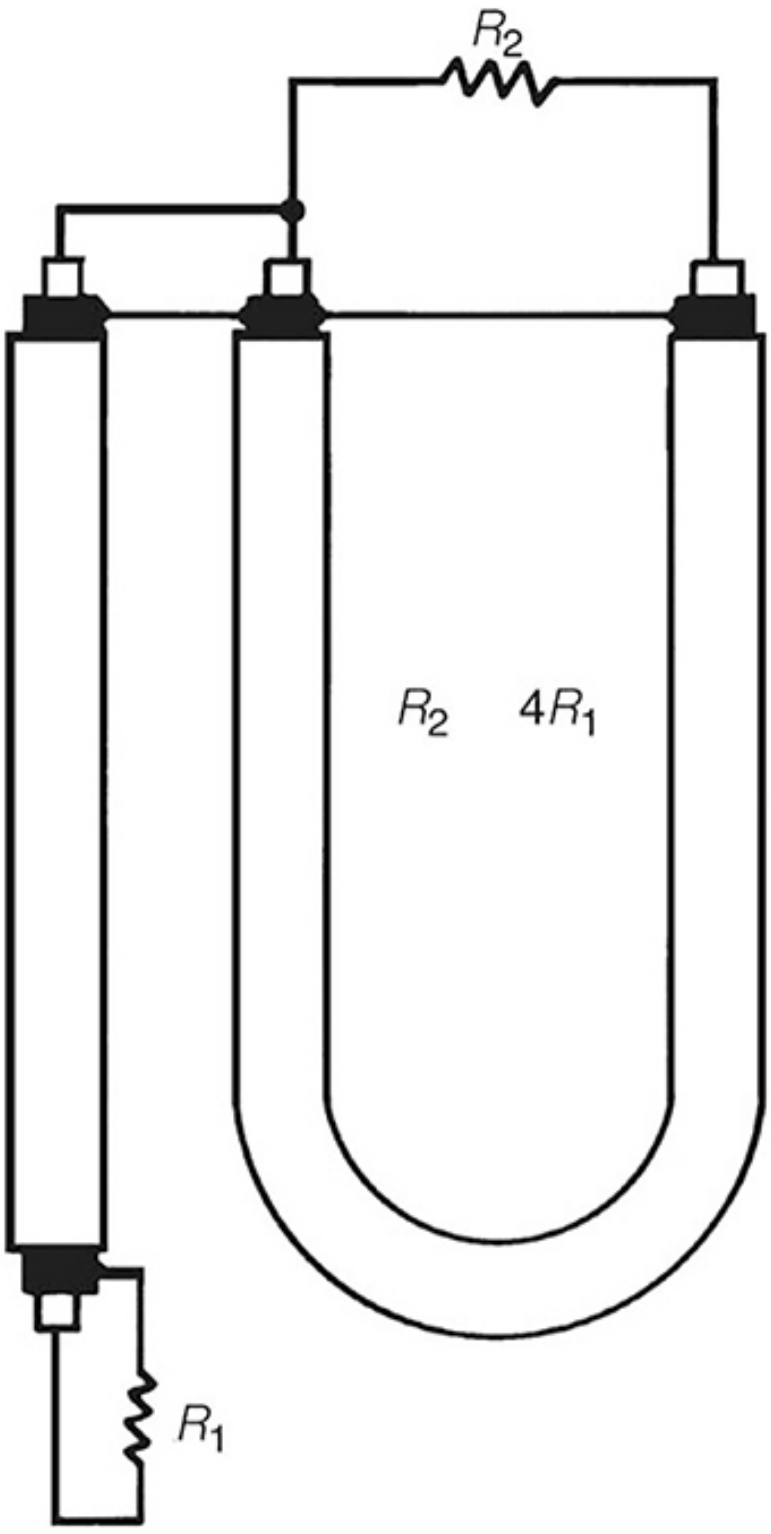


FIGURE 6.31 Coaxial balun transformer.

A balun can be used in a step-up or a step-down configuration. This type of balun can be used to transform the 300-Ω feedpoint impedance of a folded dipole down to 75 Ω for a coaxial transmission line. The radiation (or feedpoint) resistance of the dipole corresponds to
R_2
in the figure, and
R_1
represents the transmitter output impedance. For a multielement Yagi or a short (
$< \lambda/4$
) vertical monopole, one might employ the balun in a reverse configuration where now
R_1
represents the low feedpoint impedance of the antenna and
R_2
corresponds to the
Z_0
of a higher-impedance transmission line.

6.9.3 Stub Matching

The input impedance of a length of lossless transmission line is

$$Z_{\text{in}} = Z_0 \left[\frac{Z_L + jZ_0 \tan(\beta l)}{Z_0 + jZ_L \tan(\beta l)} \right] \quad (6.32)$$

where
Z_L
is the load impedance,
Z_0
is the characteristic impedance of the transmission line,
$\beta = 2\pi/\lambda$
, and *l* is the length of transmission line.
Z_{in}

is often complex—that is, having both a resistive term and a reactive part—even for a purely resistive load.

An important characteristic of transmission lines comes into view if we further examine Eq. (6.32). Suppose we are free to adjust the length (*l*) of the line so that the tangent functions in Eq. (6.32) are zero. Since that occurs when the argument *βl* is an integer multiples of *π* radians, *l* must be an integer multiple of

$\lambda/2$

. When

$l = n\lambda/2$

for integer values of *n*,

$Z_{\text{in}} = Z_L$

.

The input impedance can be found as a function of feedline length for a specific load using software or the graphical Smith chart technique of Sec. 6.10. When the load on a nearly lossless length of line is either a short circuit or an open circuit, the input impedance formula can be simplified. For a short circuit at the load end of the line,

$$Z_{\text{in}} = Z_0 \left[\frac{0 + jZ_0 \tan(\beta l)}{Z_0 + j0} \right] = jZ_0 \tan\left(\beta l\right) \quad (6.33)$$

For a short-circuit load,

Z_{in}

is purely reactive—alternating between capacitive and inductive reactance as the length of the short-circuited line is varied. For an open circuit at the far end of the line,

$$Z_{\text{in}} = Z_0 \left[\frac{1}{j \tan(\beta l)} \right] = -jZ_0 \cot\left(\beta l\right) \quad (6.34)$$

By measuring the input impedance with short and open circuit loads, the characteristic impedance and velocity factor can be determined. We multiply Eq. (6.33) by Eq. 6.34 to obtain

$$Z_{\text{in,short}}\, Z_{\text{in,open}} = [jZ_0 \tan(\beta l)][-\, Z_0 \cot(\beta l)] \quad (6.35)$$

Since
$\tan x = 1/\cot x$
, this reduces to

$$Z_{\text{in,short}}\, Z_{\text{in,open}} = -j^2\, Z_0^2 = Z_0^2 \quad (6.36)$$

We can determine the characteristic impedance of our unlabeled transmission line by taking the square root of the product of the short-circuit and open-circuit impedance measurements. From the ratio, we find that

$$\frac{Z_{\text{in,short}}}{Z_{\text{in,open}}} = \frac{jZ_0 \tan(\beta l)}{-\, jZ_0 \cot(\beta l)} = -\tan^2\left(\beta l\right) \quad (6.37)$$

Solving for the tangent of the electrical length,

$$\tan(\beta l) = \sqrt{\frac{-Z_{\text{in,short}}}{Z_{\text{in,open}}}} \quad (6.38)$$

With a known length of the line, we can find the wavenumber β and in turn the velocity factor

$$v_F = \frac{2\pi f}{\beta c} \quad (6.39)$$

Because a shorted or open length of transmission line always exhibits a pure reactance at the other end, such lines are often used to match an antenna to its transmission line. We refer to a short length of line used for this purpose as a stub. A specific reactance, either capacitive or inductive, can be obtained simply by adjusting the length of the stub.

For some applications, we can locate the stub right at the antenna feed terminals. But in most cases, in addition to canceling

the reactive part of the feedpoint impedance, we also want to change the resistive part to match the
Z_0
of the feedline. Simply adjusting the length of the stub cannot compensate for both the resistive and the reactive mismatches. A second adjustable variable is needed, the distance of the stub back toward the transmitter from the antenna feedpoint. Most often, a stub must be located at some specified distance from the antenna feedpoint.

For a variety of reasons, both electrical and mechanical, stubs are more often connected in parallel, across the feedline, than in series with one side of the feedline. In such cases, it is easiest to work with the reciprocal of the impedances involved, or admittances. Admittances of the form
$Y = G + jB$
consist of a shunt conductance *G* and a shunt susceptance *B*. The admittance is related to impedance by

$$Y = \frac{1}{Z} = \frac{1}{R + jX} = G + jB \quad (6.40)$$

To convert
$1/(R + jX)$
into admittance, we multiply both top and bottom by
$R - jX$
, the complex conjugate of
$R + jX$
. This gives us

$$\left[\frac{1}{R + jX}\right]\left[\frac{R - jX}{R - jX}\right] = \frac{R - jX}{R^2 + X^2} \quad (6.41)$$

The conductance and susceptance are then

$$G = \frac{R}{R^2 + X^2} \quad (6.42)$$

and

$$B = \frac{-jX}{R^2 + X^2} \quad (6.43)$$

These formulas relate resistance and reactance to conductance and susceptance.

To design a stub matching network for an antenna feed, we choose a length of line connected to the antenna so that the admittance looking into the line and antenna is

$$Y_{\text{in}} = Y_0 + jB \quad (6.44)$$

We then attach a stub in parallel with the line that has admittance

$$Y_{\text{stub}} = -jB \quad (6.45)$$

The length of the stub is designed using either Eq. (6.33) or Eq. (6.34) for a short-circuited or open-circuited stub to achieve the needed susceptance $-jB$. With the stub in parallel with the length of line, the total admittance is the sum of the line admittance and the stub admittance, which is $Y_{\text{in}} + Y_{\text{stub}} = Y_0$, representing the desired impedance match. The Smith Chart can be used to design stub matching networks graphically (Sec. 6.10).

A half-wavelength section of transmission repeats any load impedance, whether matched or not. Thus, for a single frequency or band of frequencies, a transmission line that is any integer multiple of $\lambda/2$ long can be used to remotely measure and monitor the feedpoint impedance of an antenna. The two caveats are:

6.9.4 Quarter-Wave Matching Section

A quarter-wavelength line exhibits a transforming property, converting low-resistance loads to higher ones, and capacitive reactances to inductive. The best way to understand and use this feature is with the Smith chart (explained in detail in Sec. 6.10) or a similar tool. One important application of this property is to transform a load impedance to a higher (or lower) main transmission line impedance. For instance, suppose an antenna for 10 MHz has a feedpoint impedance of 100 Ω resistive. It's not a very good match for 52-Ω coaxial cable, but a *λ*/4 section of 75-Ω RG-59 or RG-11 coaxial cable for television systems will transform the antenna feedpoint impedance from 100 to 52 Ω.

Impedances for the *λ*/4 transformer are calculated for resistive loads using

$$Z_L Z_0 = Z_S^2 \quad (6.46)$$

where

Z_L

is the load or antenna feedpoint impedance,

Z_S

is the characteristic impedance of the *λ*/4 section, and

Z_0

is the characteristic impedance of system transmission line we are matching to. Quarter wave matching sections work at a specific design center frequency. The length of the matching section is 1/4 of the wavelength in the transmission line, not the free space wavelength. The *λ*/4 transmission line transformer, shown in Fig. 6.32, is sometimes called a *Q*-section matching network.

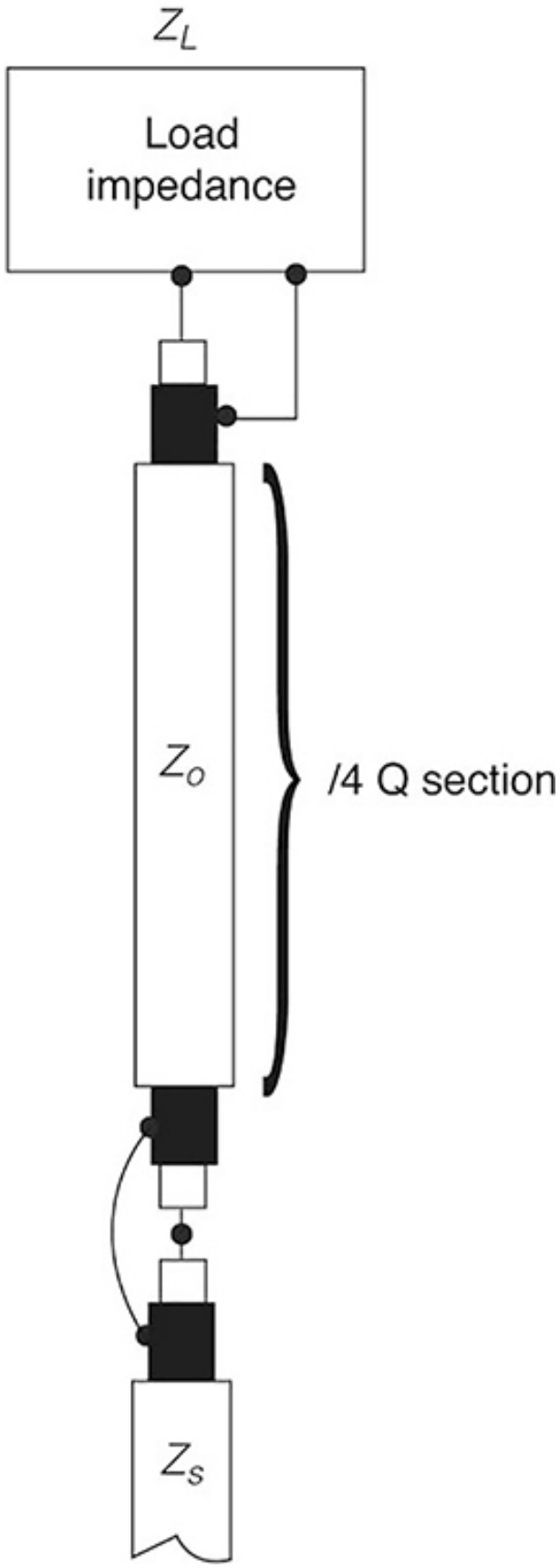

Z_L
Load
impedance
Z_o
/4 Q section
Z_s

FIGURE 6.32 Quarter-wavelength *Q* section.

Example 6.8 A 50-Ω source must be matched to a load impedance of 36 Ω. Find the characteristic impedance required of a *Q*-section matching network.

Solution

$$\begin{aligned} Z_S &= \sqrt{Z_0 Z_L} \\ &= \sqrt{(50\ \Omega)\ (36\ \Omega)} \\ &= 42\ \Omega \end{aligned}$$

For quarter-wave matching with coaxial lines, we must have a piece of transmission line with the required characteristic impedance. With open-wire transmission line systems, on the other hand, it is easy to achieve the correct impedance for the matching section. We use Eq. (6.46) to find a value for Z_S and then calculate the conductor spacing and diameter needed for a short length of a custom transmission line. Armed with information about the diameter of common wire sizes, the conductor spacing can be computed with

$$S = d \times 10^{(Z_S/276)} \quad (6.47)$$

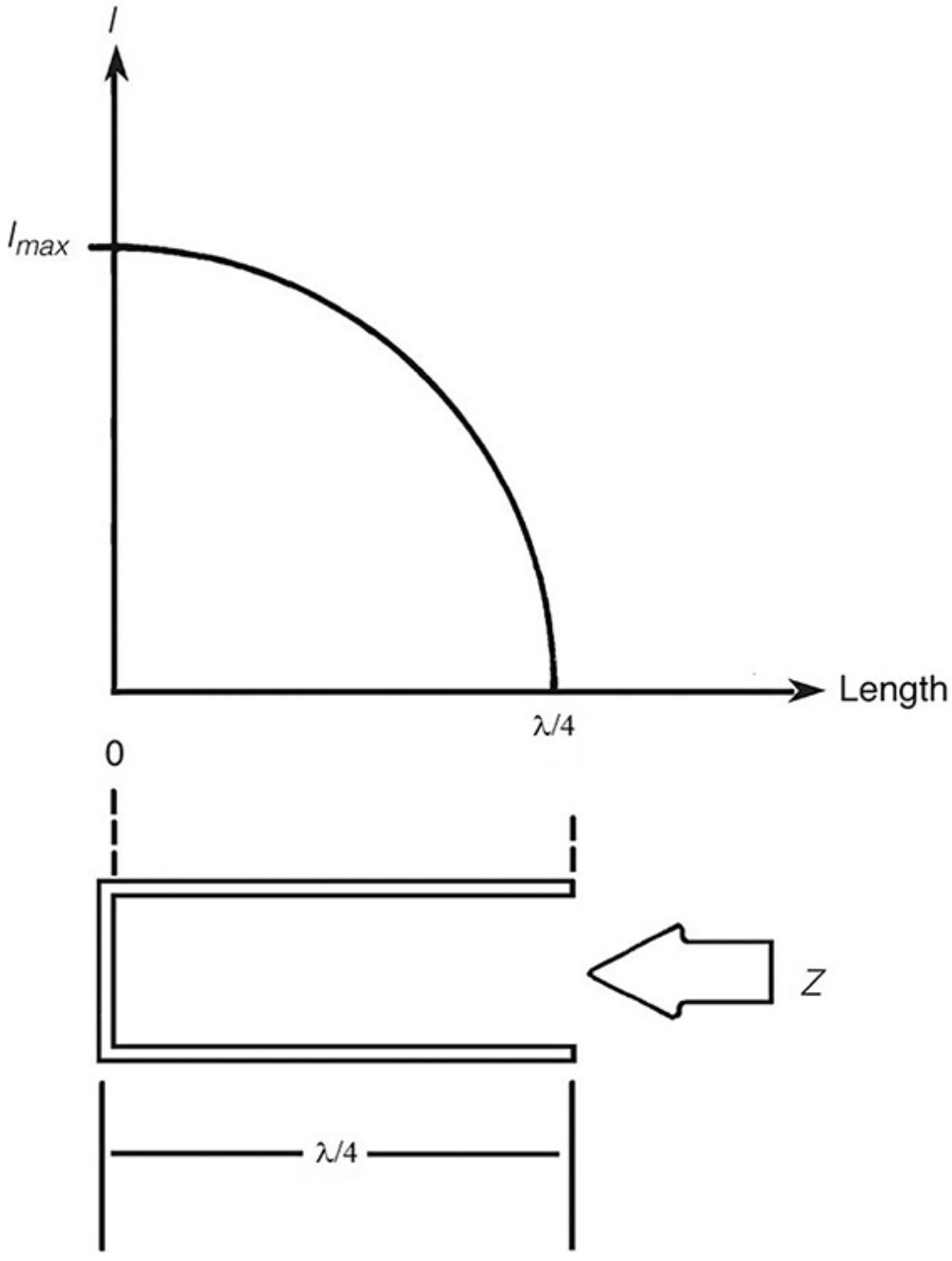
where *d* is conductor diameter expressed in same units as *S* and Z_S is the desired impedance for the matching section. Assuming air dielectric is used, the length of the quarter-wave section in feet can be calculated using 246/*f*, where *f* is in megahertz.

A quarter-wavelength shorted stub is a special case of the stub concept that finds particular application in microwave circuits. Figure 6.33 shows a quarter-wave stub and its current distribution.

The current is maximum across the short circuit load and zero at the terminals on the other end of the line. Because

$$Z = V/I$$

, when the current is zero the impedance becomes infinite. A quarter-wavelength stub has an infinite impedance at its resonant frequency and acts as an insulator. This concept may be hard to swallow, but the stub is a “metal insulator” at its design frequency.



F IGURE 6.33 A quarter-wavelength shorted stub transforms a short circuit load to an infinite input impedance.

6.9.5 Series Matching Section

The series matching section or series section transformer is a generalization of the *Q*-section that permits us to build an impedance transformer that overcomes the *Q*-section's limitations. According to *The ARRL Antenna Book*, with appropriate choices of transmission lines this form of transformer is capable of matching any load resistance between about 5 and 1200 Ω. In addition, the transformer section does not have to be located at the antenna feedpoint.

Figure 6.34 shows the layout of the series matching section. Three lengths of coaxial cable, L_1, L_2, and L_3, serve to form the entire feedline connecting the antenna to the transmitter. Lengths L_2 and the line to the transmitter (L_3, which is any convenient length and doesn't appear in the equations) have the same characteristic impedance, usually 75 Ω, while section L_1 has a different impedance. Only standard, easily obtainable values of impedance are used here.

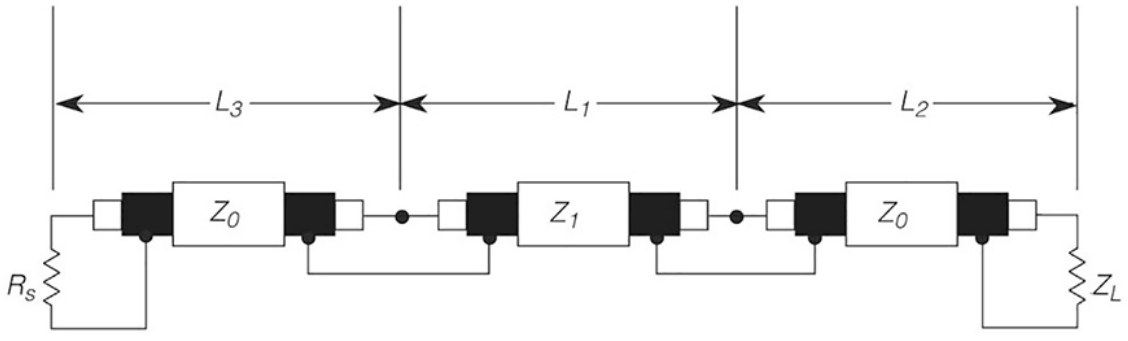


FIGURE 6.34 Series matching section.

The design procedure for this transformer consists of finding the correct lengths for L_1 and L_2. We must know the characteristic impedance of the two lines (75 Ω and 50 Ω, respectively, in this example), along with the complex feedpoint impedance of the antenna, $Z_L = R_L + jX_L$.

The first task in designing the transformer is to define normalized impedances:

$$N = \frac{Z_1}{Z_0} \quad (6.48)$$

$$R = \frac{R_L}{Z_0} \quad (6.49)$$

$$X = \frac{X_L}{Z_0} \quad (6.50)$$

Adopting ARRL notation and defining $A = \tan L_1$ and $B = \tan L_2$, the following equations can be written:

$$\tan L_2 = B = \pm\sqrt{\frac{(R-1)^2 + X^2}{R\{N - (1/N)^2\} - (R-1)^2 - X^2}} \quad (6.51)$$

$$\tan L_1 = A = \frac{\{N - R/N\}B + X}{R + XNB - 1} \quad (6.52)$$

The lengths of L_1

and
L_2
determined from Eqs. (6.51) and (6.52) are in electrical degrees. The physical lengths of the series sections are

$$L_1 = \frac{L_1(\text{degrees})\ \lambda_1}{360} \quad (6.53)$$

$$L_2 = \frac{L_2(\text{degrees})\ \lambda_2}{360} \quad (6.54)$$

where
λ_1
and
λ_2
are the wavelengths for each section of cable.

Although the sign of *B* can be selected as either positive or negative, the positive sign is preferred because a shorter matching section results. If the term inside the square root sign is negative, *B* is an imaginary number and the section cannot be constructed with the two-line impedances selected. The sign of *A* is fixed, but if it turns out to be negative, we add 180 degrees to the result. Series-section transformers can also be designed graphically with the help of the Smith charts of Sec. 6.10.

6.10 The Smith Chart

The mathematics of transmission lines can be cumbersome, especially when dealing with complex impedances and nonstandard situations. In 1939, Phillip H. Smith published a graphical tool for solving these problems. The Smith chart is a powerful learning tool for understanding and visualizing transmission line effects. Network analyzers used to measure impedances can be configured to show results on the Smith chart.

6.10.1 Smith Chart Components

The Smith chart is shown in Fig. 6.35. The Smith chart consists of a complex plane (used for reflection coefficients) overlaid with orthogonal circles that intersect each other at right angles (representing impedances). We will consider some of the Smith chart's key landmarks.

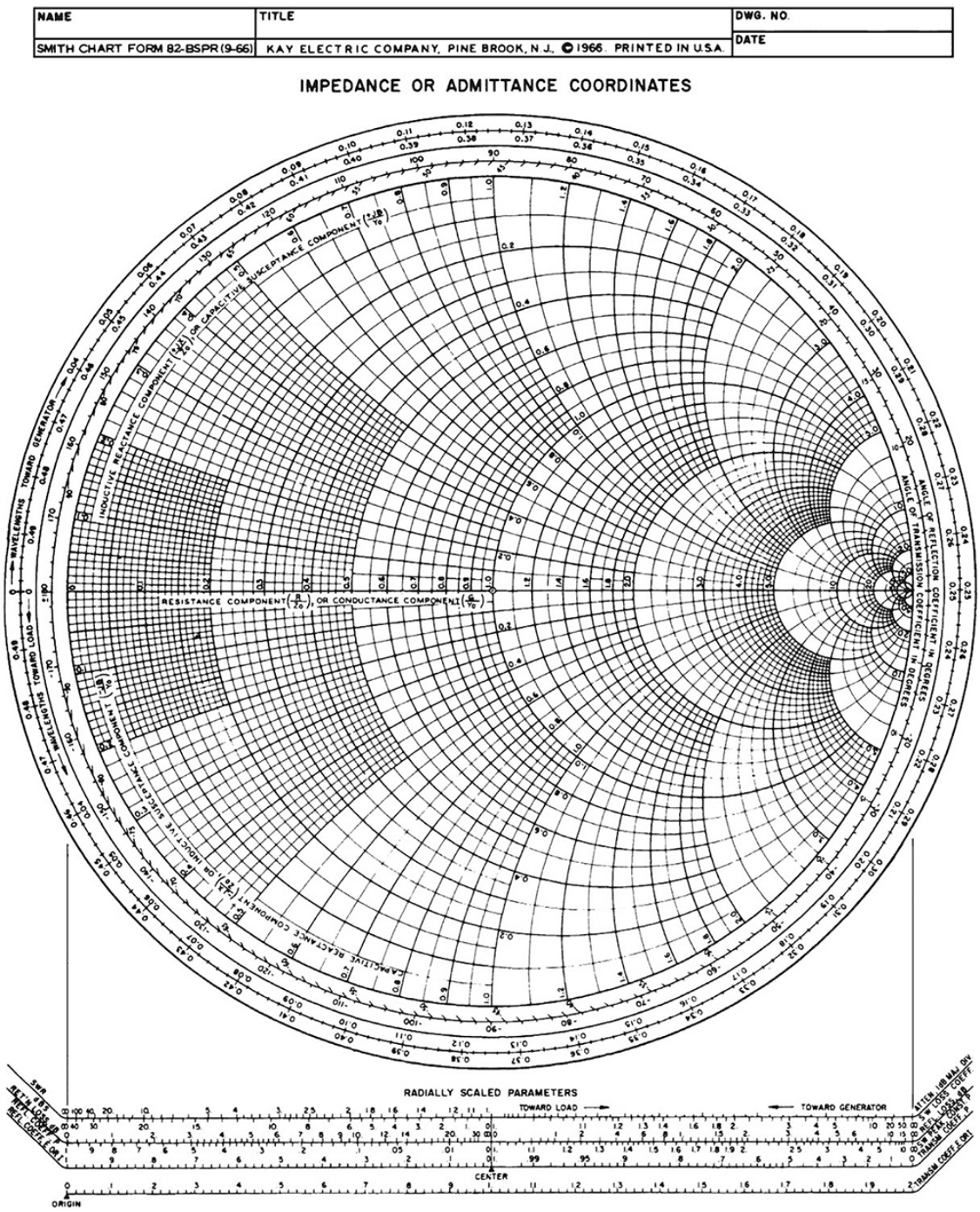


FIGURE 6.35 Smith chart.

6.10.2 Normalized Impedance

A complex impedance contains both resistance and reactance, and can be expressed in the form

$$Z = R + jX \quad (6.55)$$

where *R* is the resistive component and *X* is the reactive component of impedance. Inductive reactance (X_L) is positive (+) and capacitive reactance (X_C) is negative (−).

To make the Smith chart work for any characteristic impedance, we normalize impedances by the characteristic impedance of the line before plotting on the Smith chart. If the load impedance in a transmission line system is Z_L, the normalized load impedance is

$$z_L = \frac{Z_L}{Z_0} \quad (6.56)$$

The normalized impedance is a dimensionless complex number. By employing normalized impedances, you can use the Smith chart for virtually any practical combination of system impedances or load and source impedances, whether resistive, reactive, or complex.

The characteristic impedance Z_0 used to normalize impedances is called the system impedance. When there are different types of transmission lines, the system impedance is usually chosen to be the characteristic impedance of the predominant transmission line or the output impedance of the transmitter. For most microwave and RF systems the system impedance is standardized at 50 Ω. The system impedance has a normalized value of $z = Z_0/Z_0 = 1.0$ and is at the center of the chart (point A in Fig. 6.36).

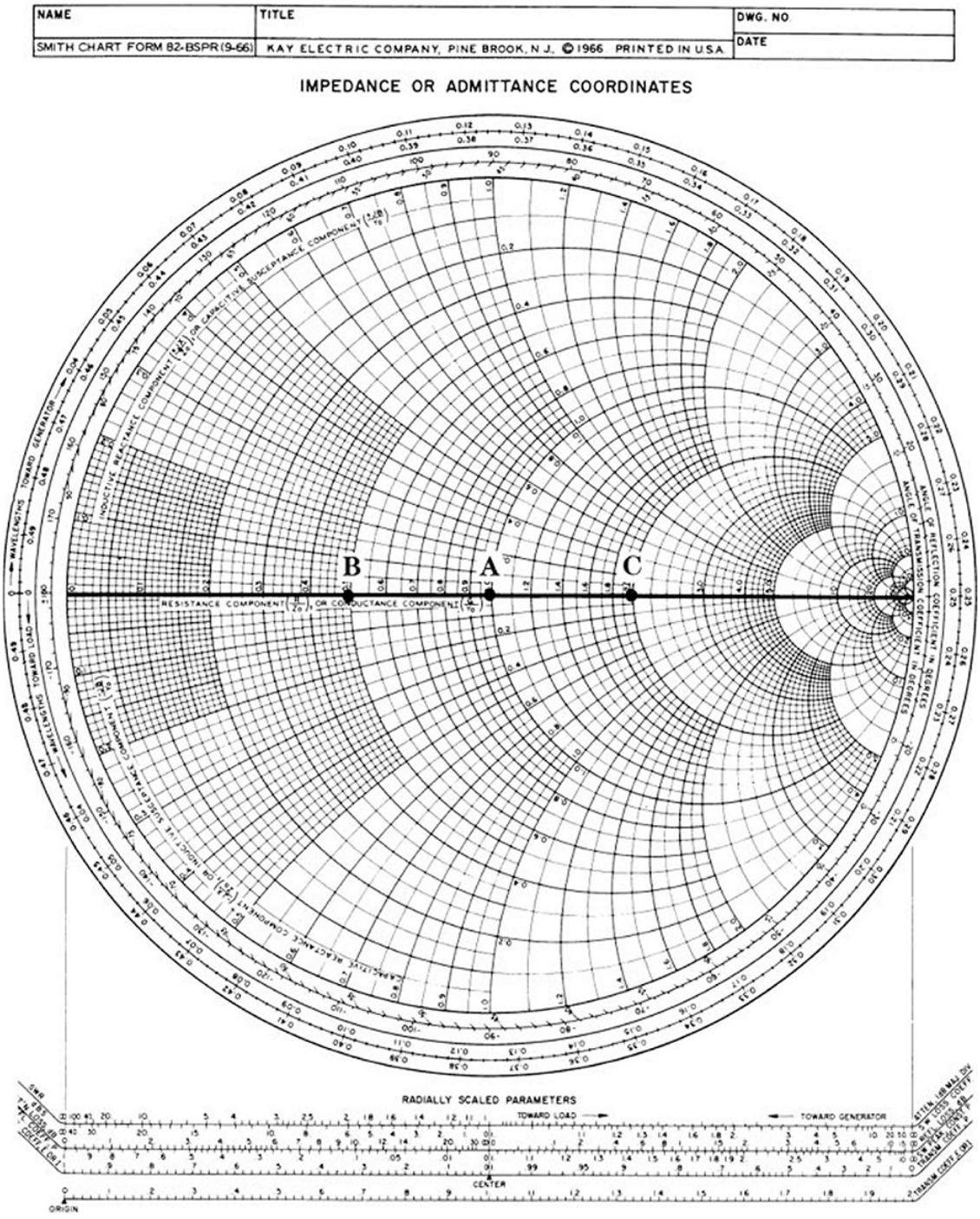


FIGURE 6.36 Pure resistance line.

Conversion of the normalized impedance to actual impedance values is done by multiplying the normalized impedance by the system impedance. For example, if the resistive component of a

normalized impedance in a 50-Ω system is 0.45, then the actual impedance is

$$\begin{aligned} \mathrm{Z} &= zZ_0 \\ &= (0.45)(50\,\Omega) \quad (6.57) \\ &= 22.5\ \Omega \end{aligned}$$

6.10.3 Pure Resistance Line

Figure 6.36 highlights the horizontal axis of the Smith chart. This line is called the pure resistance and corresponds to all real impedances of the form
$Z = R + j0$
. Points along the pure resistance line represent all possible impedances where
$X = 0$
.

To the left of the center of the Smith chart on the pure resistance line are decimal fractional values used to denote impedances less than the system impedance. In a 50-Ω transmission line system with a 25-Ω load impedance, the normalized value of the load impedance is 25 Ω⁄50 Ω, or 0.50 (point B in Fig. 6.36). Points to the right of 1.0 are greater than 1 and denote impedances that are higher than the system impedance. In a 50-Ω system connected to a 100-Ω resistive load, the normalized impedance at the load end of the transmission line is 100 Ω⁄50 Ω, or 2.0; this value is shown as point C in Fig. 6.36.

6.10.4 Constant Resistance Circles

All impedances with a specific value of the resistance (i.e., the real part of the complex impedance) line in a circle called a constant resistance circle. In other words, if we fix the resistance and change the reactance, the point on the Smith chart moves around a constant resistance circle. Several of these circles are shown highlighted in Fig. 6.37. These circles are tangent to the point
$R = \infty$

at the right-hand extreme of the pure resistance line.

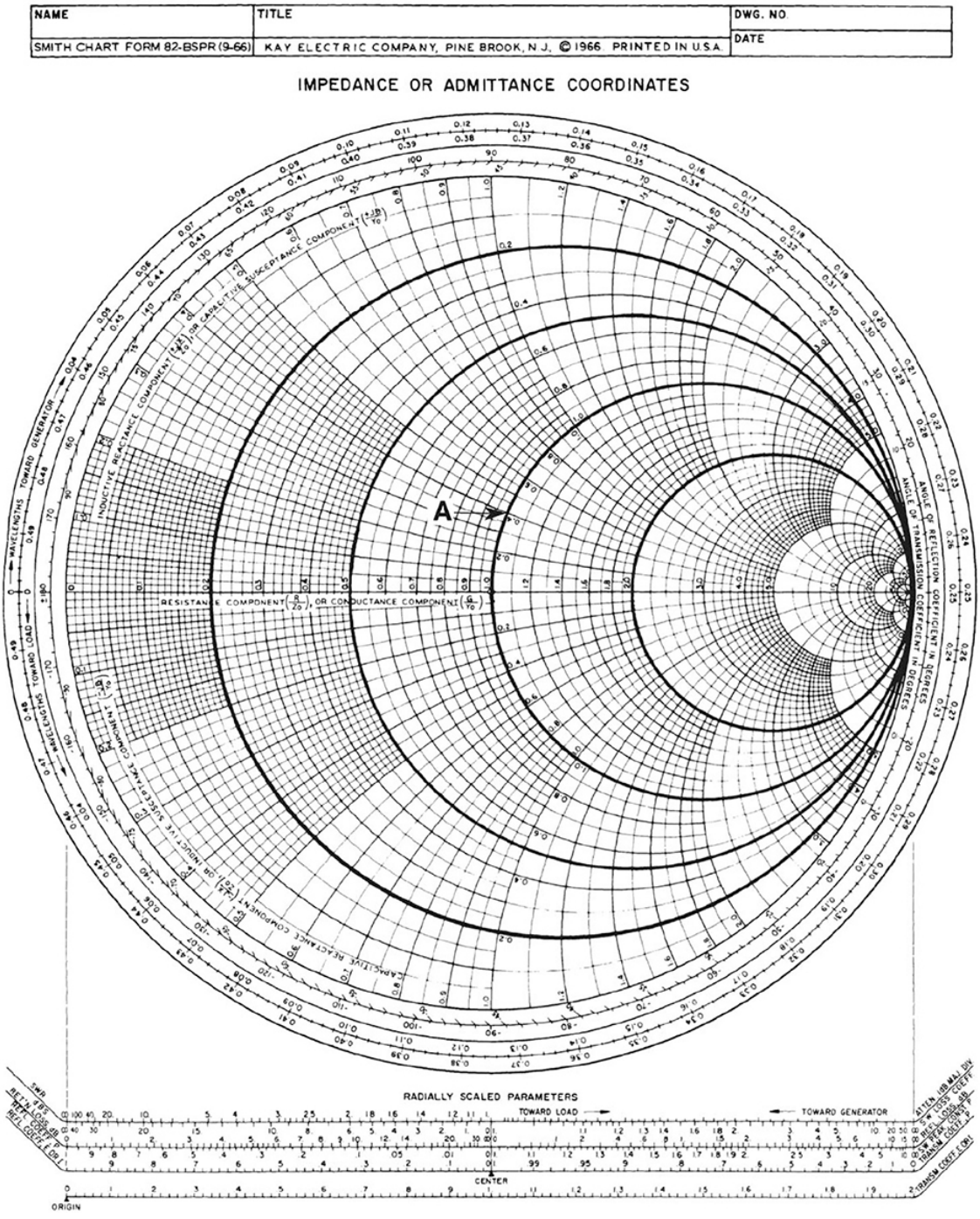


FIGURE 6.37 Constant resistance circles.

When we construct complex impedances (for which X is nonzero) on the Smith chart, all points on one of these circles will have an

identical resistive component. Circle A, for example, passes through the center of the chart, so it has a normalized constant resistance of 1.0. Impedances that are pure resistances (i.e., $z = r + j0$) will fall at the intersection of a constant resistance circle and the pure resistance line, and complex impedances (i.e., $X \neq 0$) will appear at any other points on that same constant resistance circle. In Fig. 6.37, circle A passes through the center of the chart, so it represents all points on the chart with a normalized resistance of 1.0. This particular circle is sometimes called the unity resistance circle.

6.10.5 Constant Reactance Circles

Constant reactance circles are highlighted in Figs. 6.38 and 6.39. Circles (or circle segments) above the pure resistance line (Fig. 6.38) represent inductive reactance ($X > 0$), while circles (or segments) below the pure resistance line (Fig. 6.39) represent capacitive reactance ($X < 0$). Circle A corresponds to a normalized reactance of 0.80.

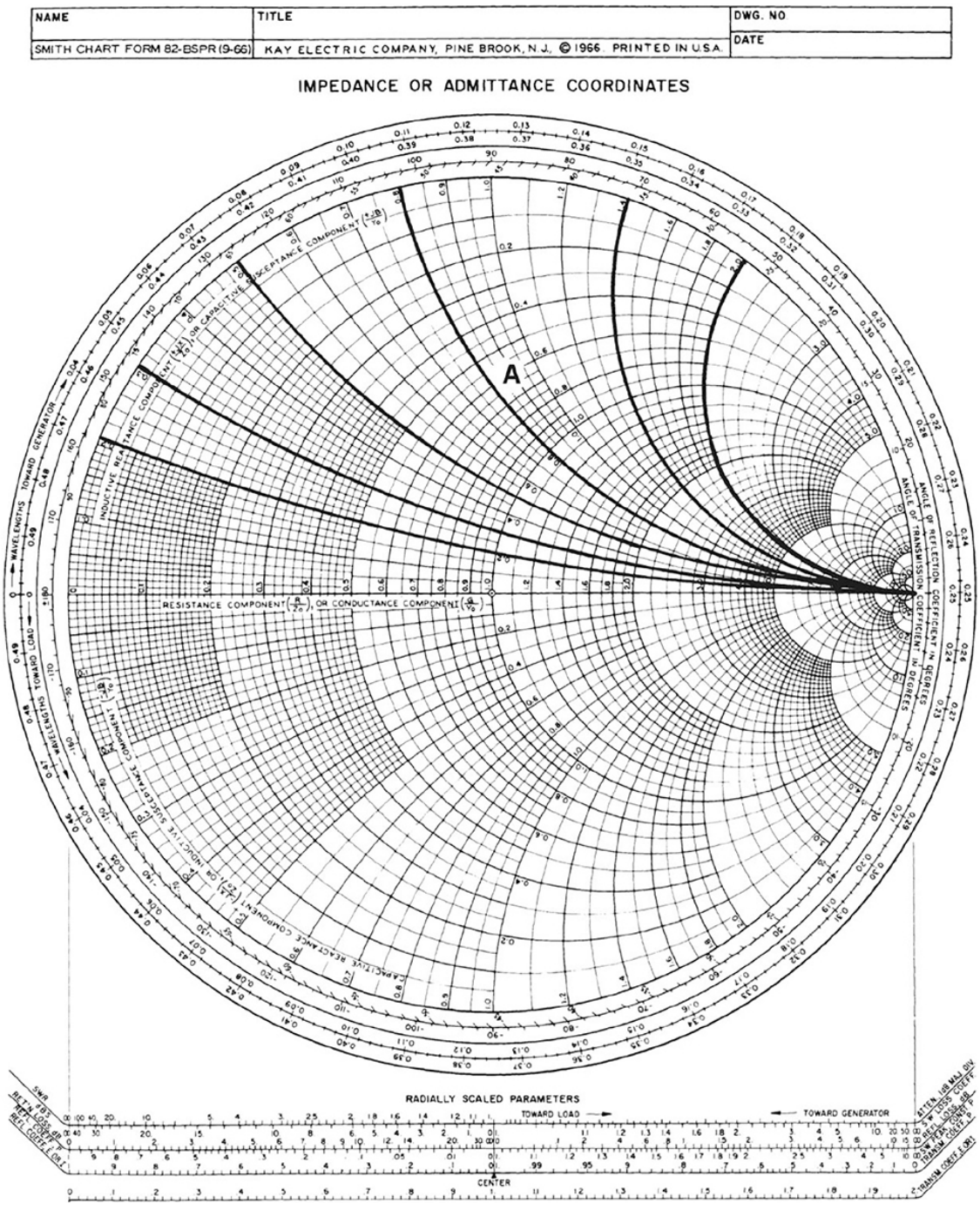
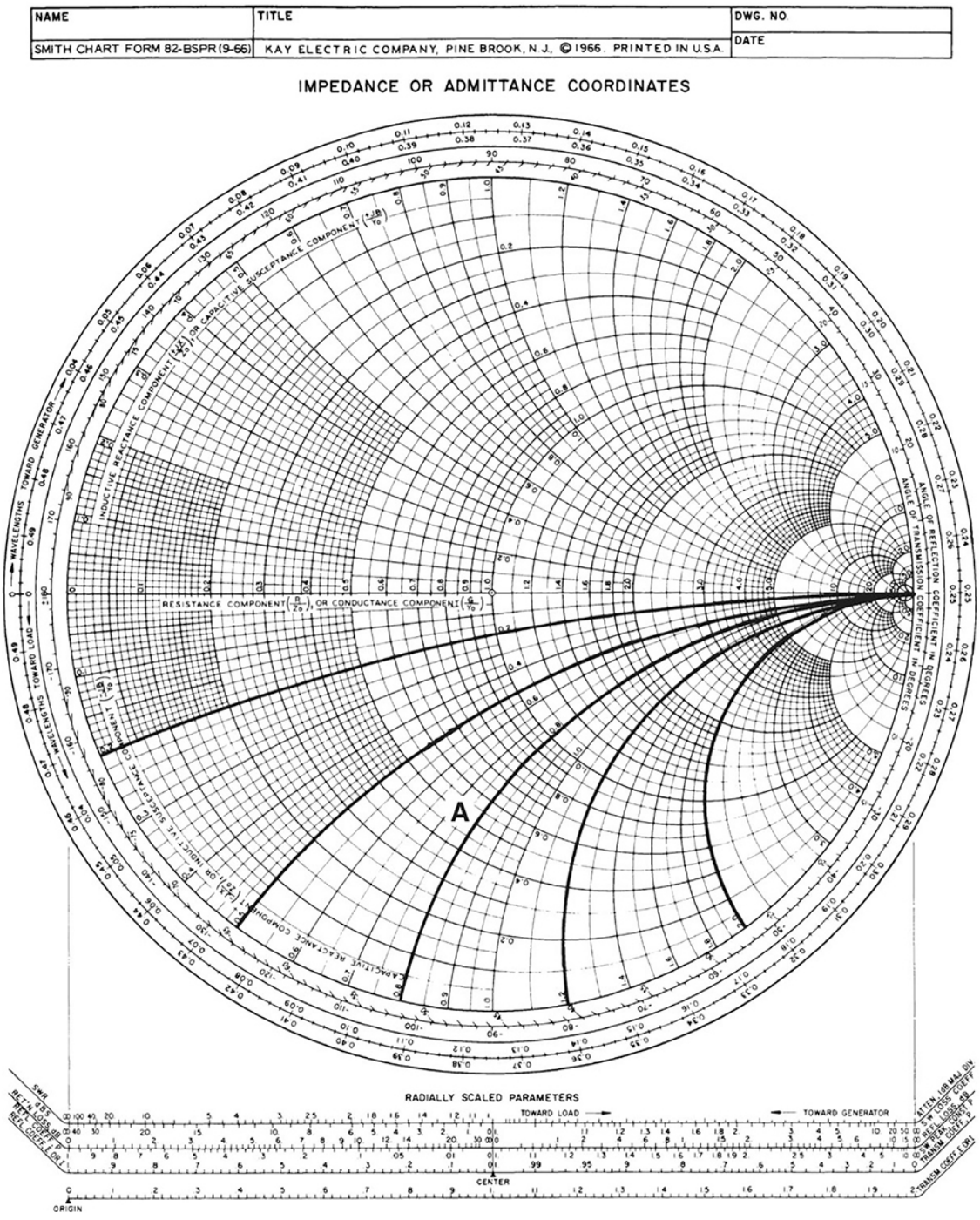


FIGURE 6.38 Constant inductive reactance lines.



F IGURE 6.39 Constant capacitive reactance lines.

The outermost circle of the Smith chart is called the pure reactance circle. Points along the outer circle represent reactance only. The real resistance is zero everywhere on this outer circle and all impedances along it are of the form

$$z = 0 \pm jx$$

.

6.10.6 Plotting Impedance and Admittance

Let's use Fig. 6.40 to see how to plot any impedance and admittance on the Smith chart. Consider an example in which system impedance
Z_0
is 50 Ω, and the load impedance is
$Z_L = 95 + j55\ \Omega$
. The load impedance normalizes to

$$\begin{aligned} z_L &= \frac{Z_L}{Z_0} \\ &= \frac{95 + \ j55\ \Omega}{50\ \Omega} \quad (6.58) \\ &= 1.9 + j1.1 \end{aligned}$$

To locate this point on the Smith chart, we find the real value, 1.9, along the pure resistance line. Since the reactive part is positive, we move up along the constant resistance circle until we reach the constant inductive reactance line marked 1.1 near the outer circle of the Smith chart. This point represents the normalized impedance
$1.9 + j1.1$
.

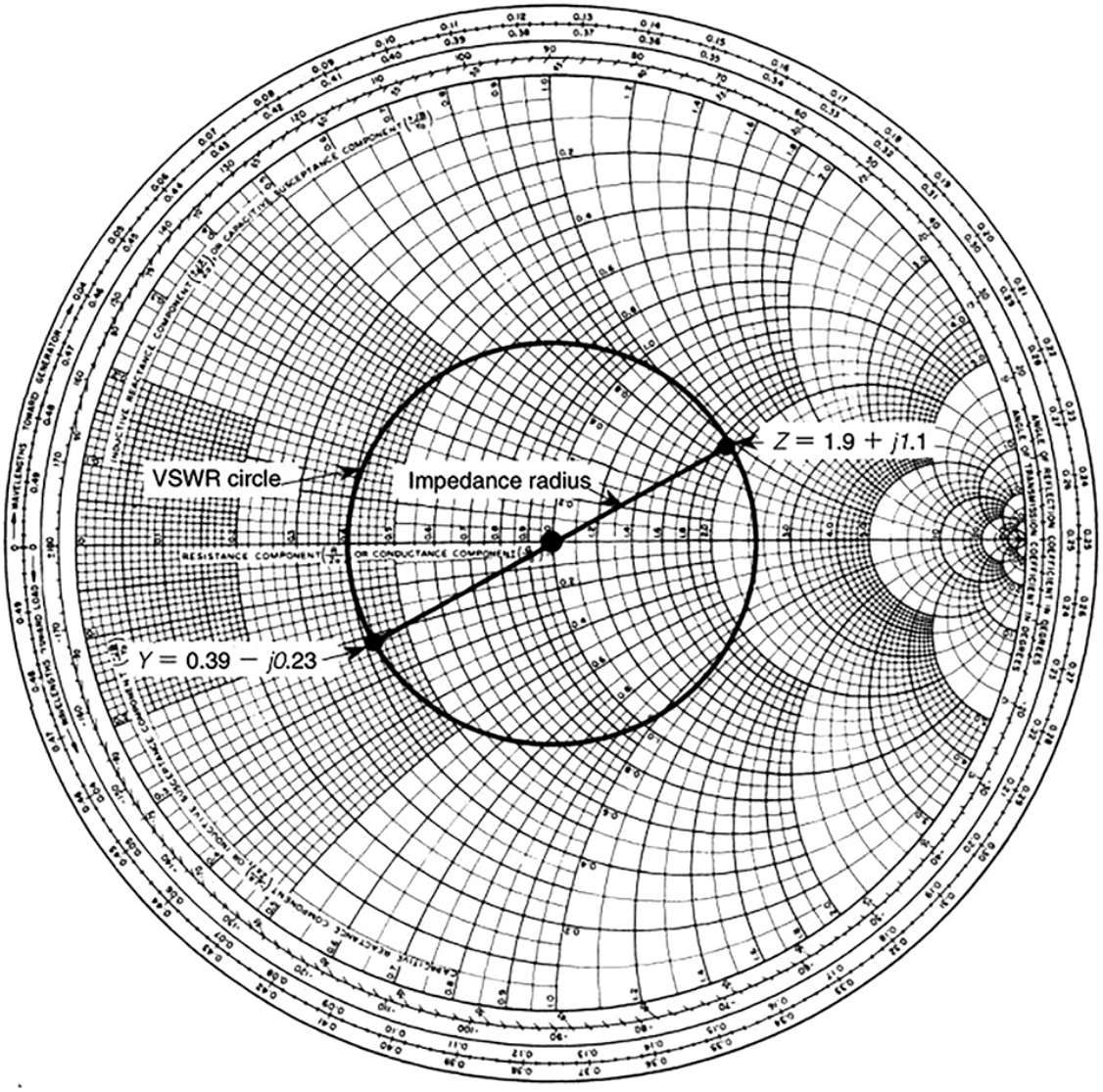


FIGURE 6.40 VSWR circle, impedance radius, and impedance to admittance conversion.

Once we have the impedance marked on the Smith chart, we can add other useful graphical features. An impedance radius is constructed by drawing a line from the point represented by the normalized load impedance,

$1.9 + j1.1$

, to the point represented by the normalized system impedance (1.0, 0.0) in the center of the chart. A circle, called the VSWR circle, is constructed by drawing a circle with this radius centered on the location (1.0,0.0).

We can also use the Smith chart to work with admittances. The normalized admittance is

$$y = \frac{1}{z} \quad (6.59)$$

On the Smith chart, the admittance is located exactly opposite, or 180 degrees from, its corresponding impedance. If the radius line connecting the center of the Smith chart to the impedance point is doubled to become a diameter, the coordinates of the point at the opposite end of the line from the original impedance point describe the normalized admittance corresponding to the reciprocal of the original normalized impedance. By extending a simple radius to a diameter, we have solved Eq. (6.59) graphically.

Example 6.9 To show the value of the Smith chart in easily solving real problems, let's first find the admittance by hand. Using the normalized impedance of Eq. (6.58), the normalized admittance is

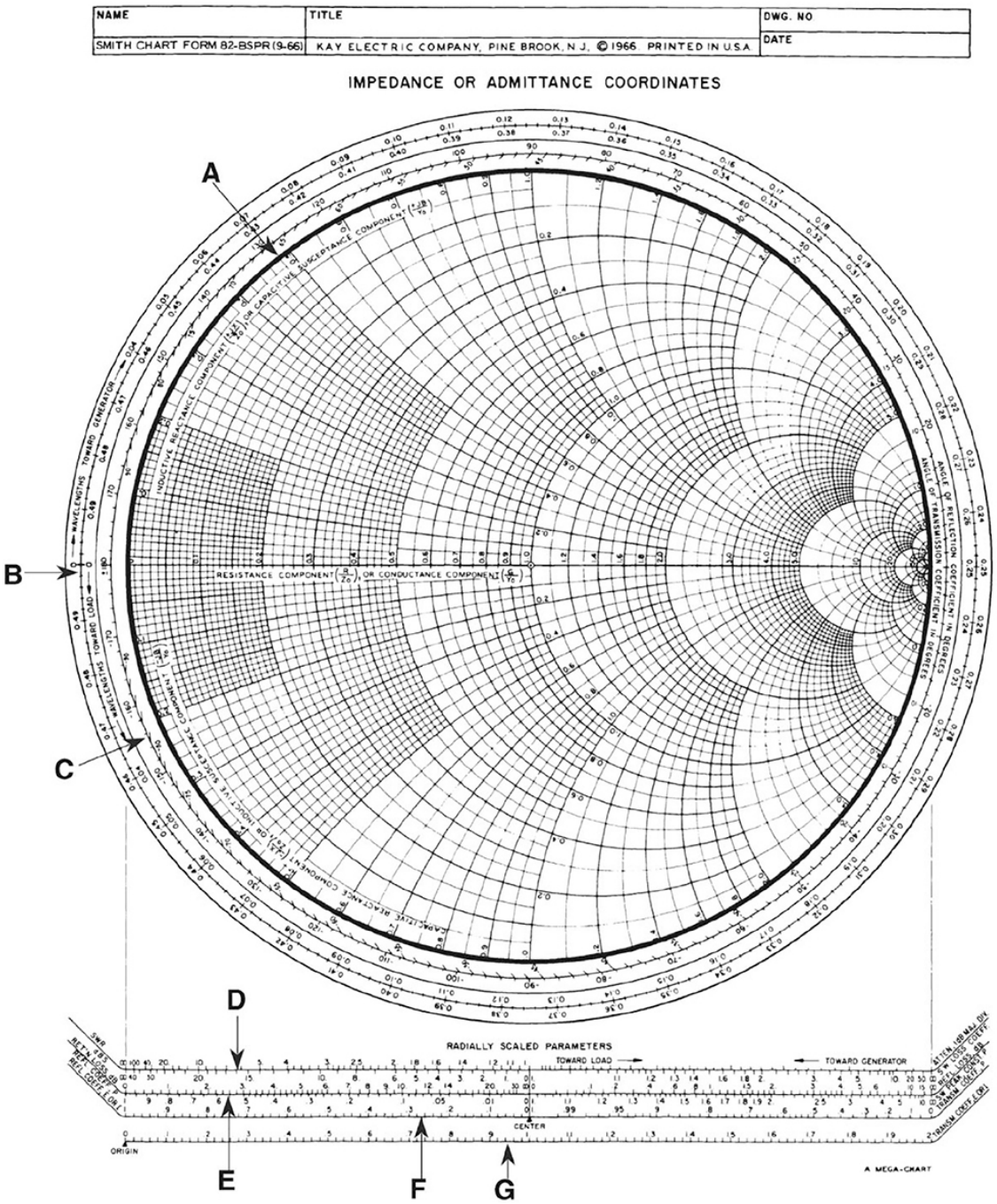
$$y = \frac{1}{z} = 0.39 - j0.23$$

Compare this with the coordinates of the left-hand end of the line in Fig. 6.40. Extend the impedance radius through the 1.0,0.0 center point until it intersects the VSWR circle again. If we read the values for the constant conductance circle and the inductive susceptance line that intersect at this point from the Smith chart, they are 0.39 and -0.23, as expected.

6.10.7 Transmission Line Distance Scale

The standard Smith chart contains three concentric calibrated circles on the outer perimeter of the chart. As we move along a transmission line, the impedance moves around the VSWR circle. The distance in wavelengths moved along the transmission line is given by Circle B in Fig. 6.41. Clockwise motion is toward the generator or transmitter and counterclockwise is toward the load or

antenna. Circle B is important when transforming an impedance at one point on a transmission line to the resulting impedance at another point some distance away.



F IGURE 6.41 Pure reactance circle (A), transmission line distance scale (B), angle of transmission coefficient (C), and radially scaled parameters (D-G).

There are two scales on the wavelength circle (B in Fig 26.4*c*). Both have their zero origin at the left-hand extreme of the pure resistance line and one complete revolution represents a distance of one half-wavelength along the line. The scales are calibrated linearly from 0 through 0.50 such that these two points are identical with each other on the circle. Starting at the zero point and traveling 360 degrees around the circle brings one back to zero, which represents one half-wavelength or 0.5λ distance along the line. The wavelength scales span the same range (0 to 0.50λ) but are marked in opposite directions. The outer scale is calibrated clockwise and it represents wavelengths toward the generator; the inner scale is calibrated counterclockwise and represents wavelengths toward the load.

The angle of transmission coefficient and angle of reflection coefficient scales are shown in circle C in Fig. 6.41. These scales report the relative phase angle between reflected and incident waves at each point on the line resulting from a given load impedance. Recall from transmission line theory (see Chap. 6) that a short circuit ($R = 0$) at the load end of the line results in the reflected voltage headed back toward the generator being 180 degrees out of phase with the incident voltage at the reflection point, and an open line (i.e., infinite impedance) results in the reflected voltage headed back toward the generator being in phase (i.e., 0 degrees) with the incident voltage at the load. On the Smith chart, a short circuit at the load corresponds to the left-hand end of the pure resistance line, where 180 degrees is marked on circle C. An open circuit at the load corresponds to the right-hand end of the pure resistance line, where 0 degrees is marked. The upper half-circle is calibrated 0 to $+180$ degrees, and the bottom half-circle is calibrated 0 to -180 degrees, reflecting how the load impedance is transformed into either inductive or capacitive reactance, respectively, depending on

how far back from the load the observation point has moved along the transmission line.

6.10.8 Radially Scaled Parameters

There are six scales laid out on four lines (D through G in Fig. 6.41 and in expanded form in Fig. 6.42) at the bottom of the Smith chart. These scales are called the radially scaled parameters. With these scales, we can determine such factors as VSWR as a ratio or in dB, return loss in decibels, the magnitude of the voltage or current reflection coefficient, and the power reflection coefficient.

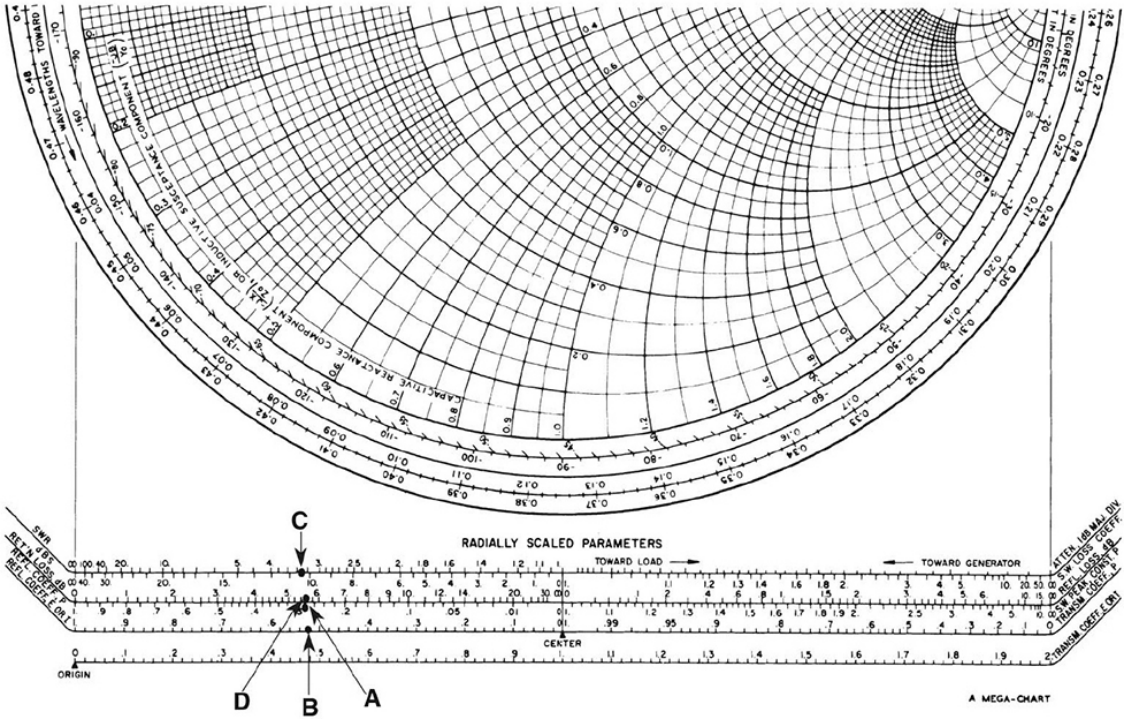


FIGURE 6.42 Radially scaled parameters.

As discussed in detail earlier in this chapter, the reflection coefficient Γ is defined as the ratio of the reflected wave to the incident (or forward) wave. The reflection coefficient is

$$\Gamma = \frac{V_{\text{ref}}}{V_{\text{fwd}}} \quad (6.60)$$

Power is proportional to the square of voltage or current, so the reverse to forward power ratio or power reflection coefficient is $\rho = |\Gamma|^2$

.

Example 6.10 10W of microwave RF power is applied to a lossless transmission line; 2.8W of the 10W total incident power is reflected from the mismatched load. Calculate the reflection coefficient, VSWR, and return loss.

Solution

The power reflection coefficient is

$$\rho = \frac{P_{\text{ref}}}{P_{\text{frd}}} = \frac{2.8\ \text{W}}{10\ \text{W}} = 0.28$$

The magnitude of the voltage reflection coefficient Γ is found by taking the square root of the power reflection coefficient, so in this example it is equal to 0.529. These points are plotted at A (for power) and B (for voltage) in Fig. 6.42.

The voltage standing wave ration (VSWR) on a transmission line is related to the reflection coefficient at the load end of the line by

$$\text{VSWR} = \frac{1 + |\Gamma|}{1 - |\Gamma|} \quad (6.61)$$

For the example values,

$$\text{VSWR} = \frac{1 + \sqrt{0.28}}{1 - \sqrt{0.28}} = 3.25$$

or, in decibel form,

$$\text{VSWR (dB)} = 20 \log_{10} \text{VSWR} = 20 \log_{10}(3.25) = 10.2\ \text{dB}$$

These points are plotted at C in Fig. 6.42.

The loss of power in a transmission line because of reflection from the load is called return loss. Return loss is found from

$$\mathrm{RL} = -20 \log_{10} |\Gamma| \quad (6.62)$$

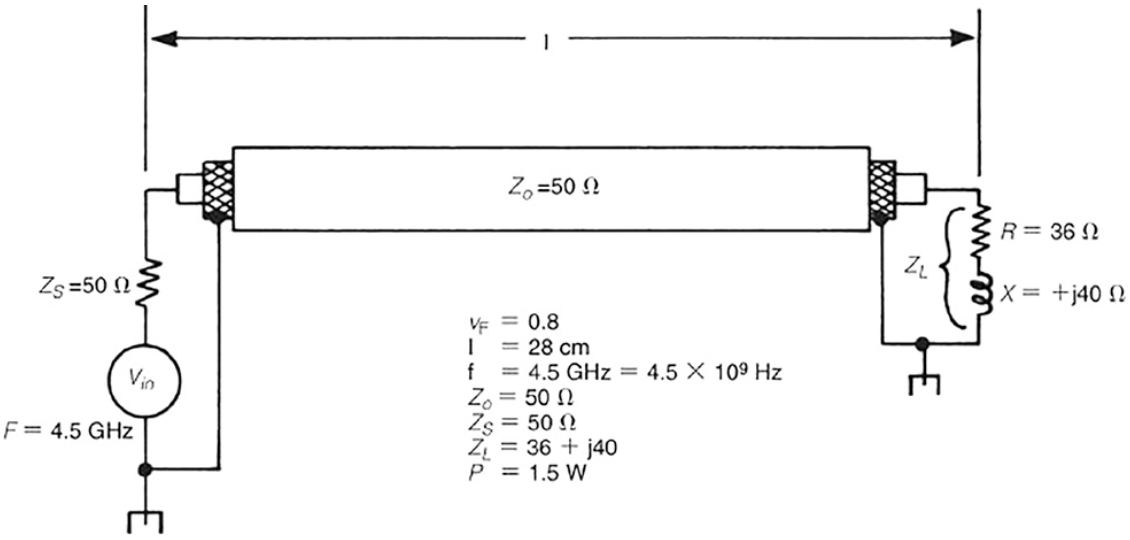
and, for our example in which
$|\Gamma|^2 = 0.28$
,

$$\mathrm{RL} = -10 \log_{10} 0.28 = 5.53 \text{ dB}$$

This point is shown as D in Fig. 6.42.

6.10.9 Transmission Line Calculations

Figure 6.43 shows a 50-Ω transmission line connected to a complex load impedance
$Z_L = 36 + j40 \ \Omega$
. The transmission line has a velocity factor
v_F
of 0.80. The length of the transmission line is 28 cm. The generator, a voltage source in this example with open circuit voltage
V_{in}
, is operated at a frequency of 4.5 GHz and produces a power output of 1.5 W. We will use the Smith chart to find the input impedance, input admittance, VSWR, return loss, and power absorbed by the load for this transmission line system.

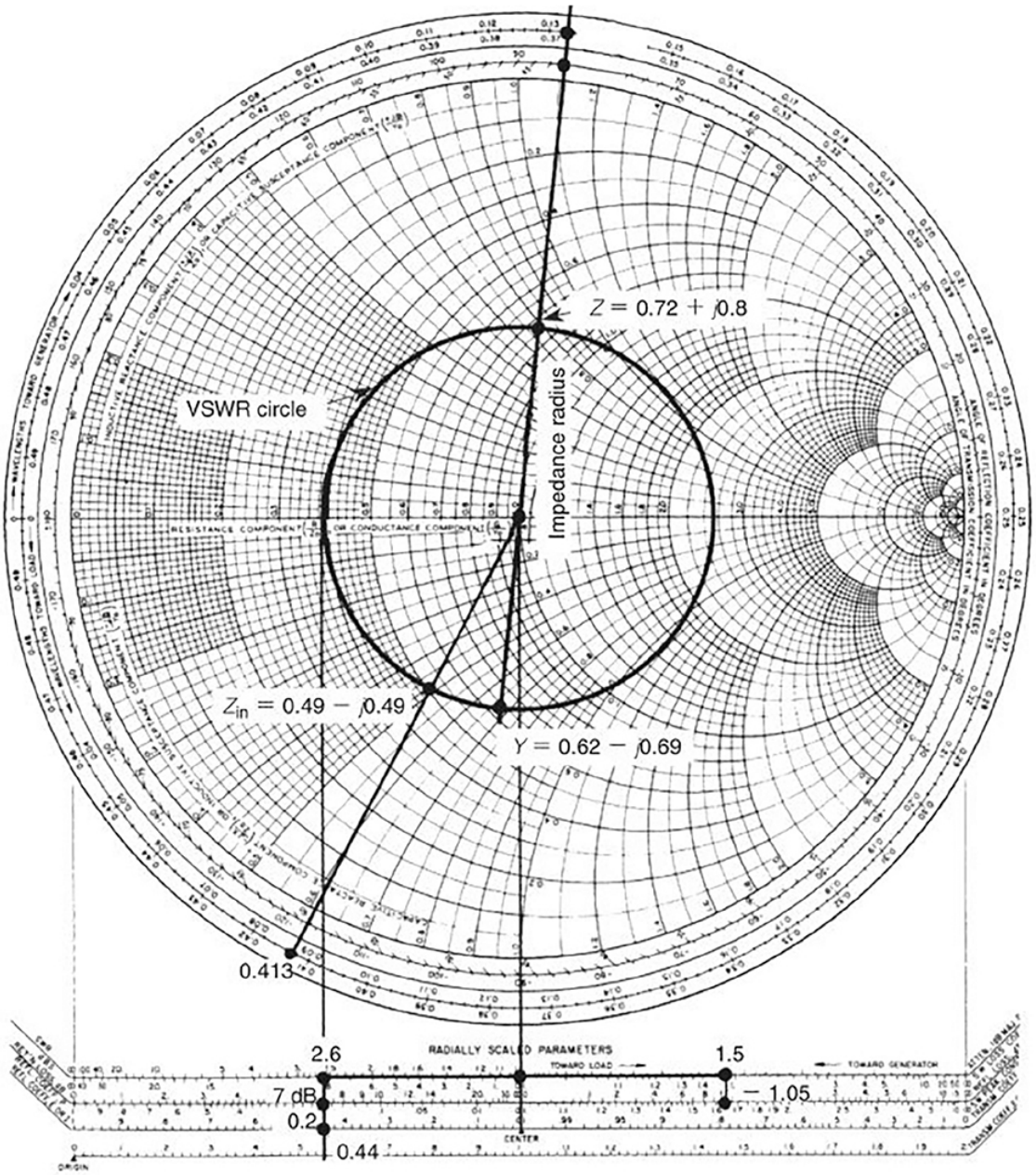


F IGURE 6.43 Transmission line and load circuit.

First, we normalize the load impedance by dividing it by the system impedance (in this case, $Z_0 = 50\ \Omega$):

$$z_L = \frac{36 + j40\ \Omega}{50\ \Omega} = 0.72 + j0.8$$

We locate the circle of constant resistance that goes through the point (0.72,0.0) in Fig. 6.44 and follow it until it intersects the $+j0.8$ constant reactance circle. This point represents the normalized load impedance $z_L = 0.72 + j0.80$.



F IGURE 6.44 Smith chart for Example 6.10 (transmission line parameters).

Next we construct a VSWR circle centered on the point (1.0,0.0) and having an impedance radius equal to the line connecting 1.0,0.0 (at the center of the chart) and the

$0.72 + j0.8$

point.

At a frequency of 4.5 GHz, the wavelength assuming a velocity factor of 0.80 is

$$\lambda = \frac{cv_F}{f} = \frac{(3 \times 10^8 \text{ m/s})(0.80)}{4.5 \times 10^9 \text{ Hz}} = 5.3 \text{ cm}$$

A half-wavelength on the line is 2.65 cm. The line is 28 cm/5.3 cm or 5.28 wavelengths long. With a straightedge, we draw a line from the center (1.0,0.0) through the load impedance and extend it to the outermost circle. It should intersect that circle at 0.1325 (or 0.3675). Because one complete revolution around this circle represents a half-wavelength, 5.28 wavelengths from this point represents 10 revolutions plus 0.28 wavelengths more. The residual 0.28 wavelength figure is added to 0.1325 to obtain $0.1325 + 0.28 = 0.413$. We locate the point 0.413 on the circle, making sure to use the same markings that the original line intersected at 0.1325, and place a mark there. We draw a line from 0.413 to the center of the circle and see that it intersects the VSWR circle at $0.49 - j0.49$. This is the normalized input impedance z_{in} looking into the line.

To find the actual impedance represented by the normalized input impedance we have read off the chart, we unnormalize the impedance by multiplying the result by Z_0 to obtain

$$Z_{\text{in}} = (0.49 - j0.49)(50\ \Omega) = 24.5 - j24.5\ \Omega$$

This is the impedance that must be matched at the generator end of the transmission line.

The admittance of the load is the reciprocal of the load impedance and is found by extending the impedance radius through the center of the VSWR circle until it intersects the circle again on

the other side (as an impedance diameter). The normalized input admittance is found to be
$y = 0.62 - j0.69$
. Confirming the solution mathematically,

$$y = \frac{1}{z} = \frac{1}{0.72 + j0.80} = 0.62 - j0.69$$

The VSWR at the generator end of the line is found by transferring the "impedance radius" of the VSWR circle to the radial scales below the circle chart. At the leftmost extension of the radius circle, drop a line vertically to the topmost of the radially scaled parameters at the bottom of the chart. For this example, we find an input VSWR of approximately 2.6:1. In decibel form, the VSWR is 8.3 dB (next scale down from VSWR), and this is confirmed by

$$\text{VSWR (dB)} = 20 \log_{10} 2.6 = 8.3 \text{ dB}$$

The power reflection coefficient can be read as well and is found to be 0.2.

The return loss can found by reading the "RET'N LOSS, dB" line where the same vertical line intersects it. The value is found to be approximately 7 dB, which is confirmed with the use of Eq. (6.62):

$$\text{RL} = 10 \log_{10} |\Gamma|^2 = 10 \log_{10} 0.2 = 7 \text{ dB}$$

The angle of the reflection coefficient is found from the outer circles of the Smith chart. The line connecting the center to the load impedance (
$z = 0.72 + j0.8$
) is extended to the angle of reflection coefficient in degrees circle, and intersects it at approximately 84 degrees. In polar form, the reflection coefficient is
$0.44 \angle 84$
degrees.

The reflection loss measured the fraction of RF power reflected back down the transmission line from the load. The difference between incident power supplied by the generator (1.5 W in this example) and the reflected power is the input power or the power absorbed by the antenna:

$$P_{\text{in}} = P_{\text{fwd}} - P_{\text{ref}} \quad (6.63)$$

The reflected power is
$0.2P_{\text{fwd}} = 0.3 \text{ W}$
. The power absorbed by the load (
P_{L}
) is the difference between incident power
P_{fwd}
and the reflected power
P_{ref}
. If 0.3 W is reflected, then that means the absorbed power is
$P_{\text{in}} = 1.5 - 0.3 = 1.2 \text{ W}$
.

6.10.10 Stub Matching with the Smith Chart

We can use the Smith chart to design stub matching networks. As shown in Fig. 6.45, we can use a short circuit loaded length of line attached at some distance from the load to match a load impedance to a transmission line characteristic impedance.

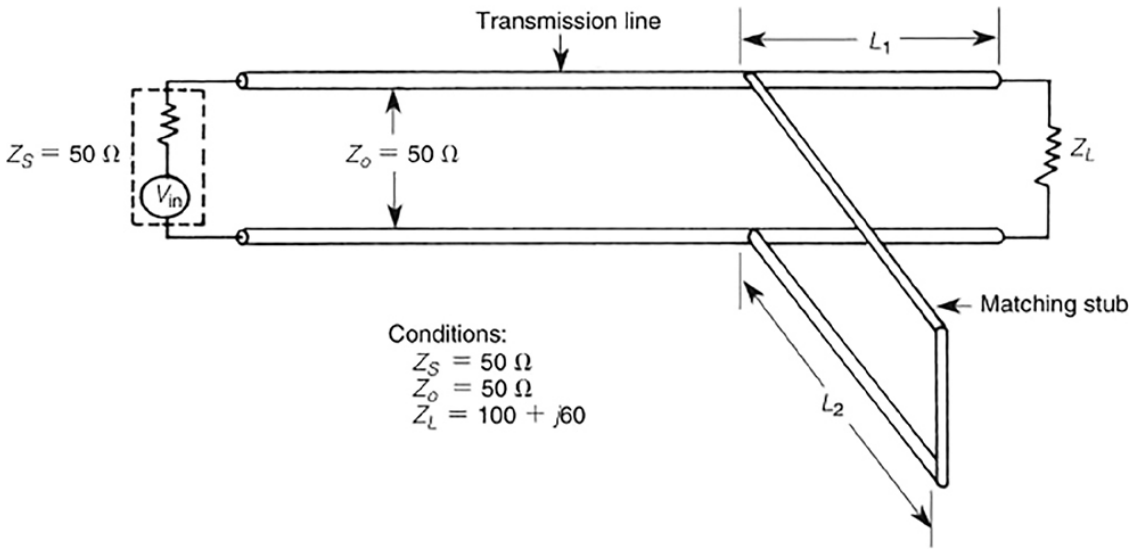


FIGURE 6.45 Stub matching example.

The stub when approximated as a lossless line exhibits a pure reactance that can vary from

$-\infty$

to

$+\infty$

, depending upon its length (and repeating every

$\lambda/2$

). A stub attached across a load impedance can cancel any possible reactive component. Because the stub must also transform the resistive part of the load impedance up or down to match the system impedance (

Z_0

), another adjustable stub parameter is required. This turns out to be the stub attachment point, or the distance

L_1

from the load back toward the transmitter or generator. Our goal is to design the length

L_1

that gives the location of the stub and the stub length

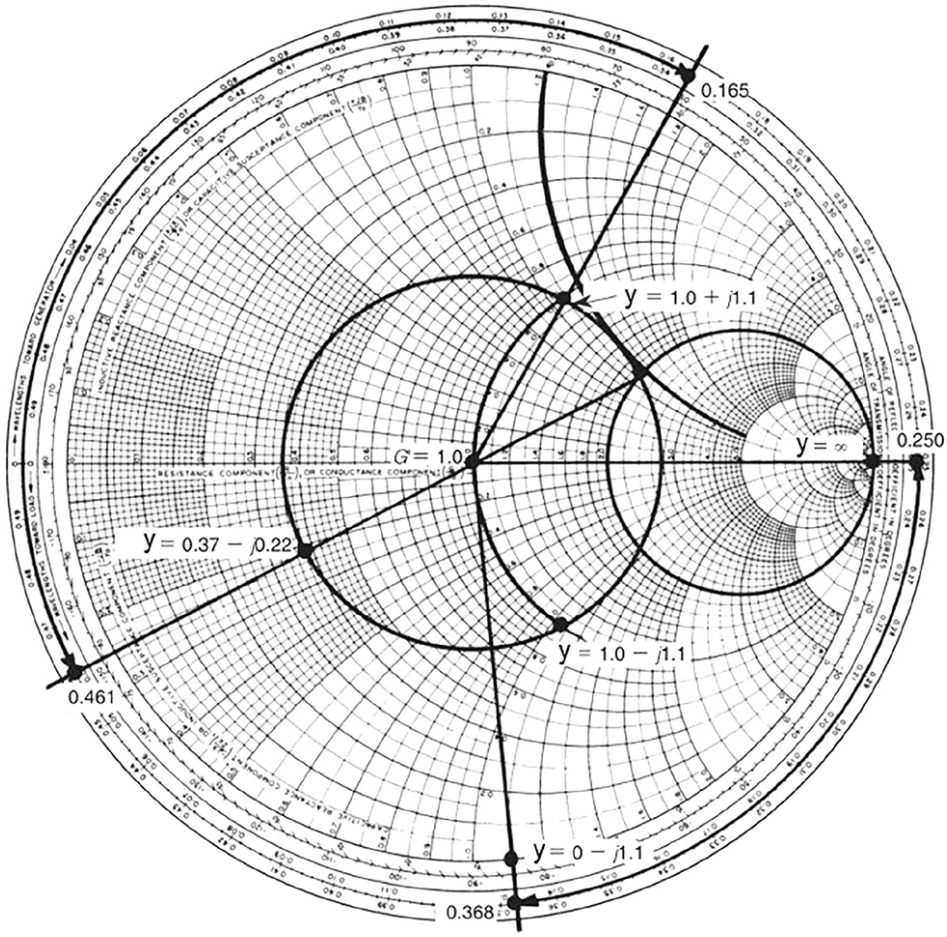
L_2

to match the load to the line at a given frequency.

Because the stub is connected in parallel with the line, it is easier to work with admittance parameters rather than impedances. We will rotate the load impedance by 180 degrees to convert it to an admittance. Once we have done this, all values on the Smith chart represent admittances. The pure resistance line becomes the pure conductance line and circles of constant resistance become circles of constant conductance.

Example 6.11 A parallel-wire transmission line is terminated by an antenna with an impedance
$Z_L = 100 + j60$
, as shown in Fig. 6.45. Find the closest location (to the load) and the length of a shorted stub that will reduce the VSWR on the transmission line to the left of the stub to 1.0:1.

Solution The load impedance is
$Z_{\mathrm{L}} = 100 + j60$
, which we normalize to
$z = 2.0 + j1.2$
. After plotting this impedance on the Smith chart in Fig. 6.46, we construct a VSWR circle and draw a diameter across it to find the load admittance,
$y = 0.37 - j0.22$
, corresponding to a shunt conductance
$g = 0.37$
in parallel with a shunt capacitance having susceptance
$b = -0.22$
.



F IGURE 6.46 Smith chart for Example 26.3 (stub matching).

The next step is to locate points where the VSWR circle passes through the unit conductance circle. At these points (there are two of them in every half-wavelength of transmission line), the transformed load impedance

y_L

has a conductance equal to that of the system transmission line along with a nonzero susceptance (i.e.,

$y_L = g_0 + jb$

. There are two such points shown in Fig. 6.46:

$1.0 + j1.1$

and

$1.0 - j1.1$

. If we choose one of these points and extend a line from the center (1.0,0.0) of the chart through the selected point to the outer circle

("WAVELENGTHS TOWARD GENERATOR"), we will be able to calculate the distance from the load where we should locate the stub.

If we choose the point

$1.0 + j1.1$

and extend a line through it from the center of the chart to the wavelength circle, we find the distance is 0.165*λ* from the left edge of the Smith chart. Next we extend the

$z \to y$

diameter we initially drew to the outer circle. Its path is from the chart center through the admittance point

$0.37 - j0.22$

to where the outer ring is marked 0.461*λ*. This point is

$0.500\lambda - 0.461\lambda = 0.039\lambda$

from the left edge of the Smith chart. The distance between these two points on the circle is the sum of 0.165*λ* and 0.039*λ*, or 0.204*λ*, so the stub location is

$L_1 = 0.204\ \lambda$

from the load.

The final step is to determine the length of the stub. This is done by finding two additional points on the Smith chart. We locate the point where admittance is infinite (far right side of the pure conductance line, where

$g = \infty$

). This corresponds to the short-circuited far end of the stub. We then locate the point where the admittance is

$0 - j1.1$

, which is the susceptance we need to place in parallel with the susceptance of the transformed load to cancel out the latter. Because the conductance component of this new point is 0, the point will lie on the

$-j1.1$

circle at the intersection with the outer circle. We draw lines from the center of the chart through each of these points to the outer circle. These lines intersect the outer circle at 0.368 and 0.250*λ*. The length of the stub is found from the shorter distance between them, so that

$L_2 = 0.368 - 0.250\ \lambda = 0.118\ \lambda$

.

Using the Smith chart we have matched a load of impedance

$Z_L = 100 + j60$

to a 50-Ω line by attaching a shorted stub of length 0.118 λ to the transmission line a distance 0.204 λ back from the load.

6.10.11 Incorporating Loss

Thus far, we have dealt with situations in which loss is either zero (i.e., ideal transmission lines) or so small as to be negligible. In the presence of appreciable loss in the circuit or transmission line, the VSWR circle becomes a spiral.

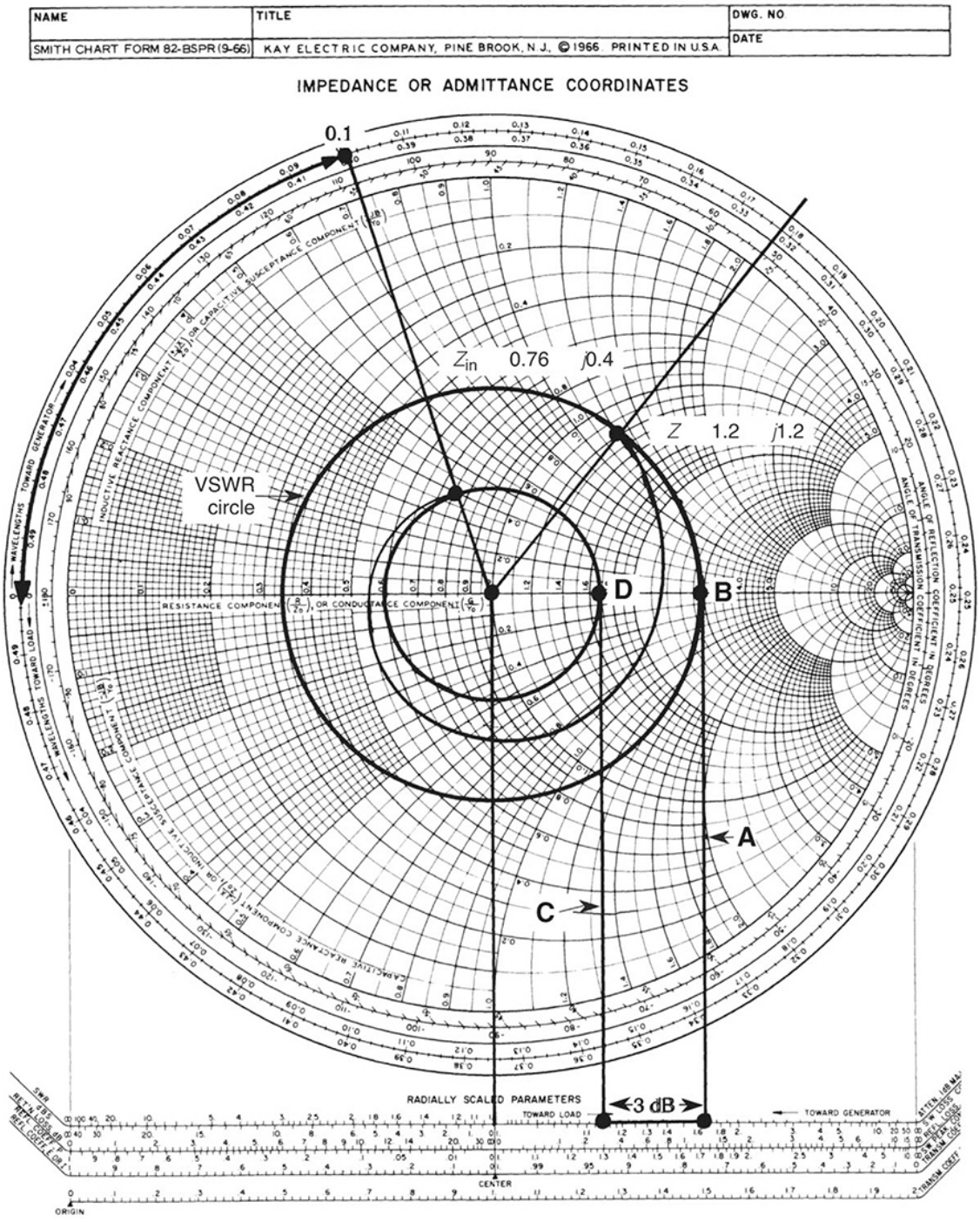
Figure 6.47 summarizes a typical situation directly on the Smith chart. Assume a lossy transmission line 0.60λ long, connected to a normalized load impedance of

$z_L = 1.2 + j1.2$

. An ideal (i.e., lossless) VSWR circle is constructed on the impedance radius represented by

$1.2 + j1.2$

. Next, a line (A) is dropped perpendicularly from the righthand point where this circle intersects the pure resistance baseline (B) to the "ATTEN, 1 dB/MAJ. DIV." line on the radially scaled parameters. A distance representing a total line loss (3 dB) is stepped off on this scale back toward the center. A second perpendicular line (C) is drawn, from the new 3-dB mark back up to the pure resistance line (D).



F IGURE 6.47 Smith chart analysis on a lossy transmission line.

The distance from the center of the Smith chart to point D, where line C intersects the pure resistance line becomes the radius

for a new circle that contains the actual input impedance of the line. The length of the line is 0.60λ, so we must step back

$$(0.60 - 0.50)\ \lambda$$

or 0.1λ as marked off on the "WAVELENGTHS TOWARD GENERATOR" outer circle. A radius is drawn from this point to the 1.0, 0.0 center point of the circular chart. The new radius intersects the smaller VSWR circle at

$$0.76 + j0.4$$

, which is the normalized input impedance (

$$z_{\text{in}}$$

) corresponding to an actual input impedance of

$$38 + j20\ \Omega$$

.

This example exaggerates the loss per unit length to clearly show the inward spiral of a lossy transmission line's VSWR curve. If a lossy line is long enough, the reflected wave is absorbed and the VSWR far from the load will be 1.0:1 regardless of the load impedance.

6.10.12 Frequency Response Plots

The Smith chart is a valuable tool for plotting impedances as a function of frequency. A complex passive network (such as a filter or an antenna) may contain resistors, inductors, and capacitors—either lumped or distributed. Both the resistive and the reactive components of the input impedance of such a network are generally functions of frequency, meaning that a point on the Smith chart representing an impedance typically moves along some arc or curve as the frequency is varied.

Figure 6.48 shows the impedance of four example antennas measured using a VNA. The general shapes of the curves are typical of practical monoband antennas. From the Smith chart representation of the input impedance, we can see how the VSWR and return loss vary for these antennas over frequency.

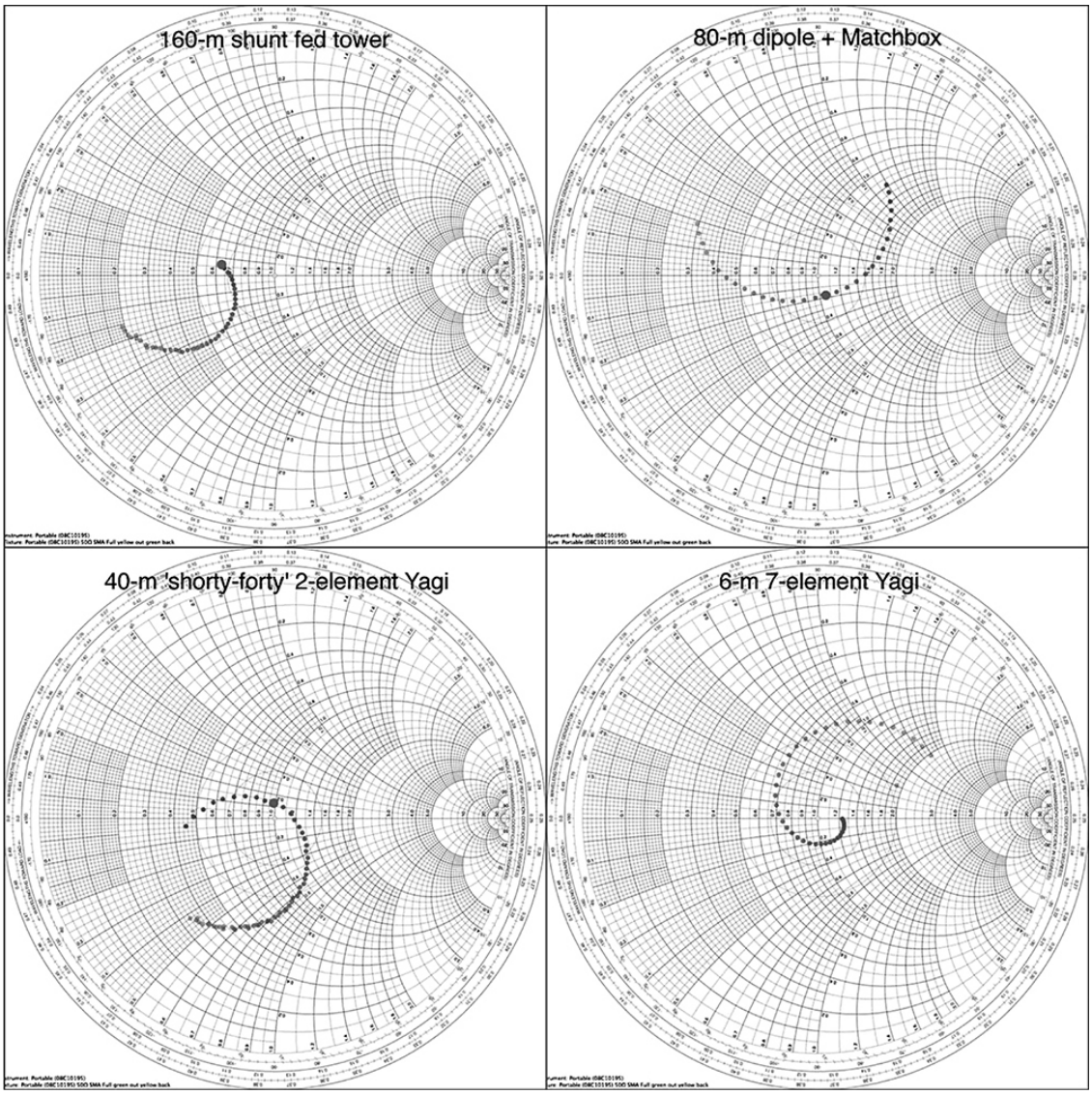


FIGURE 6.48 Antenna impedances displayed on the Smith chart.

6.11 S-parameters and Microwave Networks

Test instrument displays and component specifications for RF systems are often given in terms of S-parameters. An S-parameter is a generalization of the reflection coefficient in Eq. (6.15) to multiport circuits or devices, which are collectively called microwave networks.

An antenna is a one-port network. The transmission line stub with an open or short circuit load at one end used in impedance matching is a one-port network. A length of transmission line is a two port network. An amplifier or any other device with an input and output transmission line connection is a two-port network. Power splitters and combiners are three-port networks. Four-port networks include the quadrature hybrid or branch line coupler used to generate signals that are 90 degrees out of phase.

An *N*-port network has *N* connected transmission lines. When the network is excited with an external source at one or more ports, the sources introduce forward waves on the connected transmission lines. The network may have internal impedances, phase delays, gain, or any other type of electrical behavior. Based on the characteristics of the network, reverse waves are generated on each transmission line. The properties of the network determine the amplitudes and phases of the reverse waves.

Based on the relative amplitudes of the forward and reverse waves on the transmission lines at the network ports, the network can be characterized using S-parameters. S-parameters are defined to be the ratios of the forward and reverse waves on the transmission lines connected to each port:

$$S_{mn} = \frac{V_{\text{out},m}}{V_{\text{in},n}} \quad (6.64)$$

where
$V_{\text{out},m}$
is the amplitude of the wave coming out from port *m* and
$V_{\text{in},n}$
is the amplitude of the wave propagating into port *n*. Since the forward and reverse wave amplitudes are complex phasors that represent sinusoidal steady state voltages, S-parameters are complex and depend on the frequency.

For a two-port device like a filter, amplifier, or length of transmission line, if port 1 is considered to be the input and port 2 is the output, then

S_{11} is the input port reflection coefficient, S_{21} is the forward voltage gain, S_{12} is the reverse voltage gain, and S_{22} is the output port reflection coefficient.

We often represent S-parameters as magnitudes in dB. Most commonly, S_{11} and other S-parameters are converted to dB using

$$S_{11}\,(\mathrm{dB}) = 20\log_{10}|S_{11}| \quad (6.65)$$

A well-matched antenna reflects little of the incident power and S_{11} in dB is a negative number, ideally $-10\ \mathrm{dB}$ or smaller. An open, short, or purely reactive load reflects all incident power and the magnitude of S_{11} is 0 dB. Similarly, S_{21} is converted to dB using

$$S_{21}(\mathrm{dB}) = 20\log_{10}|S_{21}| \quad (6.66)$$

For an amplifier, S_{21} in dB is a positive number, and for a lossy device, it is negative.

Alternately, a negative sign can be introduced when converting S-parameters to dB. The return loss in Eq. (6.62) in terms of S-parameters is

$$\mathrm{RL} = -20\log_{10}|S_{11}| \quad (6.67)$$

Normally, a high return loss is desirable, as this means that most of the incident power is accepted by the network and little of the incident power is reflected. The insertion loss for a two-port network is

$$\mathrm{IL} = -20 \log_{10} |S_{21}| \quad (6.68)$$

A lossy length of transmission line will have an insertion loss that is positive in dB. For an amplifier, the magnitude of S_{12} is greater than one, so the insertion loss is negative and the device exhibits insertion gain.

S-parameters depend on the characteristic impedances of the transmission lines connected to the network ports, or the system reference impedance. Usually this is set at the standard value of 50 Ω, but occasionally other values are used. When the transmission lines have different characteristic impedances, generalized S-parameters can be used, but this is uncommon in everyday radio communications.

S-parameters can be arranged into a matrix. Properties of the network are represented in the S-parameter matrix. If the network is lossless, the S-parameter matrix is unitary. If the network is reciprocal, then the matrix is symmetric.

CHAPTER 7
Antenna Arrays

Multiple antennas or antenna elements fed from a common source of radio frequency (RF) power are collectively called an array. Similarly, multiple antennas feeding a single receiver input also form an array. The signal radiated to a distant point by an array of transmitting antennas is a combination of the signals from all the elements. Because the field radiated by each element travels a slightly different distance, the field radiated by the array in different directions is influenced by constructive addition or interference due to phase differences. For actively controlled phased arrays and electronically scanned arrays, the direction of peak radiation intensity can be controlled by shifting the phase of the drive signals to each element in the array.

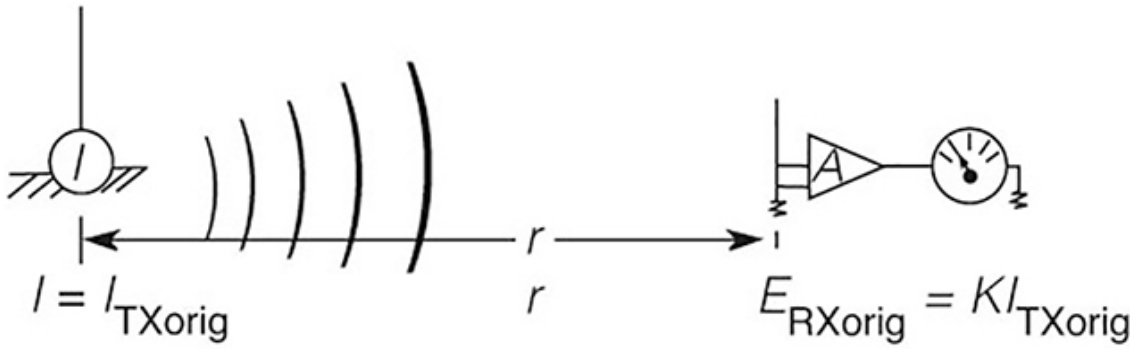
We will develop the basic concepts of arrays for a pair of simple dipoles. The concepts of array beamforming will be used to understand parasitic arrays like the Yagi and ground effects for antennas over ideal and lossy ground planes. The focus of this chapter is arrays for frequencies well below 1 GHz, but the concepts apply to arrays used for cellular systems and other applications at higher frequencies, where the arrays are physically smaller due to the shorter wavelengths but tend to employ more elements in the array.

7.1 Array Basics

We will start with a single antenna system, as shown in Fig. 7.1. The field at a point far from the radiating antenna is reduced in amplitude and delayed in phase due to the distance the field propagates from the transmitting antenna to the receiver. We will call the field radiated by the single antenna at the location of the receiver

E_{RXorig}

. The driving current into the feed terminals of the transmitting antenna is
I_{TXorig}
.



F IGURE 7.1 Received signal at a distant point from a single transmitting antenna.

Suppose, we now excite a second transmitting antenna located $\lambda/2$ from the first antenna (Fig. 7.2). To keep the total transmitted power the same, we will excite both antennas with feedpoint current $I_{\mathrm{TXnew}} = I_{\mathrm{TXorig}}/\sqrt{2}$. The principles of superposition and linearity tell us that the net electric field at the receiver will be the sum of the field arriving from each of the transmitting antennas, so that

$$E_{\mathrm{RX}} = \frac{E_{\mathrm{RXorig}}}{\sqrt{2}} + \frac{E_{\mathrm{RXorig}}}{\sqrt{2}} = \sqrt{2}E_{\mathrm{RXorig}} \quad (7.1)$$

Since power goes as the square of electric field, the power at the receiver has doubled.

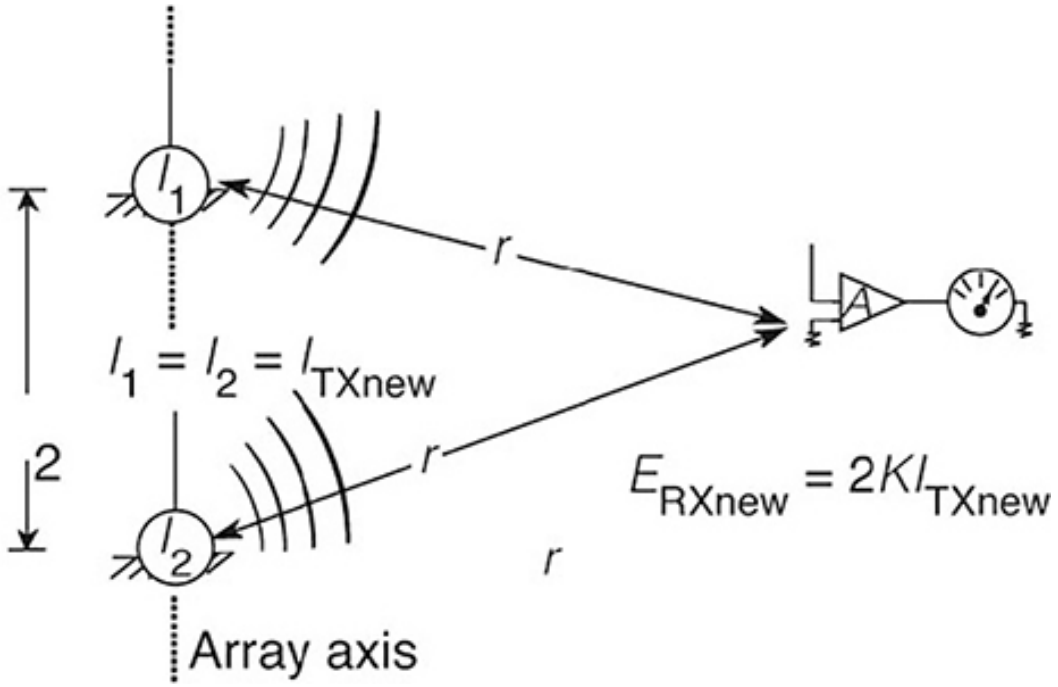


FIGURE 7.2 Received signal at a distant point from two identical transmitting antennas.

What we have created is an antenna array or more specifically a two-element all-driven phased array operated in its broadside mode. In broadside mode, the element feed currents are in phase and the direction(s) of maximum radiation are on a line at right angles (hence, broadside) to an imaginary line—called the array axis, and indicated by the vertical dotted line in Fig. 7.2—connecting the elements. Through the simple expedient of erecting a second antenna and splitting our allowed transmitter power equally between the two antennas, we have accomplished an increase in received signal strength at a distant point—an increase that would have required us to double our transmitter power (from, say, 100–200 W) if we had continued to use the original antenna.

Most antenna arrays increase the radiated power in one direction at the expense of reducing the field strength in other directions. It may be helpful to visualize the radiation field of the original antenna as a spherical balloon, and a given transmit power corresponds to blowing the balloon up to a certain size. Now wrap your two hands around the middle of the balloon and squeeze. Pushing in on part of

it causes the skin of another part of the balloon to extend, even though the total air within the balloon remains constant. Similarly, an array of antenna elements modifies the radiation field of the original antenna, making it stronger in certain directions while weakening it in others.

Let's see why this happens physically. The two transmitting verticals of Fig. 7.2 are spaced

$\lambda/2$

apart, but their feed currents are in phase. When the field from antenna A reaches antenna B, it will have traveled a half-wavelength, and, hence, its waveform will be 180 degrees out of phase from its starting point at A and also 180 degrees out of phase with the waveform of antenna B's field. As the fields from antennas A and B propagate outward along an extension of the line connecting A and B, they are essentially the same strength at every point on the line, but 180 degrees out of phase. That is to say, they almost completely cancel at all receiving points on a line connecting the centers of the two antennas. The received signal strength from this two-element phased array is, for all intents and purposes, zero anywhere along the axis of the array. The array has "stolen" power from certain directions to increase the radiated power in other directions.

If we move our receiving antenna around the compass at a constant distance from the center of the array, we find that the received signal varies from zero on the array axis to a maximum broadside to the array, as shown in Fig. 7.3. We say that this array has a null in the end-fire direction along the array axis where the radiated field is zero. The main beam of the array is in the broadside direction.

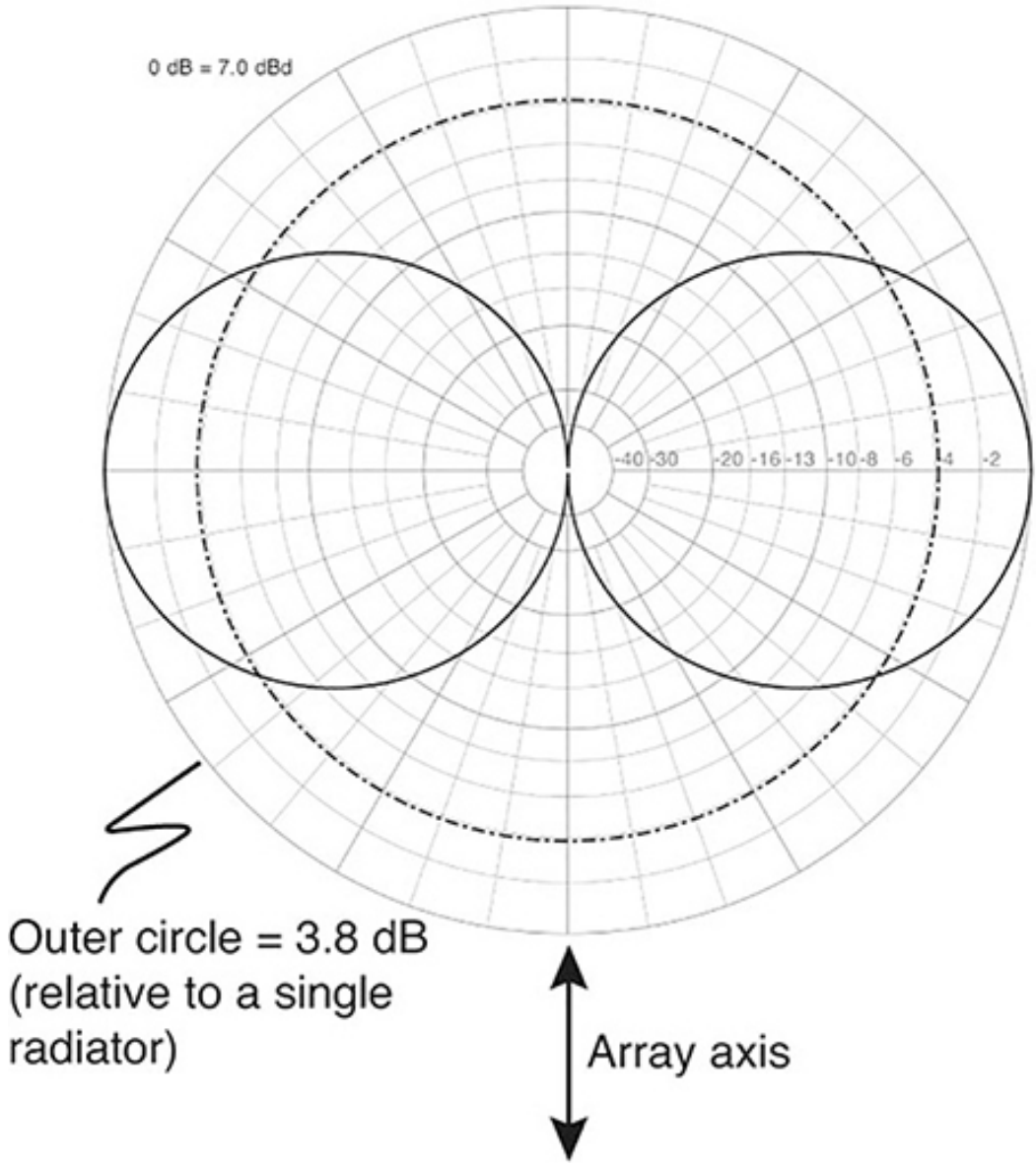


FIGURE 7.3 Array pattern for the two-element array of (B) in broadside (fed in phase) mode.

For this simple array example, we calculated a gain of 3 dB over a single element, but that is only a rough approximation because we ignored possible interactions between the two elements. If the two elements are

$\lambda/2$

apart, on-axis radiation is everywhere zero, but interaction between the two elements alters their feedpoint impedances and the resulting

gain in the desired (broadside) direction is not exactly 3 dB greater than that of a single radiating element. In fact, for $\lambda/4$ verticals spaced $\lambda/2$ apart, it's a bit higher than 3 dB; as shown in Fig. 7.3, it's 3.8 dB. For other spacings, such as $3\lambda/2$, $5\lambda/2$, etc., and for other antenna orientations (dipoles end to end, for instance) the mutual impedances of the elements can cause the gain of the array to be either greater than or less than 3 dB for a fixed input power to the array. For other than $\lambda/2$ spacing, additional sidelobes form. As we shall see in Chap. 14, the interactions between closely spaced conductors allow us to obtain gain from multielement antennas having only one element directly connected to a feedline.

If the RF signal being applied to one of the two radiating verticals in Fig. 7.2 is shifted 180 degrees so that the two antennas are driven out of phase, the pattern of Fig. 7.4 results. The pattern is rotated 90 degrees and is somewhat flatter than the broadside pattern. The main beam is now in the end-fire direction and the null is in the broadside direction. The peak gain of this array's end-fire mode compared to a single vertical is 2.3 dB—about 1.5 dB less than that of the broadside mode.

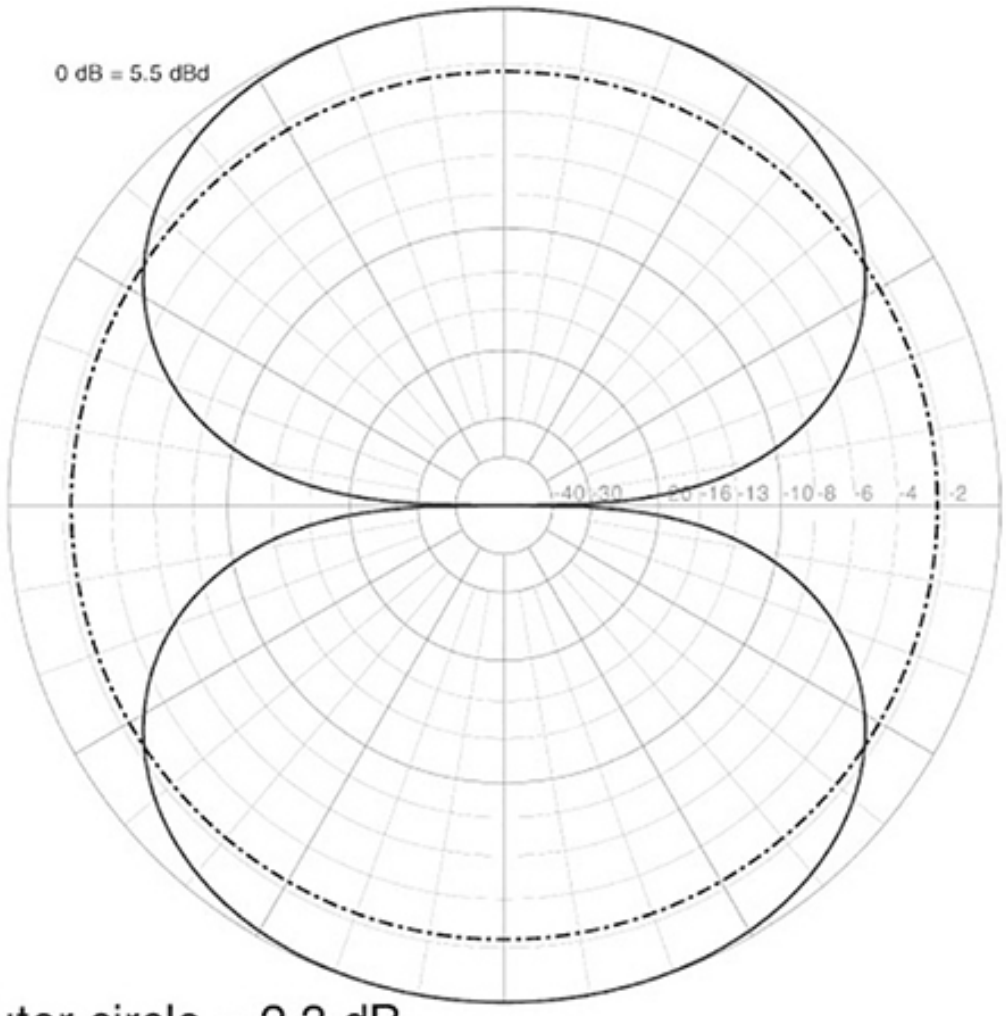


FIGURE 7.4 Array pattern for the two-element array of (B) in end-fire mode (fed 180 degrees out of phase).

7.2 Array Factor

We can build arrays from just about any kind of antenna element. To simplify array analysis, we can approximate the radiation pattern of the array by the product of the array factor, which accounts for the geometry of the array and the phases and amplitudes of the element excitations, and the pattern of one of the elements in the array.

From a purely mathematical viewpoint, the easiest arrays to analyze are those comprised of isotropic radiators because the individual elements have a very simple pattern—they radiate equally well in all directions. When two or more isotropic radiators are combined to form an array, the resulting pattern is the array factor itself. The array factor is obtained from evaluation of the geometry of the element positions, as well as the relative amplitudes and phases of the individual element feedpoint currents. For the two-element broadside array discussed here and shown in Fig. 7.2, the shape of the azimuthal array factor is given by the equation

$$\mathrm{AP} = \cos\left(\frac{\pi}{2}\cos\theta\right) \quad (7.2)$$

where θ is the angle between the array axis and a line drawn from the center of the array to a distant receiving antenna.

But isotropic radiators, as we saw in Chap. 5, are fictitious, so we are forced to form arrays out of real elements: dipoles, verticals, loops, and other basic antenna types. When we do that, the resulting radiation pattern developed by the array is the array pattern multiplied by the radiation pattern of an individual element. If we replace the

$\lambda/4$

monopoles of Fig. 7.2 with

$\lambda/2$

dipoles, we still have a two-element broadside driven array with

$\lambda/2$

element spacing. The resulting radiation pattern will be the array pattern of Fig. 7.3 multiplied by the inherent radiation pattern of a conventional dipole (Fig. 5.11).

The orientation of the dipole elements relative to the orientation of the array itself will make a big difference in the overall pattern. A rule of thumb is to use elements in such a way that their natural direction of maximum radiation coincides with at least one of the desired directions of maximum radiation for the array as a whole. Failure to do so will usually lead to arrays with patterns that are sharp or

multilobed (often in the wrong directions) and feedpoint impedances that are unstable and difficult to match.

Consider the broadside array of Fig. 7.5, where we have replaced the verticals with

$\lambda/2$

horizontal dipoles laid end to end with centers

$\lambda/2$

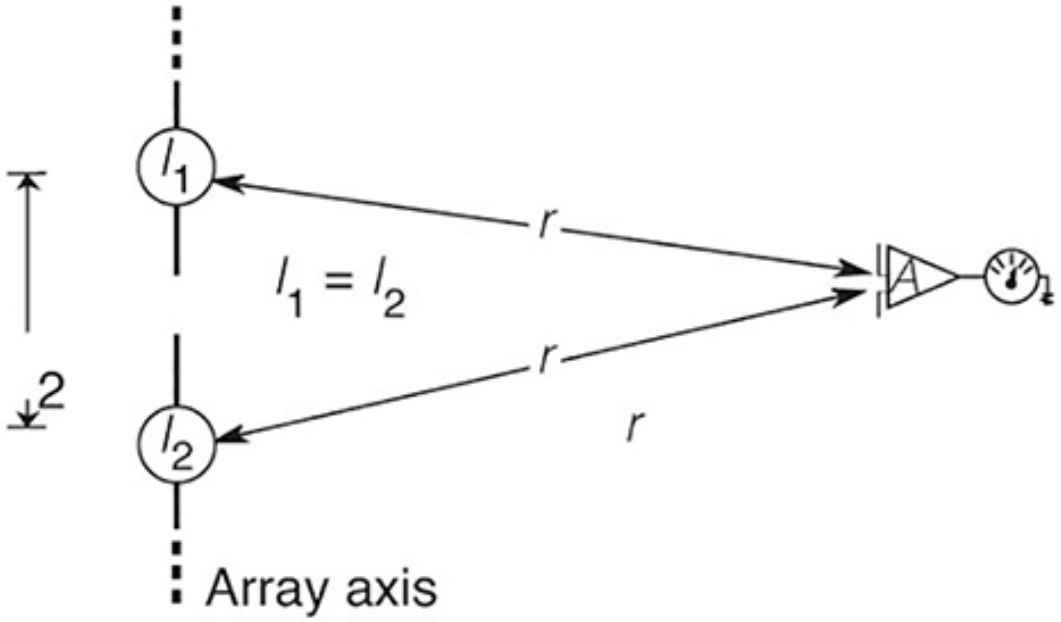
apart. The pattern for this array is graphed in Fig. 7.6. As a result of the added directionality of the

$\lambda/2$

dipoles, the directivity of the main lobe is "sharper" or narrower than the main lobe of Fig. 7.3. The peak gain of the main lobe relative to a single

$\lambda/2$

dipole is only 1.5 dB because the mutual impedances between the two dipoles cause the feedpoint current in each dipole to be less than it would be if the two dipoles were far enough apart to have negligible interaction.



F IGURE 7.5 Two-element broadside array of $\lambda/2$ dipoles spaced $\lambda/2$ apart on their axis.

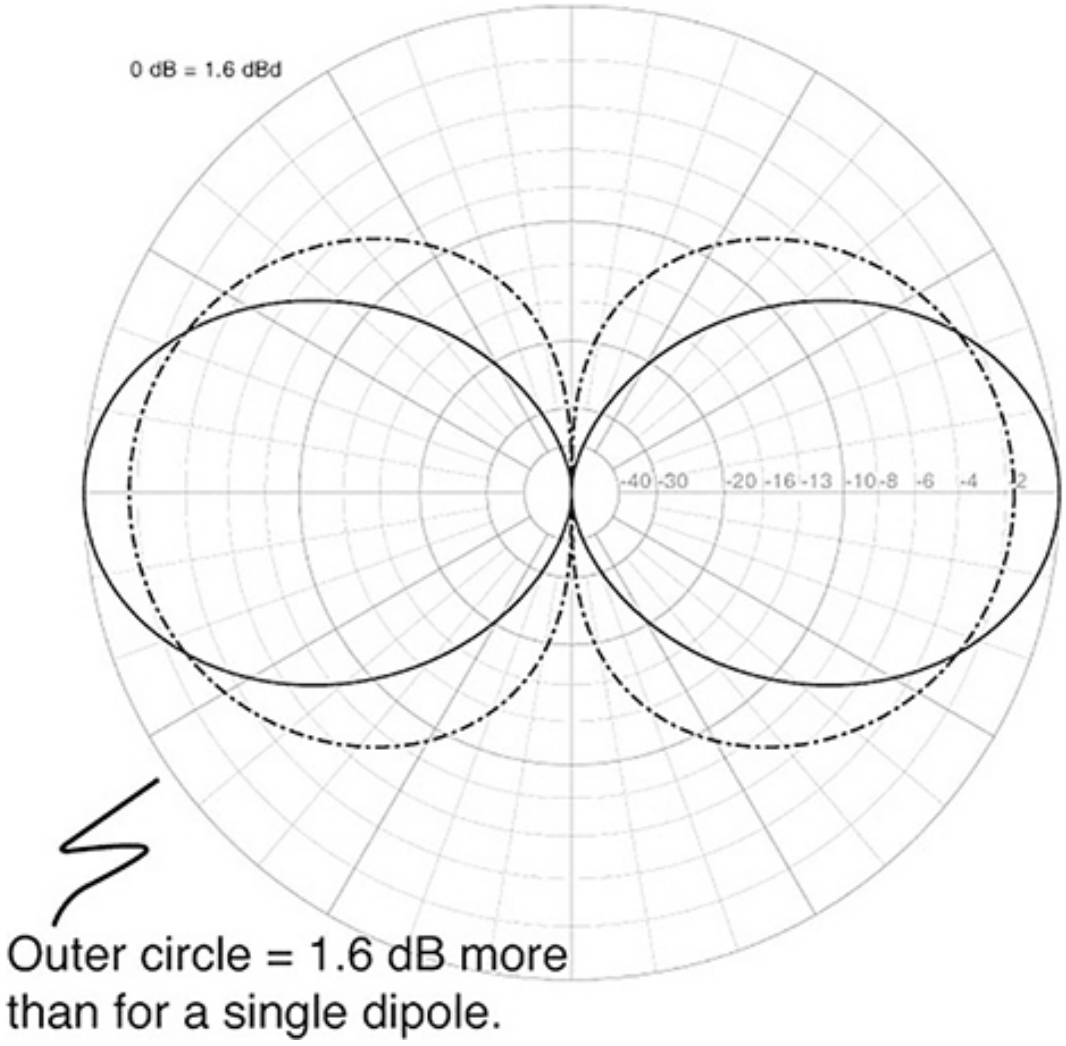


FIGURE 7.6 Overall pattern for two-element broadside array in (A).

If we place one of the dipoles $\lambda/2$ above the other, as depicted in Fig. 7.7, each dipole "sees" the other dipole differently, and their mutual impedances are different from those of Fig. 7.5 as a result. As reported in Fig. 7.8, the overall gain of this two-element broadside array is 3.8 dB more than that of a single dipole. If one is going to use two driven dipoles to maximize forward gain in a broadside array, the configuration of Fig. 7.7 is preferable to that of Fig. 7.5. In either case, the dipoles are oriented so that the direction of maximum radiation for the basic dipole

element coincides with the desired direction of maximum radiation for the array as a whole when operated in its broadside mode. Along the axis of the array, both the array factor and the element pattern bring the radiated field strength to zero. The overall pattern of Fig. 7.8 will be of particular interest to us later when we discuss the effect of ground on a dipole's pattern—especially when we vary the height of the dipole.

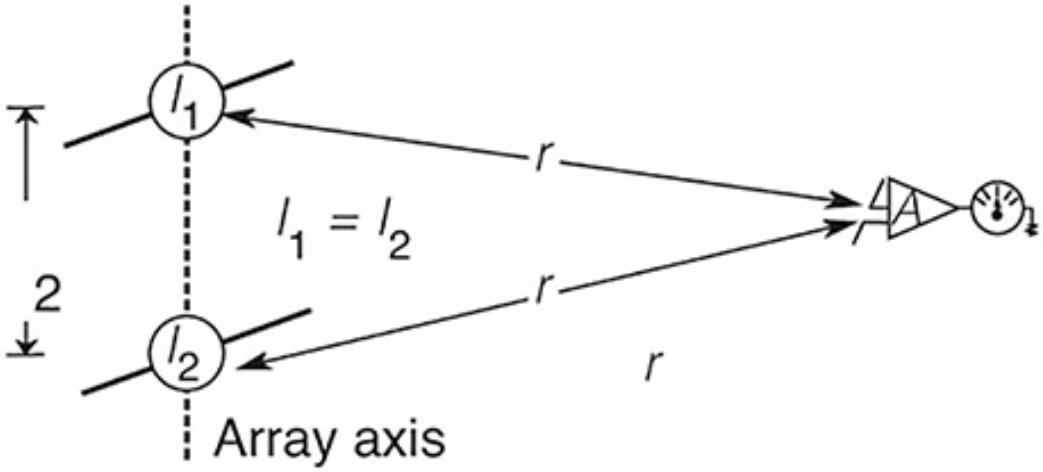
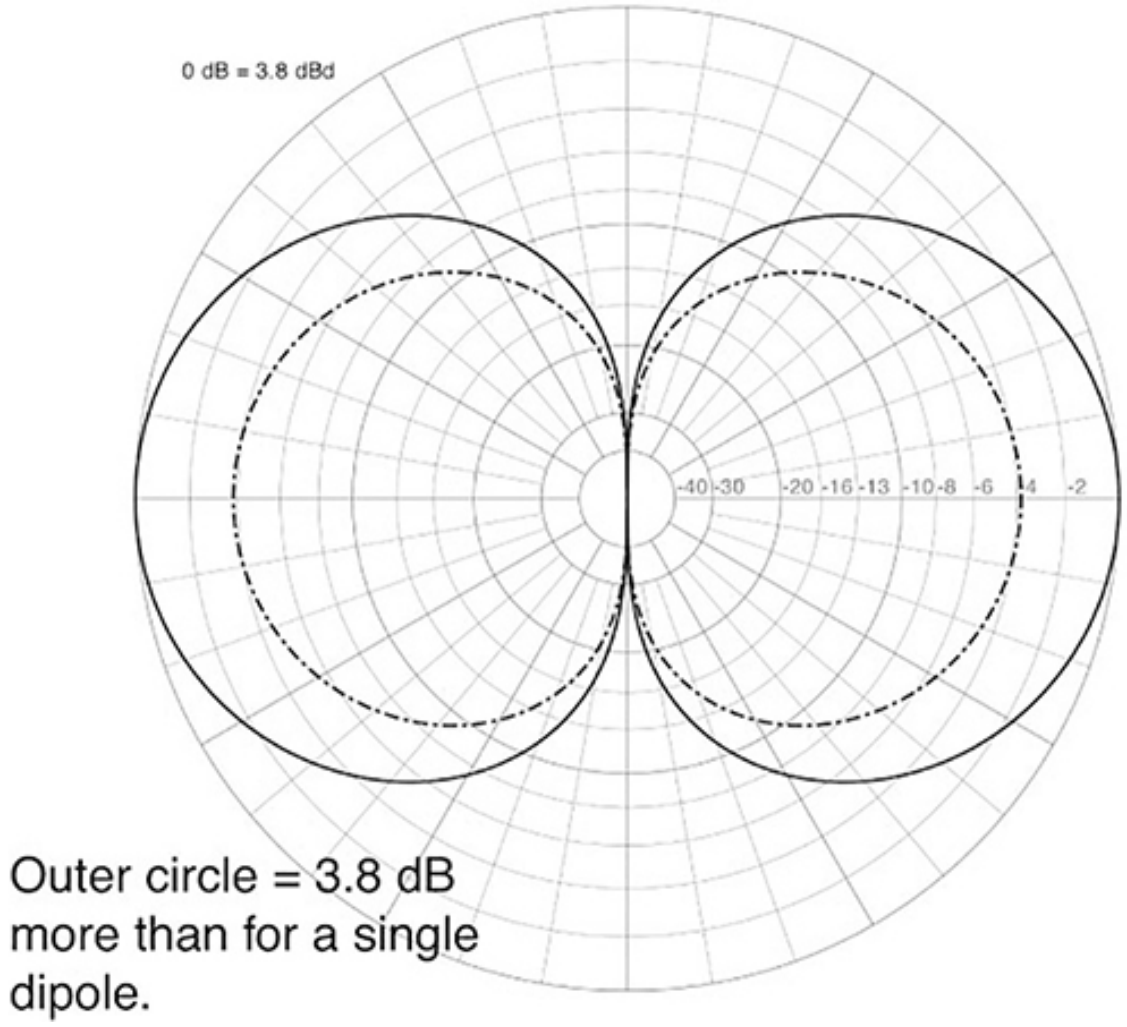


FIGURE 7.7 Two-element broadside array of dipoles spaced $\lambda/2$ at right angles to the array axis.



F IGURE 7.8 Overall pattern for the two-element broadside array in (C).

Another form of array operation is the end-fire mode. Here the objective is to maximize radiation along the axis connecting the elements. One way to do this is by spacing the elements $\lambda/2$ apart and feeding them 180 degrees out of phase. By the time the out-of-phase field from element A travels along the axis to element B it will have shifted another 180 degrees and be back in phase with the field from B, thus effectively doubling the radiated field along the axis relative to a single element. For two identical radiators spaced $\lambda/2$

apart and fed with out-of-phase excitations of equal amplitude, the shape of the array pattern is given by

$$\mathrm{AP} = \sin\left(\frac{\pi}{2}\cos\theta\right) \quad (7.3)$$

From our earlier discussion about attempting to align the inherent pattern of the elements of the array with the main lobe of the array pattern, we conclude that horizontal dipoles oriented parallel to the array axis are not the best choice for elements in this end-fire array. If we feed one dipole in the array of Fig. 7.9 with a drive current that is out of phase with the current in the other element, we obtain the azimuthal pattern of Fig. 7.10, which is normally undesirable. Along the array axis, the dipole patterns radiate very little horizontally polarized energy to the far field. At right angles to the array axis, the element spacing and drive current phase relationships conspire to create a similar null. The radiated energy is primarily at 45 degrees to either axis, and the pattern resembles a four-leaf clover.

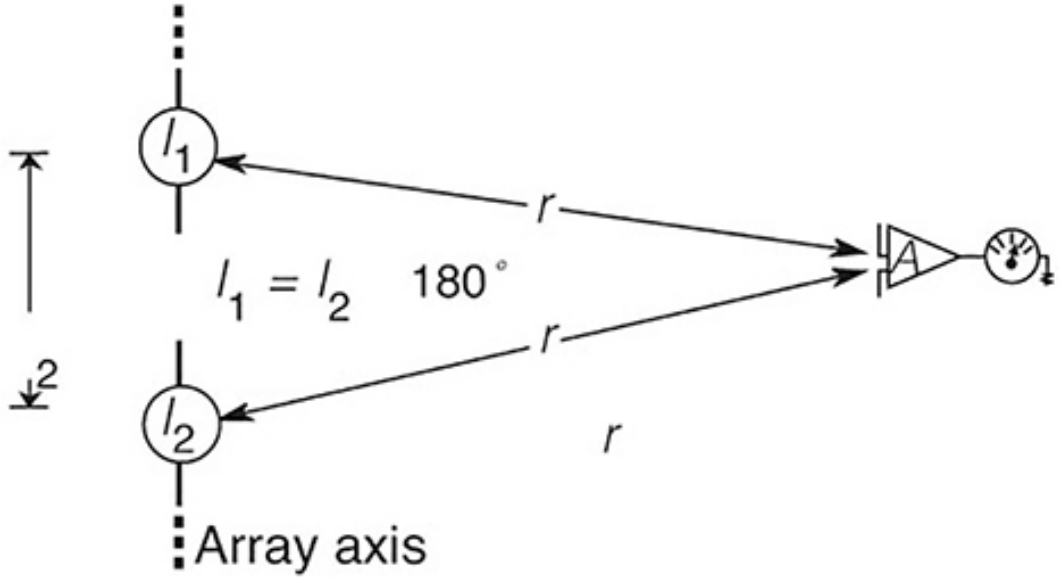


FIGURE 7.9 Two-element end-fire array of collinear $\lambda/2$ dipoles spaced $\lambda/2$ apart on the array axis.

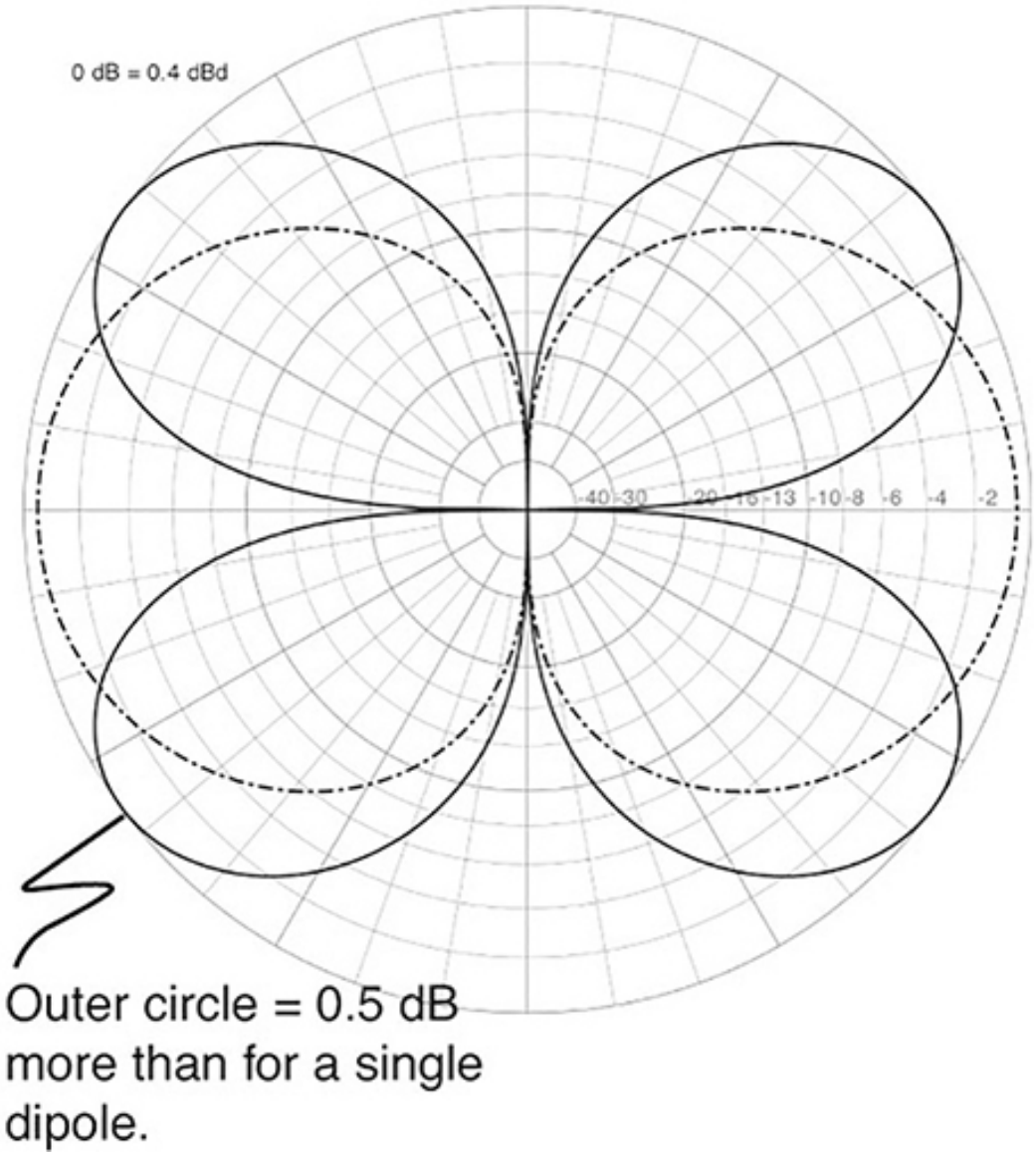
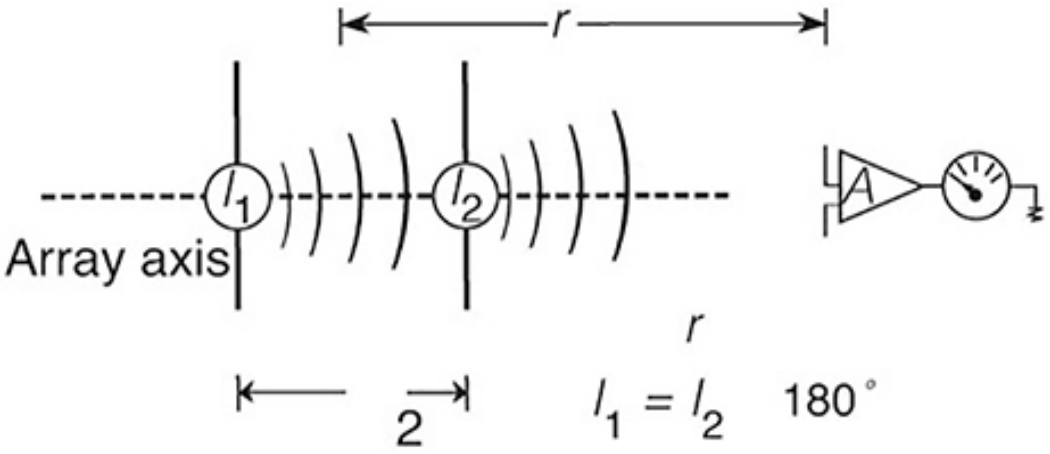


FIGURE 7.10 Overall pattern for the two-element end-fire array with collinear dipoles.

A better design is to rotate the horizontal dipoles by 90 degrees, as shown in Fig. 7.11, so that the resulting overall pattern of Fig. 7.12 has main lobes in the horizontal direction. The overall gain relative to a single antenna is a bit less than 3 dB because of the effect of the mutual impedances between the elements for this spacing and relative element orientation.



F IGURE 7.11 Two-element end-fire array of parallel dipoles spaced
$\lambda/2$
on the array axis.

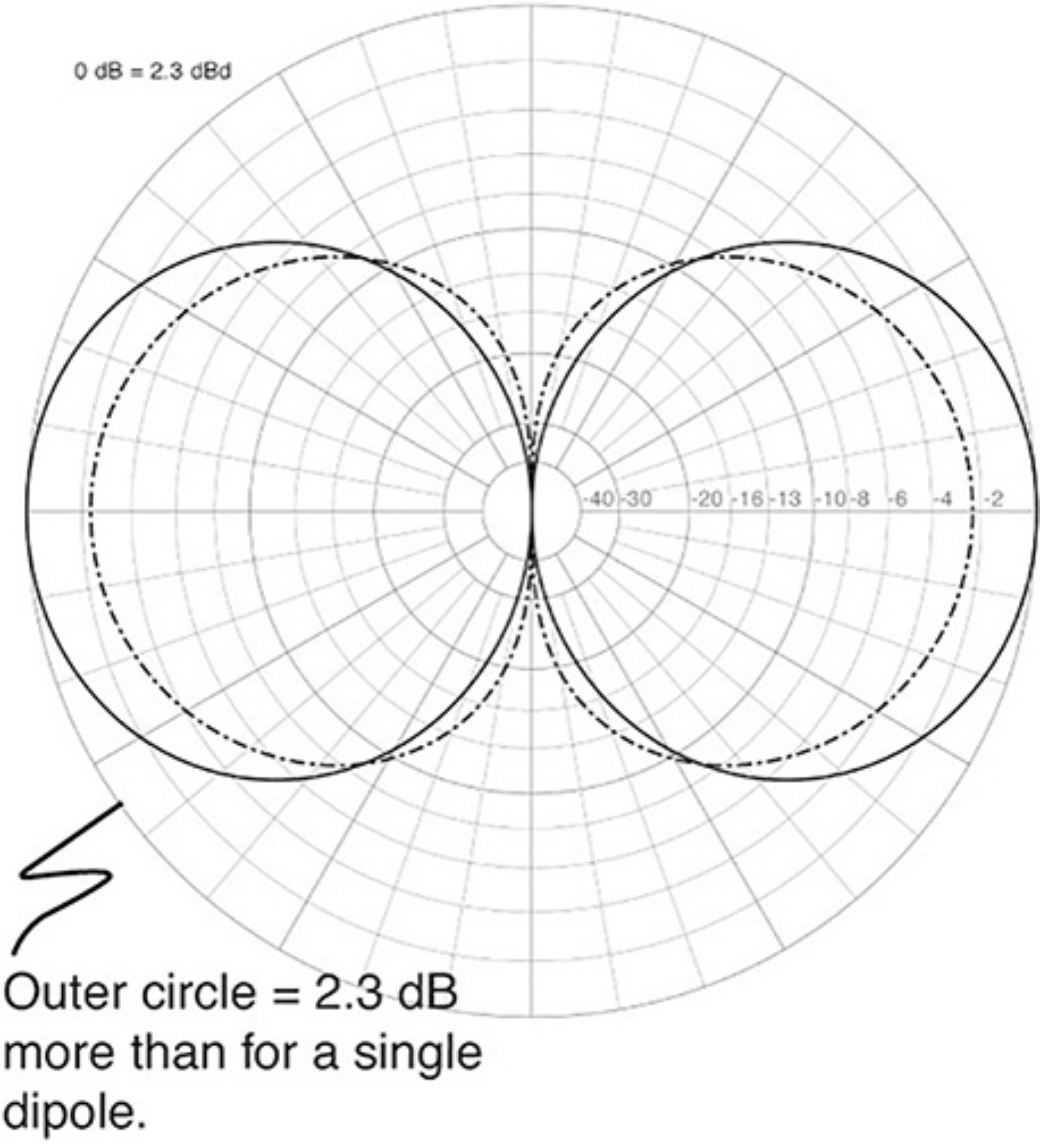


FIGURE 7.12 Overall pattern for the two-element end-fire array with parallel dipoles.

To switch a single fixed physical array between broadside and end-fire patterns, the ideal element may well be a vertical dipole or monopole, since those antenna types have omnidirectional azimuthal radiation patterns and can be used in either broadside or end-fire applications without having to be physically rotated or relocated when changing the preferred direction of maximum gain.

The driving currents for the elements may be different. AM broadcast stations have employed multielement arrays with unequal

feed currents. One common configuration uses a binomial current distribution in the vertical elements. (A five-element array might have relative element currents of 1-4-6-4-1, for instance.) Combined with very close element spacings, these arrays are capable of producing extremely high gain patterns and deep nulls or sidelobes. The latter are important in the AM broadcast band to prevent co-channel and adjacent channel interference to "protected" stations on the same frequency in other parts of the country.

Another design is the collinear array. Multiple copies of an element type are laid (or stacked, if vertical) end to end and electrically connected "heel to toe" through phasing sections. The feed is designed to force the currents in all elements to be in phase even though only one element is driven by a signal from the transmitter. Many VHF and UHF verticals and mobile whips are collinears. The short lengths of antennas for those frequencies make stacking practical from a mechanical standpoint. For a vertical, collinear stacking does not alter the shape of the azimuthal pattern. The usual objective is to sharpen the elevation pattern so that more gain in all azimuthal directions is available at low elevation angles at the expense of (usually useless) high-angle radiation. At VHF and UHF, collinear antennas are often commercially sold as a complete array in a single assembly.

As we will discuss in Chap. 10, when a half-wave dipole is operated substantially above its design frequency, f_C, it becomes a pair of collinear arrays because the wire on each side of the center insulator is now longer than $\lambda/4$. As the operating frequency rises above f_C, the maximum broadside gain of the antenna grows relative to its value at f_C because the additional wire length in the radiating element possesses a greater length × current product. Eventually, however, current reversal in adjacent half-wave segments begins to reduce the

maximum broadside gain of the array, which occurs when the total length of the wire is 5⁄4 λ, corresponding to an operating frequency near 9 MHz for an 80-m dipole. Above that frequency, the maximum broadside gain starts to decrease with frequency, and eventually the array pattern breaks up into multiple smaller lobes. Thus, the broadside gain of an 80-m

$\lambda/2$

dipole operated on 40 or 30 m is greater than the 80-m gain; on 40 m the increase is on the order of 1.5 dB. The increased gain is accomplished through a narrowing (or "sharpening") of the main radiation lobe. On 30 m, the gain of the main lobe is still about 0.5 dB greater than on 80, but it is accompanied by two additional lobes exhibiting gain equal to the main lobe but somewhat broader at their 3-dB points. A horizontal wire antenna 5/4 λ long on any band is known as an extended double Zepp (EDZ) (or, alternatively, double extended Zepp) for that band. See Chap. 8 for further discussion of the EDZ.

If one side of the dipole antenna is thrown away and the remaining side oriented vertically and fed against a ground plane as a vertical monopole, the length for maximum broadside gain is one-half of 5⁄4 λ, or 5⁄8 λ; hence, the popularity of VHF and UHF whips of that length.

The loop antenna, which we will discuss in later chapters, can be thought of an array. To see this, consider the bent dipole. This is nothing more than an ordinary

$\lambda/2$

dipole with the outer half of each side bent straight down. If a second such bent dipole is turned upside down, attached at its ends to the corresponding ends of the first dipole, and one of the two feedpoints shorted out, we have what everyone calls a loop antenna. (See Fig. 7.13.) Because of the "automatic" phase reversal in adjacent

$\lambda/2$

sections of a longwire, the currents at the centers of the two dipoles are always in phase and we have a two-element broadside array, with the elements spaced

$\lambda/2$
apart. The loop can be suspended so that it lies in a horizontal plane, a vertical plane, or some plane partway in between. If suspended vertically and only a single high support is available, the loop takes a triangular form and is called a delta loop. If multiple loops are attached to a common boom, the resulting "array of arrays" is called a quad. Regardless of its name, however, each loop is essentially two half-waves in phase, and the direction of maximum radiation for that array is broadside to the plane formed by its component dipoles.

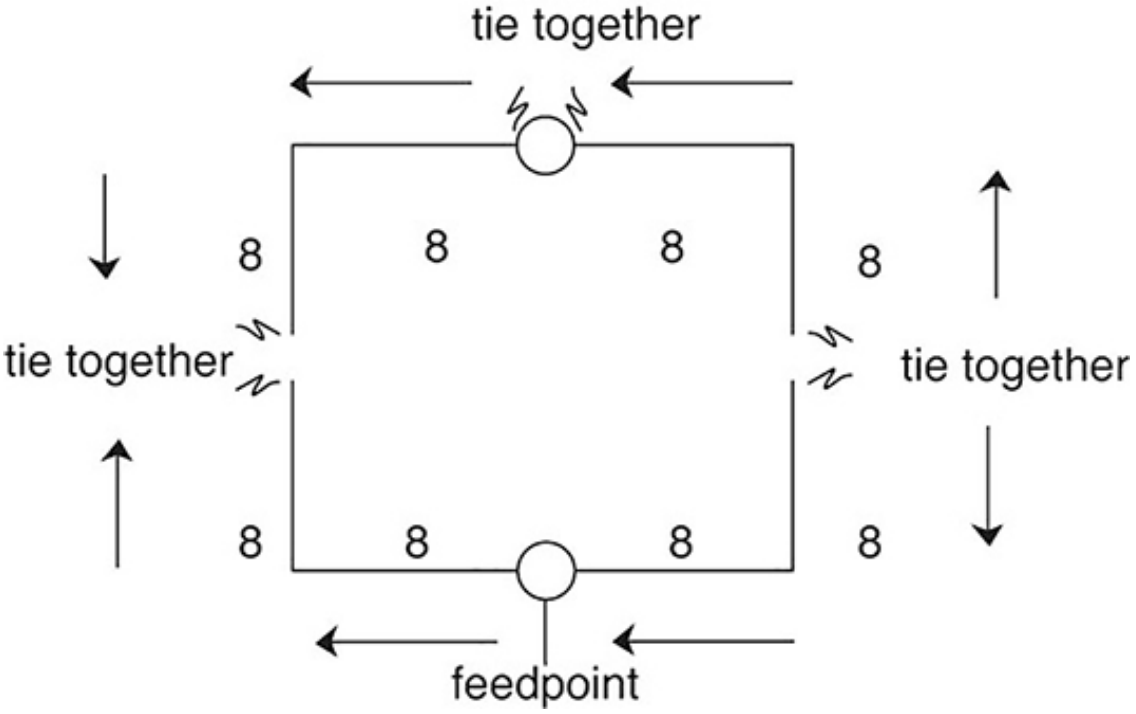


FIGURE 7.13 The full-wave loop array of two bent dipoles. Arrows show direction of current everywhere at a given instant.

Further, there is nothing magic about feeding the loop at the center of the bottom dipole. As long as the entire wire path around the loop is one wavelength, the loop can be broken and fed anywhere on that path. If a vertically suspended loop is fed in the middle of a side, it becomes predominantly vertically polarized and may well exhibit better ability to work or hear long-haul DX. If a horizontally suspended loop is fed from the "next" side, the azimuthal orientation of maximum performance will rotate 90 degrees.

Unfortunately for the HF loop, low dipoles and arrays of low dipoles are usually adversely affected by the ground beneath them. The effect of the ground is to create an image array below the antenna. Analysis of the operation of a loop antenna or array must take that into account. (Array factors multiply each other.) Ground reflection effects are discussed later in this chapter. Chapter 15 goes into loops and quads in more detail.

Thus far, we have limited the discussion of arrays to those with feedpoint currents that are either in phase or out of phase. To accomplish various goals for main beam steering and sidelobe control, any phase relationship between feedpoint currents can be used.

An array that employs quadrature phasing of element currents is the two-element cardioid array (Fig. 7.14). Two identical omnidirectional radiators (vertical monopoles, for instance) spaced $\lambda/4$ apart are fed with currents of identical magnitude phased 90 degrees apart. Visualizing how the array pattern is formed is not that much more difficult than for the broadside and in-line arrays. Suppose the current in element 1 lags the current in element 2 by 90 degrees. By the time the radiated field from element 1 reaches 2, it has accumulated an additional lag of 90 degrees due to the distance it had to cover. Thus, radiation along the axis heading in the direction of 1 to 2 is the algebraic sum (from 2 outward) of two equal fields that are 180 degrees out of phase, and the radiation field is canceled in that direction.

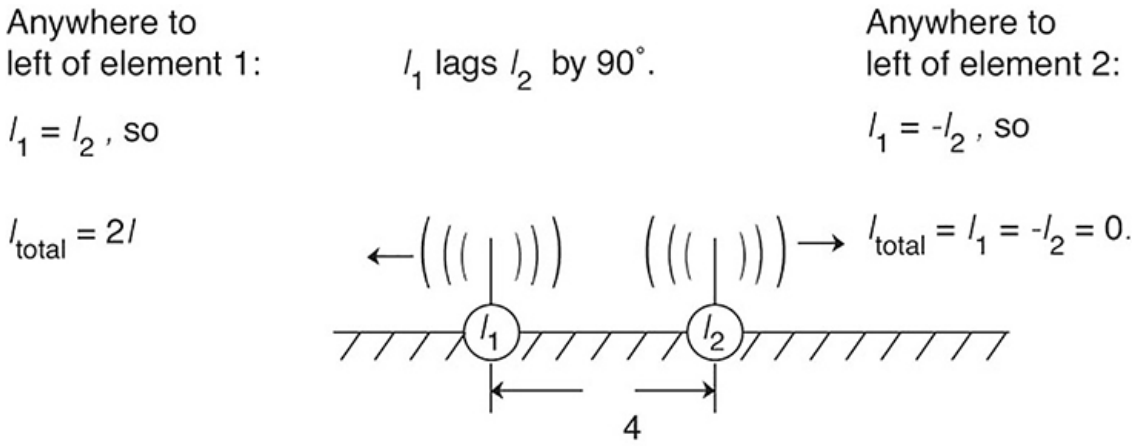


FIGURE 7.14 Two-element quadrature or cardioid array.

In the opposite direction along the axis of the two elements, heading from 2 to 1 and beyond, the radiation from the feedpoint current in 2 (which leads the feedpoint current in A by 90 degrees) loses the 90-degree lead as it covers the distance from 2 to 1. Thus, from 1 outward, radiation fields from 1 and 2 are perfectly in phase, and reinforced.

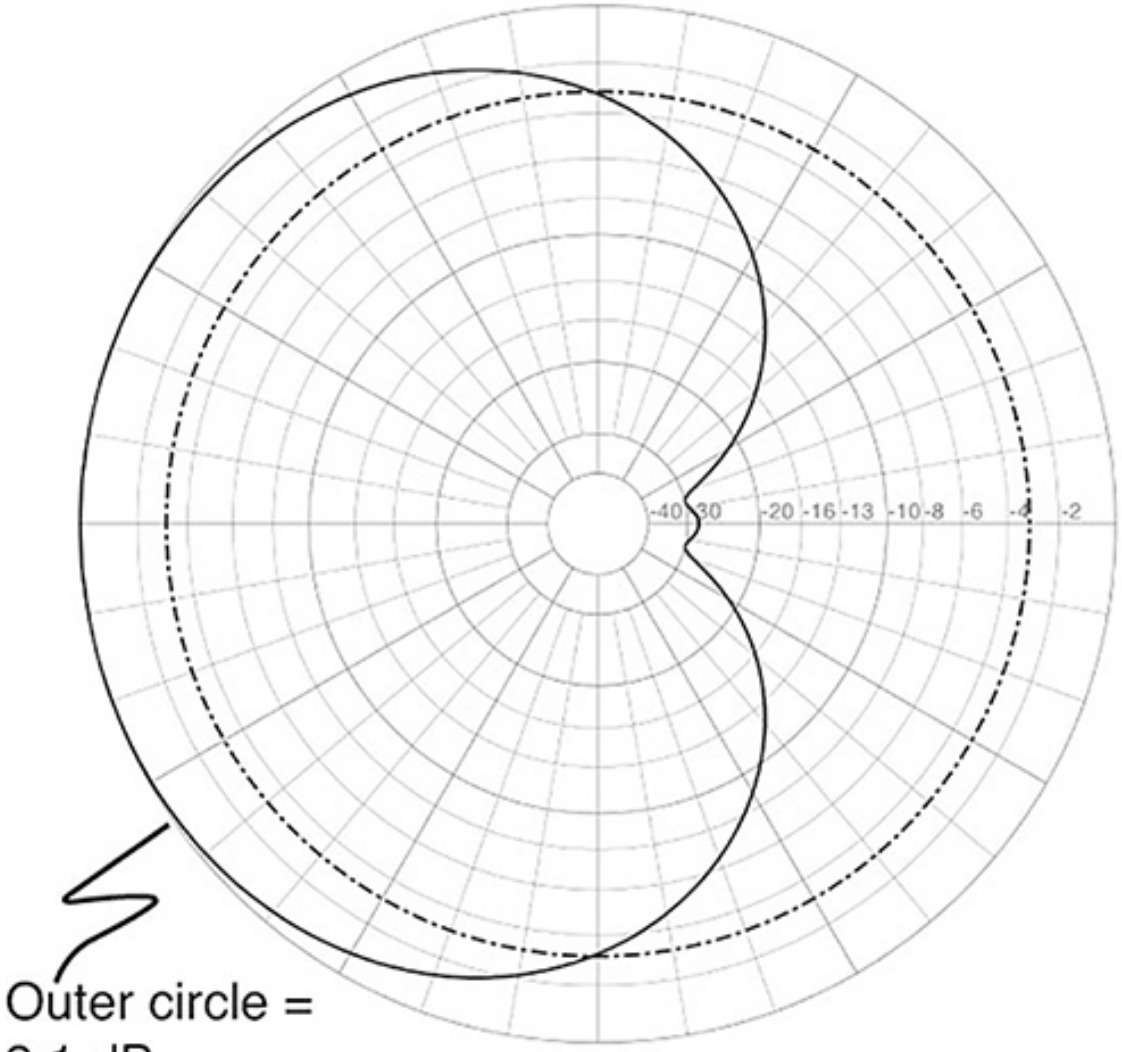
Between these two extremes, the amplitude of the radiation pattern at a distant receiving site is a smoothly changing value from maximum to minimum signal strength, symmetrically about each half-circle. The resulting equation for far-field signal strength as a function of azimuth angle is proportional to

$$1 + \cos\ \theta$$

and has the shape shown in Fig. 7.15. For a pair of

$$\lambda/4$$

vertical monopoles over perfect ground, the array has a maximum gain of 3.1 dB relative to a single identical monopole.



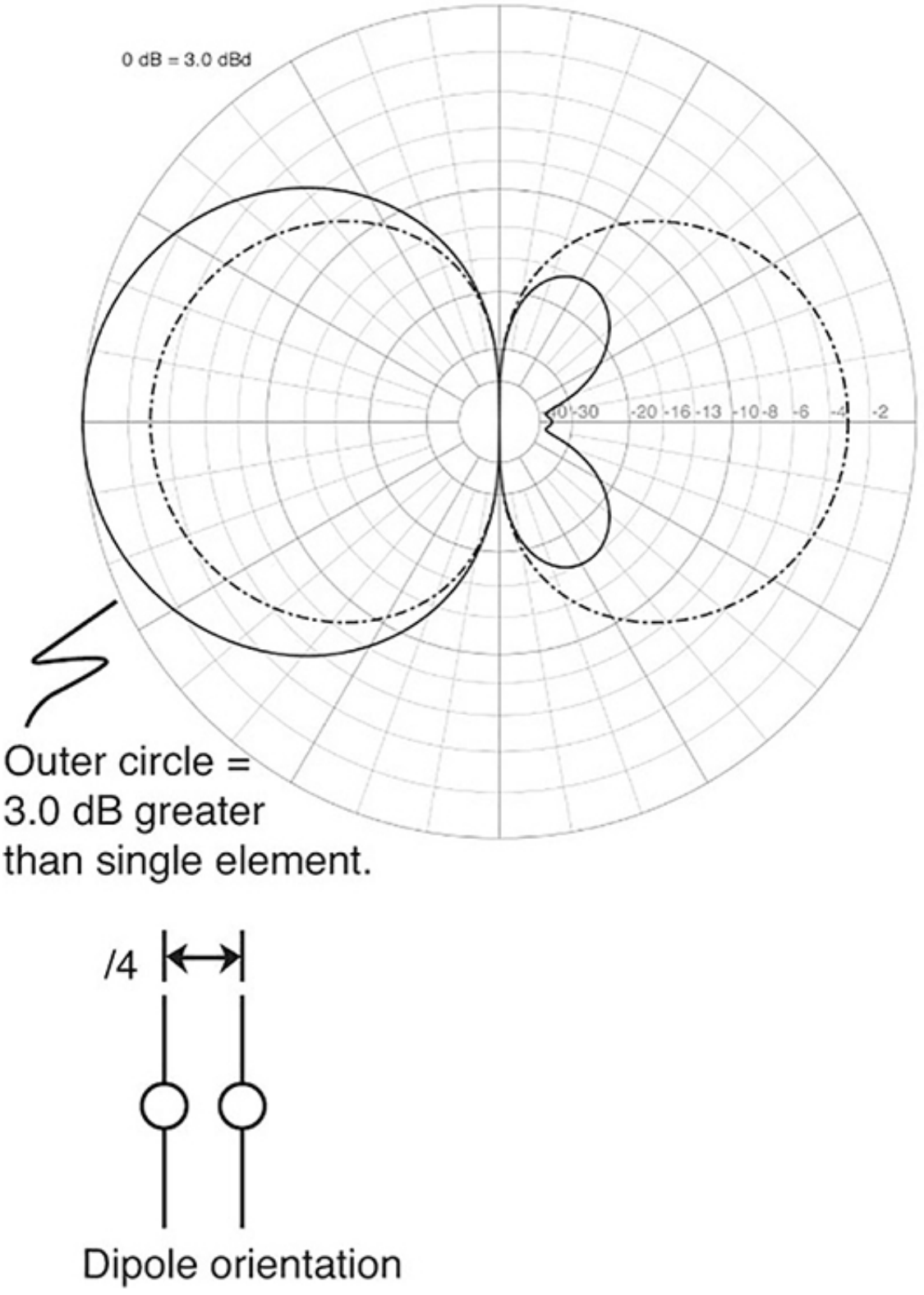
F IGURE 7.15 Cardioid pattern using omnidirectional radiating elements.

We say this two-element array produces a unidirectional cardioid pattern. The best antenna elements to use for forming a cardioid are omnidirectional radiators in the azimuthal plane—verticals, in other words—but horizontal

$\lambda/2$

dipoles at identical heights above ground and oriented so that their maximum radiation is broadside to the array pattern maximum are a

possible alternative if additional side nulls and rear fishtail lobes (Fig. 7.16) in the pattern are acceptable.



F IGURE 7.16 Cardioid pattern from two $\lambda/2$

horizontal dipoles.

7.3 Feeding Arrays

Feed systems for various members of the loop and collinear antenna families are the easiest because the second element is driven by the first element, so there is little or nothing for the user to do to make sure currents in the two elements are equal or nearly so.

In theory, a $\lambda/2$ dipole in free space has a feedpoint impedance of 73 Ω resistive. In practice, MF and HF dipoles are something other than 73 Ω and they often have a reactive component to their impedance. Even in free space, the second dipole in a loop may well cause the entire structure to have a feedpoint impedance other than 73 Ω. The exact change depends on the distance between the centers of the two dipoles. But the methods for canceling out the reactive part and transforming the real part to 50 Ω are the same as for a simple dipole. As a general rule, open-wire line, used in conjunction with an antenna tuning unit (ATU) at the transmitter end of the line or at ground level directly below the loop, is an excellent choice and usually represents less of a downward drag on the antenna. But 50-Ω coaxial cable may actually be a closer match to the loop's natural impedance.

Next in ease of implementation are bidirectional line arrays with equal element currents. Transmitter power can usually be equally divided among them simply by feeding multiple identical transmission lines in parallel. Often it will be useful to incorporate one or more impedance-transforming sections of appropriate transmission line to convert the typically low junction impedance to something higher and more appropriate for the main run back to the transmitter.

When we move to element feeding phase differences other than 0 and 180 degrees, feeding the array becomes more difficult. The impedances of the individual elements are different and the elements accept differing amounts of power from the source or passive component divider networks unless special techniques are employed

for ensuring equal power or equal current at all the element feedpoints.

As we saw in Chap. 6, transmission lines exhibit a velocity of propagation that is less than the speed of light in free space or air. As a result, a length of transmission line used to create a 90-degree delay at some operating frequency will be shorter than the physical distance between two array elements spaced $\lambda/4$ apart as measured in air or free space. If the layout of the array is such that the physical distance of each such delay line from a central combining point to each element is substantially shorter than the distances between the elements themselves, it may be possible to phase each element with a minimum transmission line length. However, frequently it is necessary to make the phasing lines longer, such as using a 3 $\lambda/4$ line to provide a 90-degree delay.

Some important points when designing phasing lines for arrays with separately fed elements:

- Construct feed lines out of the same transmission line. To ensure precise phase control, lines to individual elements of a given array should be cut from the same spool.
- Determine the phase velocity for the cable you are using. If accurate manufacturer values are unavailable, use a time domain reflectometer (TDR), a vector network analyzer (VNA), or other suitable instrument. See Chap. 30 for some instruments and techniques. This is less important for broadside arrays where all the elements are fed in phase through identical lengths of cable originating at a common junction, but even in this situation, cables between the common junction and individual elements should be obtained from the same spool or production lot.
- Measure cable lengths accurately. Even at the MF and lower HF bands, keep lengths of phasing lines within a few inches of each other.

- Ensure that matching networks between the central distribution point and each element of the array are identical and avoid unexpected phase shifts or delays in the signals being sent to the elements.

While we have framed our discussion in terms of transmitting arrays, most of these considerations are valid for receiving arrays.

7.4 Parasitic Arrays

For certain types of arrays, some of the antennas or structures in the array are not driven directly. If an object or antenna is in the vicinity of a driven antenna, it scatters the electromagnetic fields radiated or received by the antenna. By deliberately positioning one or more conductors near a simple radiating element such as a dipole, we can increase the field strength in some directions at the expense of others. The resulting configuration is known as a parasitic array.

The principal value of a parasitic array is its potential for simplifying the feed system. Parasitic arrays usually use a combination of element-to-element spacings and element dimensions to create desired array patterns without the use of phasing lines and multiple matching networks. Because of this, they are mechanically simpler to construct. For many applications, parasitic arrays are the antenna of choice in view of their reasonably high gain (6 dB or more) and ease of rotation and aiming.

Parasitic arrays can be vertically or horizontally polarized. They can be made of

$\lambda/4$

grounded elements or

$\lambda/2$

symmetrical elements constructed from copper wire (typically on 160, 80, and 40 m bands), aluminum tubing (80 through 6 m bands), or solid aluminum rod (6 m and up in frequency). Vertical arrays for 160 through 40 m are often constructed with triangular lattice tower sections so that they provide their own support mechanism.

A mathematical explanation of how parasitic arrays work is beyond the scope of this book. Instead, a short qualitative description is

provided as a guide to helping the reader develop an intuitive feel for their operation.

Picture a horizontal $\lambda/2$ dipole up in the air somewhere, fed with RF energy at its resonant frequency. With nothing around it, its radiation pattern will be the familiar donut shape of Fig. 5.11. Now suppose we place another thin metallic object, of approximately the same dimensions, near the dipole (typically within a wavelength or less) and in the dipole's broadside radiation lobe. Electromagnetic waves from the dipole will impinge upon this neighboring object, inducing currents and voltages of the same frequency along the object's surface. These currents radiate electromagnetic waves, much like the fields from the original dipole. A distant receiving antenna will pick up a signal that is the linear combination (superposition) of the signal from the dipole and the scattered signal from the metallic object. The scattered signal arrives at a distant receiving point with amplitude and phase that are different from those of the dipole. Depending on the different paths traveled by the two waves and the degree to which the radiated fields from the added conductor differ in phase from those of the dipole, the signals may add or subtract at the receiving antenna.

The superposition at a distant receiving point of waves from the two radiating conductors is identical to the process we saw previously with all-driven arrays. We have created an array despite the fact that only one of its elements is driven. We call that element the driven element and the other conductor is called a parasitic element.

With an all-driven array, the pattern and gain are controlled by judicious choice of element spacing and by individually setting the feedpoint current amplitude and phase for each element of the array. With a parasitic array, the only control we have is the ability to adjust element dimensions and interelement spacings.

If the parasitic element is identical to the dipole and is resonant at the operating frequency it is receptive to the generation of induced currents and voltages from the dipole's fields. If we slightly lengthen the parasitic element we can control the phase of the scattered signal. A small change in length does not materially reduce its ability

to pick up and reradiate a portion of the passing field from the dipole, but it does substantially alter the phase of the currents induced in it. For typical interelement spacings, the resulting field from the parasitic element opposes or partially cancels the dipole field in directions beyond the added element, such that the overall field seen by a distant receiving antenna is significantly reduced. If we instead make the parasitic element somewhat shorter than its self-resonant length at the operating frequency, the reradiated fields are again of a different phase from the dipole's fields, but the effect is usually to add to the dipole's fields at a distant receiving point beyond the added element.

In the first case we call the parasitic element a reflector and in the second we call it a director. Depending on the specific design objectives for a parasitic array, there can be multiple reflectors and/or multiple directors. Compared to all-driven arrays, interelement spacings in parasitic arrays are often smaller; 0.1λ to 0.3λ is typical.

The most common parasitic array is the Yagi-Uda beam antenna—often called Yagi, for short. It has attained such a level of popularity that an entire chapter (Chap. 14) of this book is devoted to it. But other kinds have existed for at least as long, including many wire arrays and combinations of grounded verticals. We will explore these in more detail in later chapters specifically dedicated to the different families of arrays.

7.5 Ground Effects

In the real world, few antennas are located in free space. Only at VHF and above can we raise an antenna to heights that approximate free space. When discussing antenna operation and performance, we often must incorporate the effects of nearby conductors on the antenna and the waves it radiates or receives.

When the dimensions of the nearby conductor are large enough to materially alter the radiation pattern of the antenna at the chosen operating frequency, we call it a ground plane or, more simply, ground. An ideal ground is a perfectly conducting conductor stretching infinitely far in all directions. For many scenarios, especially

those involving frequencies below VHF, real ground is far from ideal. For mobile VHF and UHF whips mounted atop a vehicle, a vehicular metal roof is often reasonably well approximated by an ideal ground.

The performance of an antenna located above ground is affected in three ways:

- Like any other conductor in the vicinity of a basic antenna, a ground plane functions as a parasitic element by reradiating the EM fields originating at the antenna. The resulting signal strength far from the antenna is thus a combination of signals received directly from the antenna as well as via the ground plane. For an ideal ground plane, it is as if there is an exact copy of the radiating antenna below the ground plane, either in phase or out of phase with the above-ground antenna depending on the polarization. This is referred to as an image.
- Some fraction of the reradiated energy from the ground plane returns to the antenna. This alters the current and voltage relationships at the antenna and changes the input impedance. Depending on the spacing between the antenna and the ground plane, the radiated field strength for a given excitation level from the transmitter may be greater or smaller than in the free-space case.
- If the ground is not perfectly conducting, a portion of the RF energy that impinges upon it will be dissipated through loss in the resistive component of the ground. As we discuss in Chap. 11, ground-mounted vertical monopoles are more susceptible to this effect than horizontal dipoles because the return current for the vertical comes back to its feedpoint through lossy earth in the region immediately surrounding the vertical.

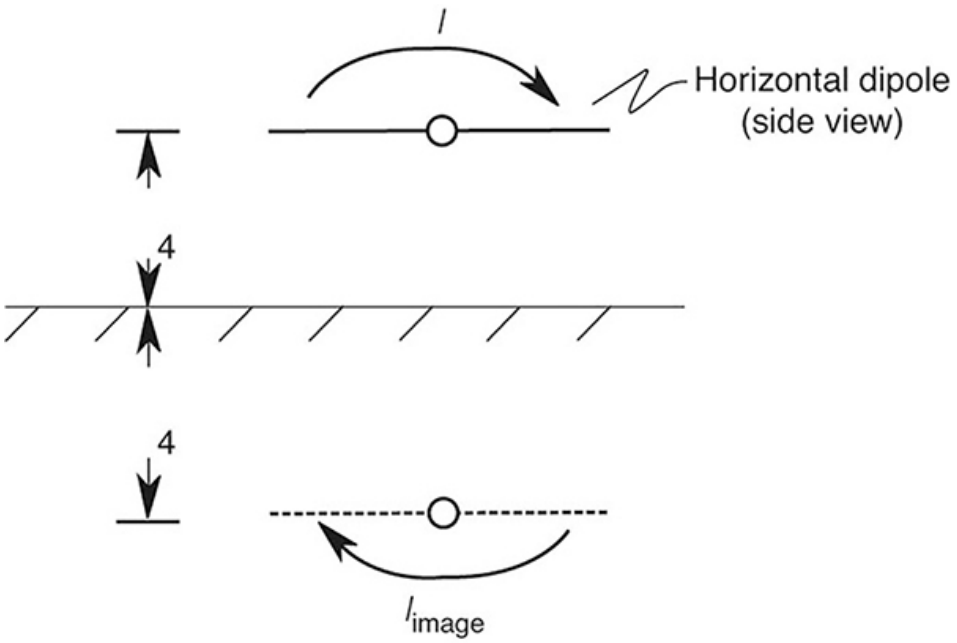
The effect of an ideal, perfectly conducting ground plane can be understood through electromagnetic boundary conditions. A perfectly conducting material cannot support a voltage differential between any two points on or inside the material. Perfectly conducting means the resistance between any two points on the conductor is zero, so

any voltage differential across those points would imply the existence of an infinitely large current flowing between them. (Anyone who has ever dropped a wrench across the two terminals of a car battery has a good grasp of this concept!) Since a voltage differential is the result of an E-field between the two points, there must be no E-field tangential to the surface of the perfect conductor.

If a wave strikes an ideal ground plane, in order for the E-field to be zero along the surface of the ground, a reflected wave must be produced. For an antenna above an ideal ground plane, a horizontal E-field from the antenna must be matched by identical fields of equal amplitude and opposite polarity reflected from the ground. The reflected wave appears to be produced by an image antenna, similar to the original antenna but a mirror image located same distance below the ground plane as the original antenna is above it. The horizontal components of its currents, voltages, and fields are equal in amplitude but opposite in direction (or phase) to those of the original antenna.

The image antenna concept allows us to predict the effect of a ground plane on the radiation pattern of a nearby antenna simply by removing the ground plane and replacing it with the image antenna. We can determine the effect of a nearby ground plane on an antenna's pattern by removing the ground plane and treating the antenna and its image as a two-element array. Since we've already seen how to deal with two-element arrays whose elements are fed out of phase and in phase, we now can use the same visualizations to see what happens to far-field patterns from an antenna over perfectly conducting ground.

Example 7.1 What is the broadside elevation response for a horizontal $\lambda/2$ -dipole of Fig. 7.17 at a height of $\lambda/4$ above ground?



F IGURE 7.17 A horizontal dipole and its image antenna.

Solution The dipole is
$\lambda/4$
above ground, so the image antenna is
$\lambda/4$
below the surface of the ground. Figure 7.18 shows the signal paths
r_1
and
r_2
from the two dipoles to a distant receiver. Because the original dipole is horizontal with respect to the ground plane), the image dipole is also horizontal but the current is 180 out of phase with respect to the original dipole.

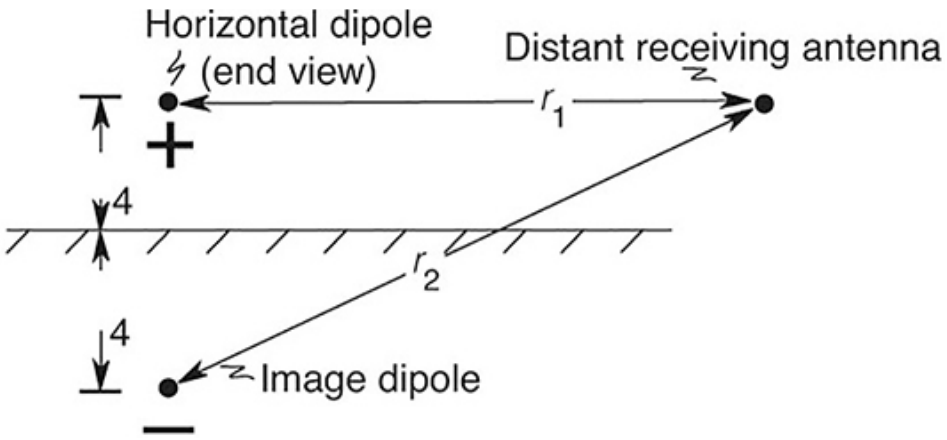
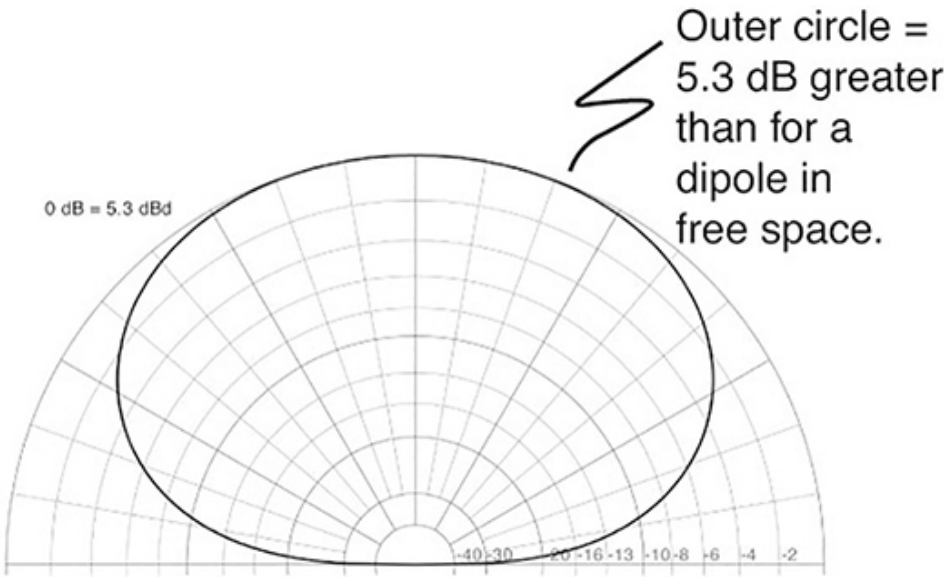


FIGURE 7.18 Combined fields from antenna and image at a distant receiver.

The spacing between the original dipole and the image antenna is twice $\lambda/4$, or $\lambda/2$. Because the antenna excitations are out of phase, the combined signal strength at any far-field receiving point that is equidistant from them is zero. This occurs in the horizontal direction along the horizon.

Further, because the wave from the image antenna takes a half-cycle to travel the half-wavelength straight up toward the original dipole, it arrives in perfect phase at the other antenna, so the far field strength is at a maximum in the vertical direction. At elevation angles between vertical (90 degrees) and the horizon (0 degrees), the far field strength takes on intermediate values, as shown in Fig. 7.19.



F IGURE 7.19 Total received field strength versus elevation angle broadside to dipole of (A).

Because the main beam for a horizontal dipole above a ground plane is in the vertical direction, this antenna is sometimes called a cloud burner, and it is ideal for near-vertical incidence skywave (NVIS) communications. The broadside pattern is quite strong until the vertical elevation or takeoff angle is below 30 degrees (the 3-dB point). For the 80 and 40 m bands, a

$\lambda/2$

dipole at a height of roughly

$\lambda/4$

(35 ft) is a good choice.

For takeoff angles above 30 degrees, the signal at the receiver is more than double that of the original antenna in free space. The image antenna—which is really a way of looking at the ground reflection effect—has increased the received signal strength of our original antenna by nearly 6 dB relative to the free-space case.

The degree of reinforcement or cancellation at any given elevation angle above an arbitrary horizontally polarized antenna depends on the height of that antenna above ground. As the height of the antenna above ground is increased, eventually the phase shift of the wave from the image antenna on its way to distant receiving points

increases to where the maximum radiated field is no longer straight up in the air. With increasing height above ground, the elevation angle corresponding to maximum radiated field strength begins to drop. No matter how high the antenna is raised, as long as it is above a ground plane, the horizontally polarized radiation at a 0 degree elevation angle (i.e., the horizon when flat land is involved) will always be zero because the image antenna and the real antenna are exactly out of phase on these equidistant paths.

We now consider a vertically polarized dipole not in contact with the earth, as shown in Fig. 7.20. Boundary conditions at the ground plane make the image antenna's vertically polarized E-field in phase with the vertical component of the original antenna's radiated field. The ground beneath a vertical

$\lambda/2$

dipole is equivalent to an identical dipole having the same vertical orientation and a standing current that is exactly in phase with the drive current of the aboveground dipole. Figure 7.23 at the end of the chapter consists of a series of graphs depicting the ground reflection factor for vertically polarized antennas as a function of their height above ground.

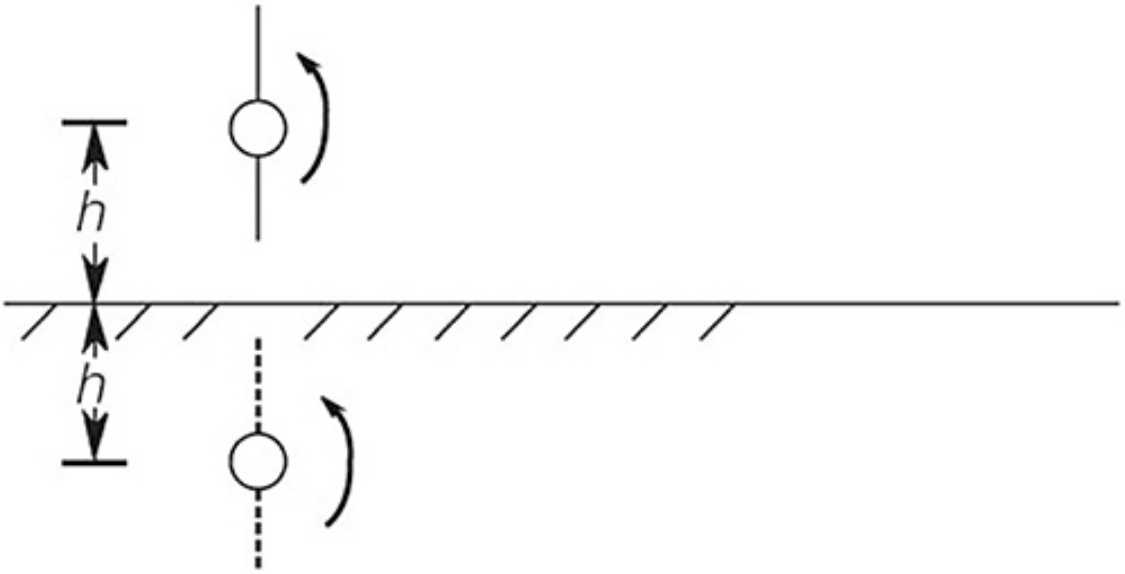


FIGURE 7.20 Image antenna for vertical dipole not in contact with the ground.

For tilted antennas, or antennas otherwise having a mix of both horizontal and vertical polarization, visualizing the two cases separately may be the easiest way to understand the overall effect of nearby ground on the radiated signal. If needed, a tilted antenna can be modeled using software.

For the specific case of a ground-mounted monopole, the image antenna is best visualized as an identical but upside-down vertical with its upper end touching the base of the real antenna. If the current flow at some instant is downward in the real monopole, it is also downward in the image monopole. Because the vertical component of the two currents is in phase, the current distribution in the two verticals appears continuous across the ground plane; consequently, the effect of the ground-plane boundary conditions is to cause the image antenna to act like the other half of a vertical dipole. Like any other dipole, the standing current is maximum at the center of the resulting antenna structure and in phase across the entire combined length of the two pieces, as suggested by Fig. 7.21.

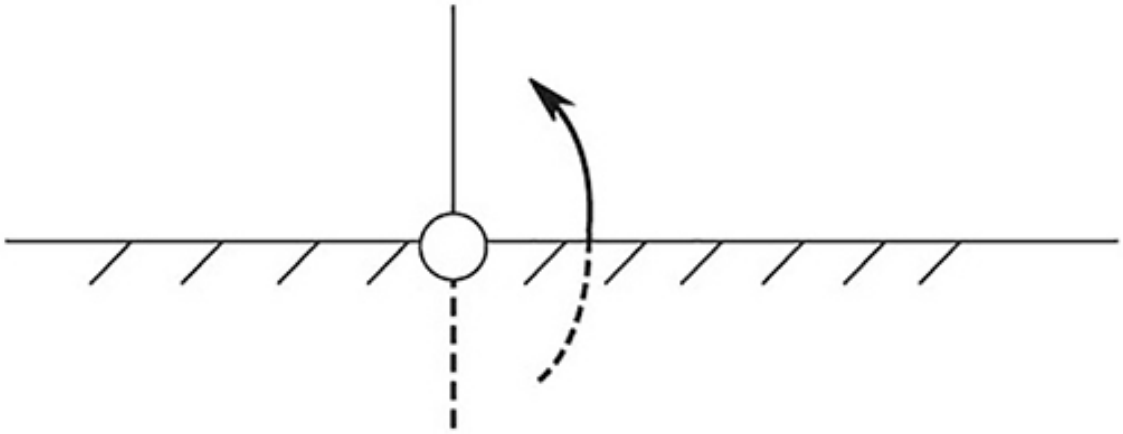


FIGURE 7.21 Image antenna for ground-mounted vertical monopole.

Because the radiated E-fields from the real and image vertical antennas are in phase and at right angles to the ground plane, there is no cancellation of the radiated field at a 0 degree elevation angle. In the presence of a perfectly conducting ground plane, a vertical enjoys the same 6-dB boost in received signal strength at certain

wave angles as do horizontal dipoles, but—because its E-field is at right angles to the ground plane—this boost occurs at 0 degrees. The ground-mounted vertical is unsurpassed in its ability to produce a strong signal at the horizon and very low elevation angles. As discussed in Chap. 3, the losses inherent in real earth cause the distant radiated field from any vertical to be zero near the horizon. Despite that, numerous expeditions have demonstrated that ground-mounted monopoles at the edge of ocean saltwater have an advantage over virtually any other antenna for intercontinental communications on the MF and HF bands.

7.6 Lossy Ground

The discussion of ground reflection effects thus far assumes perfectly conducting ground. For average or poor grounds, the effect is to partially fill in some of the nulls in the elevation pattern (so they seldom are deeper than perhaps 15 dB) and to reduce the maximum amplitude of the array factor in its favored directions, especially as the frequency increases. Lobes above the lowest angle lobe are reduced even further.

Especially at the lower elevation angles, ground characteristics that are important are not the ones for the ground directly beneath the antenna; rather, it is the ground many wavelengths from the antenna that establishes the efficiency of the reflection that can add as much as 6 dB to the signal compared to the same antenna in free space. For instance, for a horizontal dipole or Yagi at height λ, a 10-degree takeoff angle is reinforced by the ground reflection 5.7λ away; on 14 MHz that's about 400 ft! At 1 degree (the "sweet spot" in the minds of many DXers), the first bounce is nearly a mile away.

With real earth beneath an antenna, the effective ground can lie some distance beneath the physical surface. How far beneath will depend on the operating frequency and the actual characteristics of the earth at that frequency, but distances of a few inches to a few feet are common.

Figures 7.22 and 7.23 at the end of this chapter display the ground reflection factors for horizontally polarized and vertically polarized

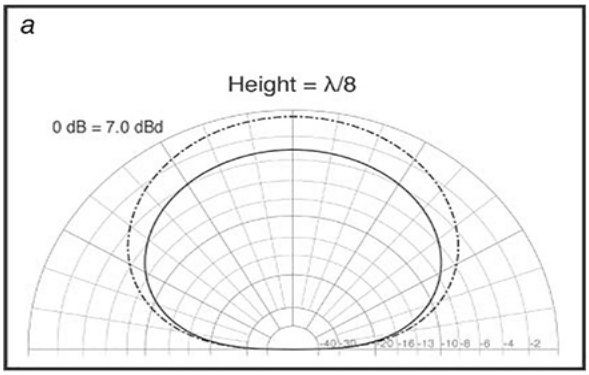
antennas, respectively, at selected heights above ground. (The charts of Fig. 7.22 are for an azimuth broadside to the axis of the dipole.) Each individual chart shows, for a given height, both the theoretical elevation pattern over perfectly conducting ground (broken line) and the equivalent pattern (solid line) over average ground (

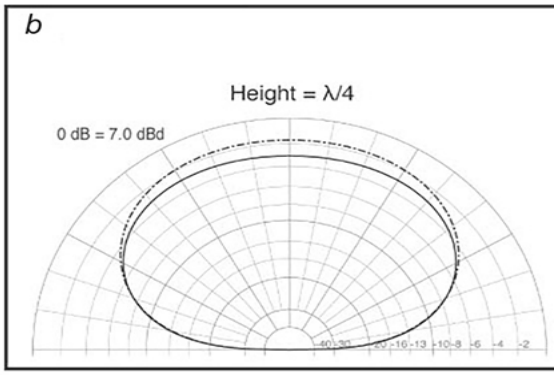
$$\sigma = 0.005\ \mathrm{S/m}$$

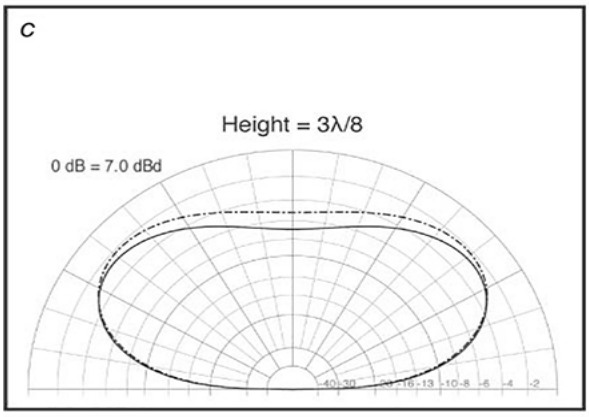
and

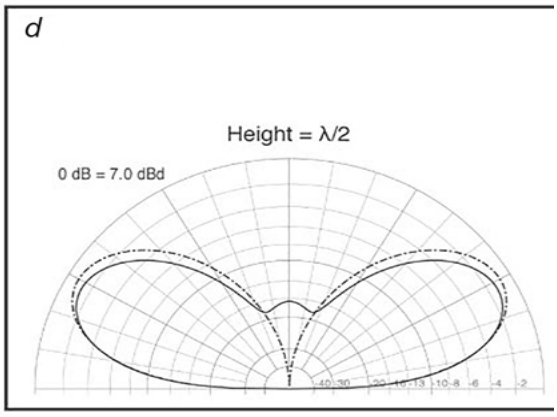
$$\epsilon_r = 13$$

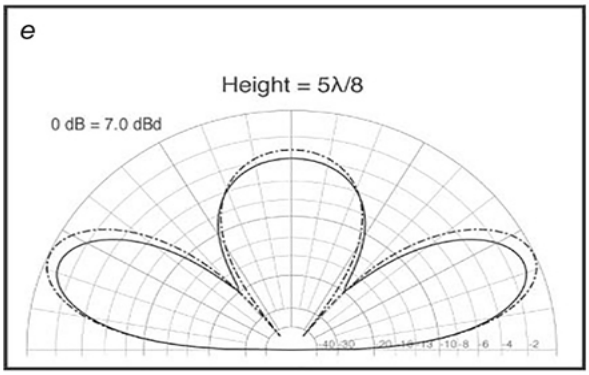
). In all of these charts, the outer ring of the plot corresponds to 7.0 dBd, or a maximum signal strength that is 7.0 dB greater than would be obtained from an identical dipole in free space.

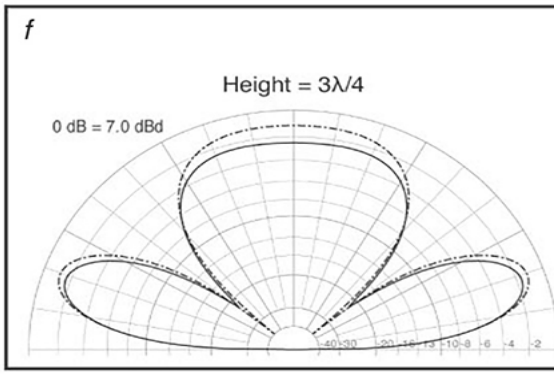

a
Height = λ/8
0 dB = 7.0 dBd

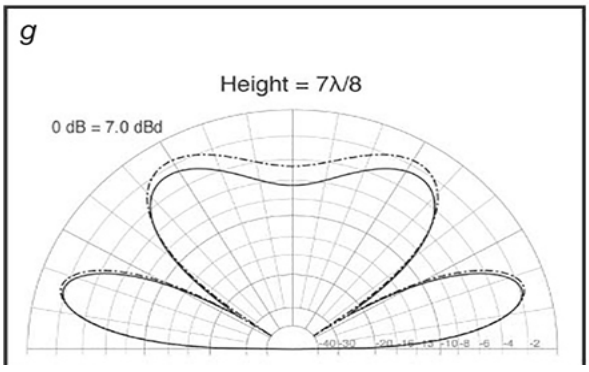

b
Height = λ/4
0 dB = 7.0 dBd

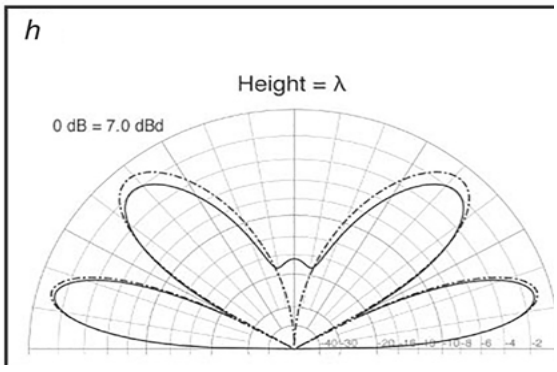

c
Height = 3λ/8
0 dB = 7.0 dBd


d
Height = λ/2
0 dB = 7.0 dBd


e
Height = 5λ/8
0 dB = 7.0 dBd

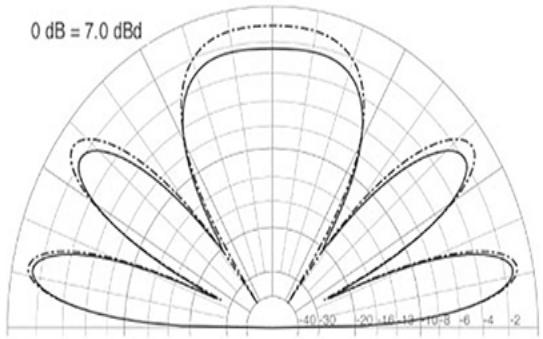

f
Height = 3λ/4
0 dB = 7.0 dBd


g
Height = 7λ/8
0 dB = 7.0 dBd


h
Height = λ
0 dB = 7.0 dBd

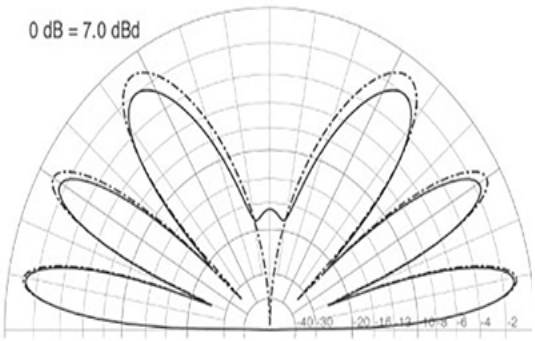
i

Height = 5λ/4



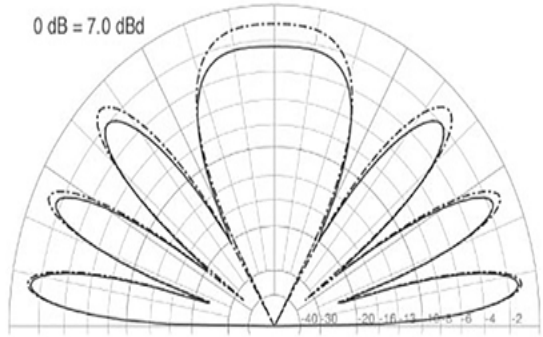
j

Height = 3λ/2



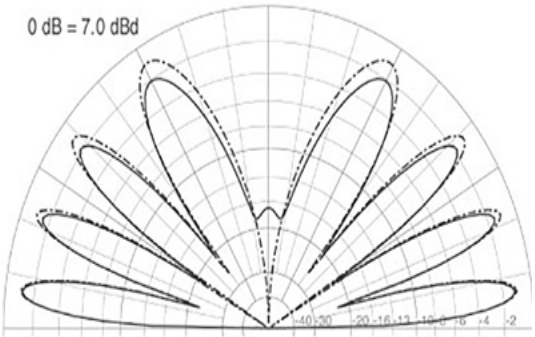
k

Height = 7λ/4



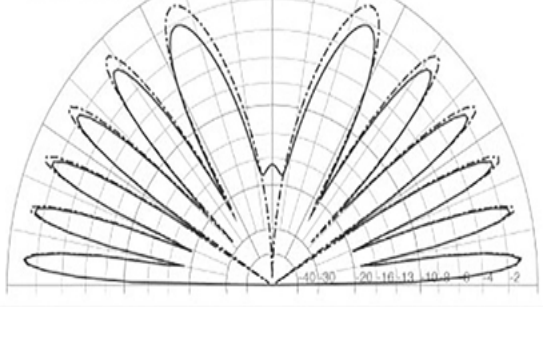
l

Height = 2λ



m

Height = 3λ



n

Height = 4λ

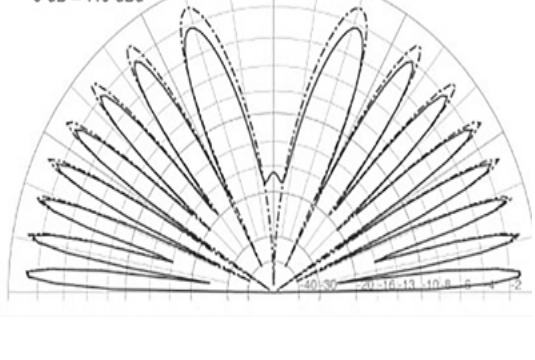
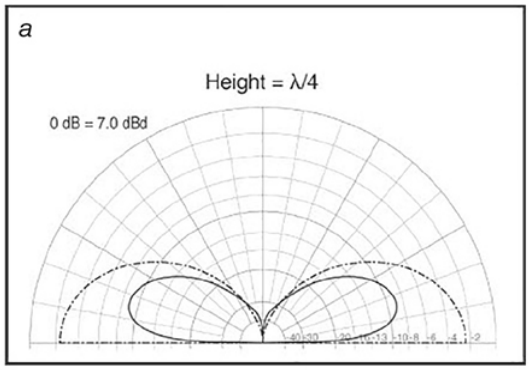
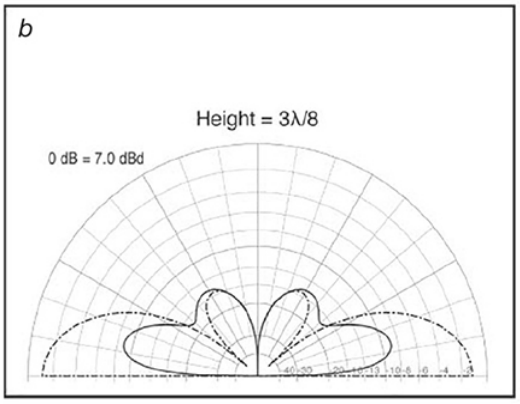
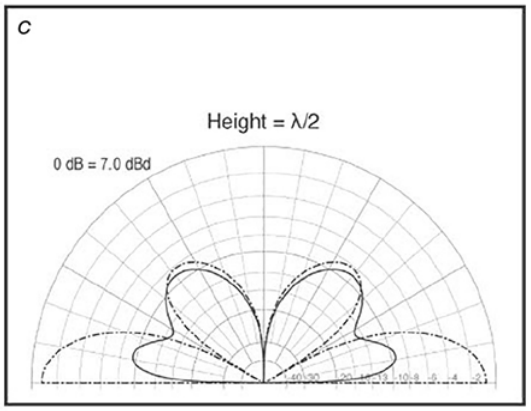
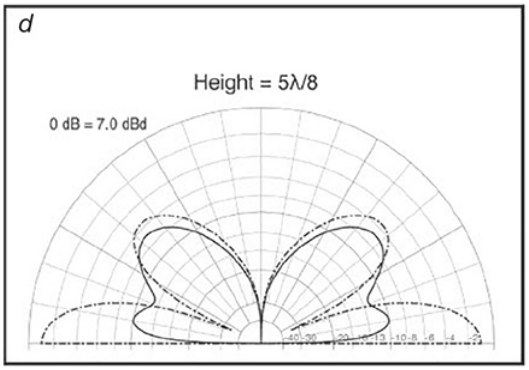


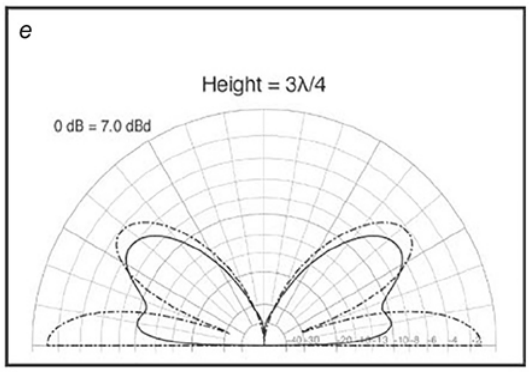
FIGURE 7.22 Broadside ground reflection factors for horizontal $\lambda/2$ dipoles at selected heights above perfect (broken line) and average (solid line) ground. (*Continued*)

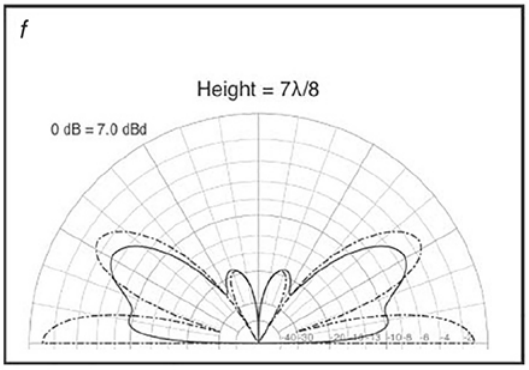

a
Height = λ/4
0 dB = 7.0 dBd


b
Height = 3λ/8
0 dB = 7.0 dBd


c
Height = λ/2
0 dB = 7.0 dBd


d
Height = 5λ/8
0 dB = 7.0 dBd


e
Height = 3λ/4
0 dB = 7.0 dBd


f
Height = 7λ/8
0 dB = 7.0 dBd

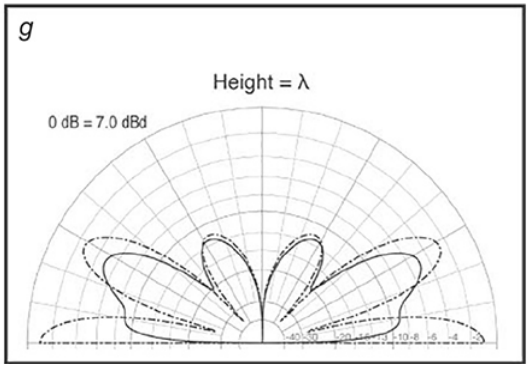

g
Height = λ
0 dB = 7.0 dBd

FIGURE 7.23 Ground reflection factors for vertical $\lambda/2$ dipoles at selected heights above perfect (broken line) and average (solid line) ground.

It can be useful to determine the optimal height for maximizing an antenna's utility over the intended range of communications distances. Especially for individuals planning to erect MF or HF antennas, Figs. 7.22 and 7.23 can help with this.

The charts of Figs. 7.22 and 7.23 assume flat ground under and near the antenna. The effect of ground that gradually slopes down as one gets farther and farther from the antenna is to enhance the low angle signal strength in the same manner that hoisting the antenna higher can. Conversely, ground that slopes up tends to mimic lower antenna heights. As discussed in Chap. 3, terrain analysis programs can help identify or predict the effect of a specific terrain contour on the performance of a specific antenna installation.

Example 7.2 An amateur radio operator's favorite band for casual operating is 20 m, but she also spends time DXing on 15 and 10 m when those bands are open. She is planning to purchase a triband Yagi antenna but is trying to decide whether to put up a 35-ft house-bracketed tower or a guyed 70-ft tower to support the antenna.

Solution From what we know about the characteristics of those bands throughout most years of a typical sunspot cycle, we don't expect them to be open for short skip very often, so she should try to select a height that favors elevation angles below 50 degrees on 20 m, somewhat lower angles on 15, and even lower angles on 10. A height of $\lambda/2$ on any band provides a very broad lobe centered on 25 degrees—right in the middle of her desired range of angles. On 20 m, $\lambda/2$ corresponds to 35 ft, which is just about the practical minimum if the antenna is to be above nearby obstructions. On 10 m, a triband beam at 35 ft is a full λ high and the peak amplitude of the main lobe has moved lower in takeoff angle even as a second lobe, slightly weaker than the

main lobe, appears in the elevation pattern, as seen in Fig. 7.22*h*. Of course, average ground will act more like a dielectric and less like a perfect mirror at 28 MHz, so the nulls will partially fill in. Except at times of extremely high sunspot activity or sporadic E-skip, it's very unlikely that the ionosphere will support high-angle F-layer paths at 28 MHz, so the most useful elevation angles on 10 m are almost always lower than those of 15 or 20.

If our operator is interested primarily in working very long haul DX on 20 m, she might want to consider heights closer to λ (70 ft) at the risk of inserting some nulls at useful angles in her 15-and 10-m elevation patterns. By raising her triband beam from 35 to 70 ft over average ground, she gains not only another decibel of forward gain at the peak of the main lobe, but the elevation angle of that peak drops from 25 degrees to 14 degrees. At the low elevation angles that are so useful for long-haul DXing, the resultant overall improvement from doubling the height of her beam from

$\lambda/2$

to λ is about 5 dB—equivalent to tripling her transmitter output power! In return, a null appears at 30 degrees (reducing her signal for some short-haul elevation angles on 20 m) and the elevation patterns on 15 and 10 break up into multiple lobes, albeit with lower peak elevation angles. In short, a height of 70 ft or as close to it as she can get will put most of her output power in the right range of takeoff angles on the band where she will need it the most—20 m—at a cost of some higher-angle nulls on all three bands.

PART II

Building-Block Antennas

CHAPTER 8
Dipoles
CHAPTER 9
Large Wire Loop Antennas
CHAPTER 10
Multiband and Tunable Wire Antennas
CHAPTER 11
Vertically Polarized Antennas

CHAPTER 8
Dipoles

Expensive antennas are not always required for effective radio communications. Many of the world's strongest signals and most competitive amateur radio enthusiasts employ simple but effective antennas that can be erected by inexperienced people without any special tools or equipment. The key to success in these cases is knowledge of a few key fundamental concepts, coupled with intelligent planning and siting of the antenna based on the user's objectives.

8.1 The Half-Wave Dipole

Leading other antenna types in performance versus simplicity is the half-wave dipole. The half-wave dipole can consist of nothing more than a length of wire. An understanding of how a $\lambda/2$ dipole functions is perhaps the single most important piece of knowledge to take away from this book. The dipole forms the basic building block for countless other types of antennas: the Yagi, the quad, the collinear, the loop, and others. Whether used individually or in concert with one or more additional dipole-like elements, the half-wave dipole excels at efficiency of radiation and ease of matching to feedlines.

Occasionally you may hear also the dipole called a Hertzian dipole or doublet. While usage has evolved over time, the term "Hertzian dipole" is usually applied to electrically short dipole antennas. Historically, the terms "dipole" and "doublet" referred to two elements of charge, one positive and one negative, driven to opposite ends of an infinitesimally short Hertzian dipole by a signal source. As noted in Chap. 5, many engineering texts on antennas use the Hertzian dipole

as their starting point for mathematical derivation of the characteristics of all dipole structures, including half-wave dipoles.

The half-wavelength dipole is an electrically balanced antenna consisting of a single conductor one half-wavelength long. Most commonly the wire is cut in the middle and a signal source injected there, as shown in Fig. 8.1. This configuration is the ubiquitous center-fed halfwave dipole.

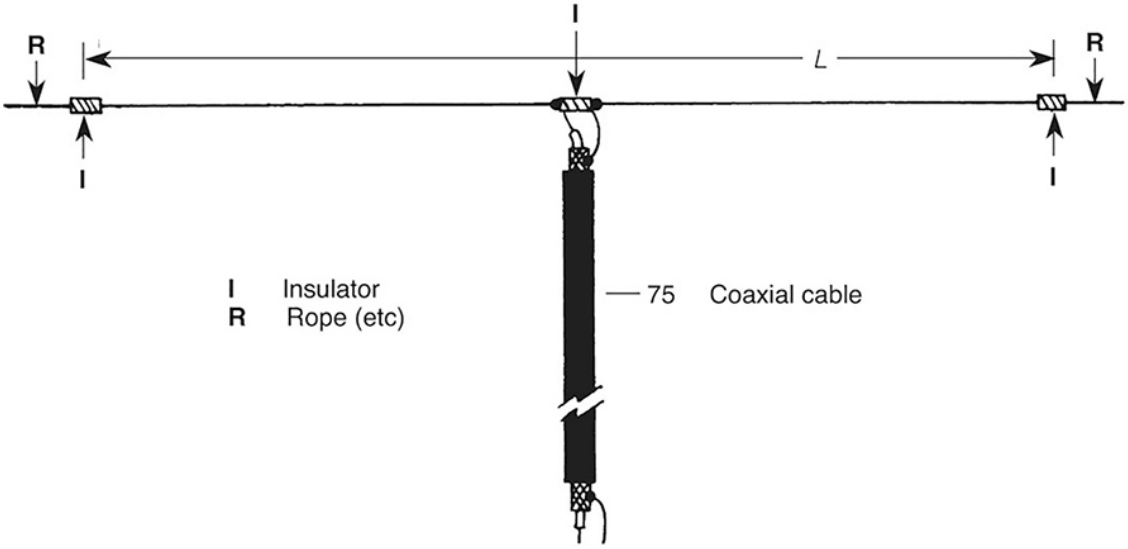


FIGURE 8.1 Simple half-wave dipole antenna.

8.1.1 Dipole Construction

The half-wave dipole is so fundamental to radio communications that it is discussed in many places throughout this book. The dipole can be mounted vertically as described in detail in Sec. 11.1 or horizontally as in Sec. 14.4.1.

While the dipole is used in many different configurations across radio communications bands, dipoles are usually constructed from copper wire or other thin conductor. Because the velocity of propagation of an electromagnetic wave along a wire is slightly less than in free space, the half-wave dipole resonates at a frequency where it is about 5 percent shorter than a half-wavelength in free space. The free-space half-wavelength

$$\frac{\lambda}{2}\,(\text{feet}) = \frac{492}{f(\text{MHz})} \quad (8.1)$$

or, in metric terms,

$$\frac{\lambda}{2}\,(\text{meters}) = \frac{150}{f(\text{MHz})} \quad (8.2)$$

where *f* is the operating frequency in megahertz. For thin dipoles made from a conductor having a diameter or cross-sectional side much smaller than the dipole length (i.e., the usual case for dipoles), the length of a resonant half-wave dipole antenna is

$$L\,(\text{feet}) = \frac{468}{f(\text{MHz})} \quad (8.3)$$

Because of end effects and interactions with supports and other structures, this length is approximate. Trial and error may be required to arrive at a length that minimizes the reactive part of the feedpoint impedance.

Example 8.1 Calculate the approximate physical length for a half-wavelength dipole operating on a frequency of 7.15 MHz, the middle of the U.S. 40-m band.

Solution

$$\begin{aligned} L &= \frac{468}{F} \\ &= \frac{468}{7.15} \\ &= 65.45\ \text{ft} \end{aligned}$$

In practice, one might spool off about 70 ft of wire, fold it in half, and cut it at the midpoint. This leaves a foot or more of wire at each end of the two halves for threading through end and center insulators and wrapping back on itself.

Example 8.2 Find the length of a
$\lambda/2$
dipole antenna for a frequency of 24.930 MHz in the 12-m amateur radio band.

Solution

$$L\ (\text{ft}) = \frac{468}{F(\text{MHz})} = \frac{468}{24.930} = 18.77\ \text{ft}$$

The simplest way to feed this antenna is to break the dipole in the center and attach a 72-Ω coaxial cable there. The dipole then consists of two separate conductors, each of which is one-half the overall length (or, in the example given for the 12-m band, about 9.4 ft).

At HF and below, the antenna is usually installed horizontally with respect to the earth's surface or other local ground plane. The resulting electric fields at distances greater than a few wavelengths from the antenna lie parallel to the wire. The horizontal dipole produces a horizontally polarized radio wave. At VHF and above, dipoles are commonly found in either orientation.

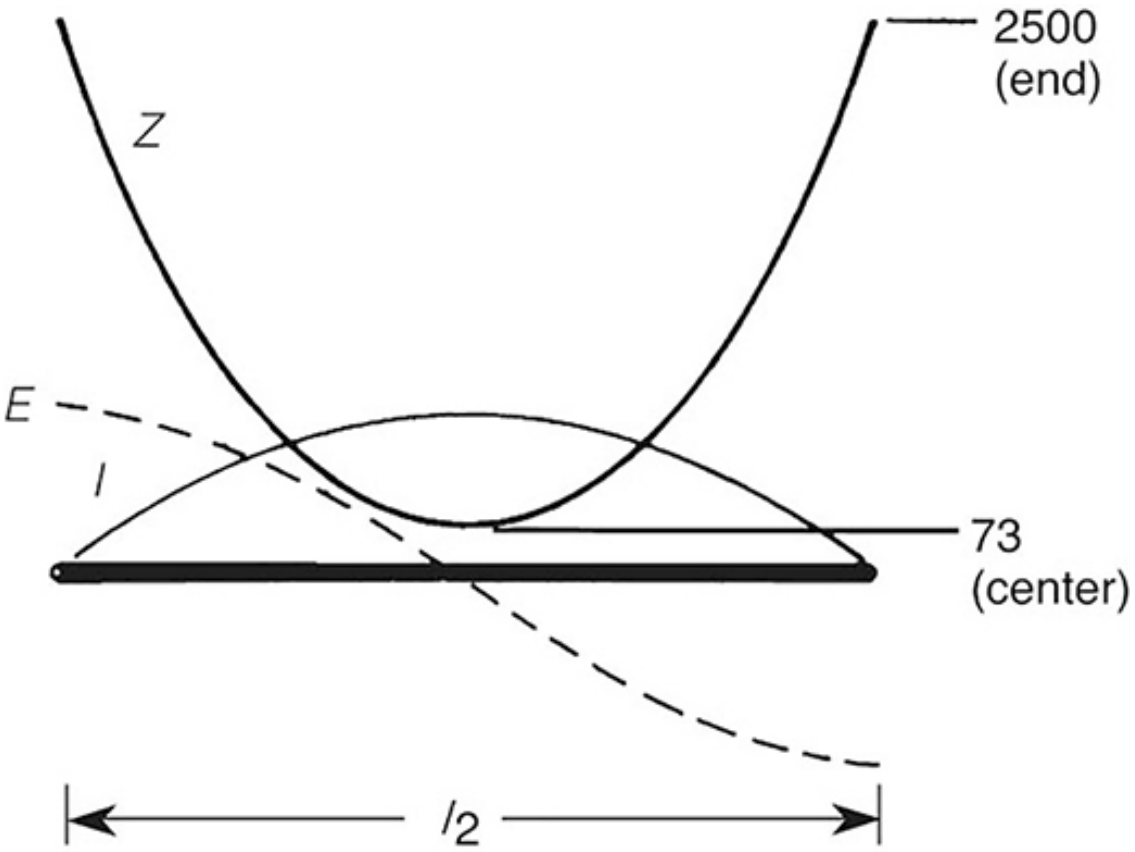
For submicrowave applications, the length of a half-wave dipole does not need to be very accurate. Local factors—nearby objects, height above ground, and the length to diameter ratio of the conductor to name a few—might make it necessary to either add or trim a short length of wire on each side to achieve exact resonance. The primary reason to aim for approximately a half-wavelength is to make the antenna wire resonant, or nearly so, at the frequency of operation so that we can minimize the reactive part of the feedpoint impedance that will have to be tuned out by the transmitter, an antenna tuning unit (ATU), or other matching network. Since interactions with structures near the antenna can influence the antenna input impedance, *in situ* tuning is often required.

The accuracy required for a half-wave dipole length also depends on the percentage bandwidth of the application. For a half-wave dipole cut to length for the center (14.175 MHz) of the U.S. 20-m band, the antenna might be used at frequencies ranging between

plus and minus 1.2 percent of the center frequency, and there will be a generally insignificant change in the feedpoint impedance over the entire band. An 80-m dipole, on the other hand, could conceivably be called upon to span a range of frequencies up to plus and minus nearly 7 percent of the band center (3.750 MHz). For some transmitters, the feedpoint characteristics will change too much for proper matching and full power transfer to the antenna over the entire range without inclusion of an adjustable matching network somewhere between the transmitter and the antenna, or without some modification to the antenna to increase its bandwidth.

8.1.2 The Dipole Feedpoint

Figure 8.2 shows the voltage (*V*) and current (*I*) distributions along a half-wave dipole. If center-fed, the feedpoint is at a voltage minimum (node) and a current maximum.



F IGURE 8.2 Current, voltage, and impedance along a half-wavelength dipole.

At resonance, the impedance at the feedpoint is purely resistive. There are two contributors to R_{ANT}:

- Ohmic losses R_{loss} that generate nothing but heat when signal is applied. Ohmic losses originate in the small but finite resistance of real conductors and connections at wire junctions and in any shunt leakage paths. In a properly constructed dipole having a length on the order of $\lambda/2$ and located in the clear, losses are usually negligible. Because of the skin effect, R_{loss} is not the same as the DC resistance of the wire as measured with an ohmmeter.

- The radiation resistance R_{rad} of the antenna. This is the value of a fictitious lumped-component resistance that would dissipate the same amount of power that radiates from the antenna.

Example 8.3 Suppose we use as an antenna a large-diameter conductor with negligible ohmic losses at the transmitter frequency. If 1000 W of RF power is applied to the feedpoint, and a current of 3.7 A RMS is measured, what is the radiation resistance?

Solution

$$\begin{aligned} R_{\text{rad}}\ (\text{ohms}) &= \frac{P}{I^2} \\ &= \frac{1000}{(3.7)^2} \qquad (8.4) \\ &= 73\ \Omega \end{aligned}$$

In this example, matching the antenna feedpoint impedance would appear to be simplicity itself because the free-space feedpoint impedance of a simple dipole is about 73 Ω, seemingly a good match to 75-Ω coaxial cable or hardline. Unfortunately, in practice, the earth and any other conducting object or surface changes the input impedance of the antenna, making the 73-Ω feedpoint impedance almost mythical.

Figure 8.3 shows a plot of approximate radiation resistance (R_{rad}) versus height above ground. The radiation resistance varies from less than 10 Ω to around 100 Ω. As we saw in Chap. 7, this is because some of the signal radiated by the dipole strikes the ground directly beneath the antenna and is reflected straight up, where it induces an additional current on the dipole. Because of the time it takes the radiated field to reach the ground and return, the induced field is shifted in phase with respect to the original signal in the dipole. This causes the ratio of voltage to current at the feedpoint to vary with the height of the antenna.

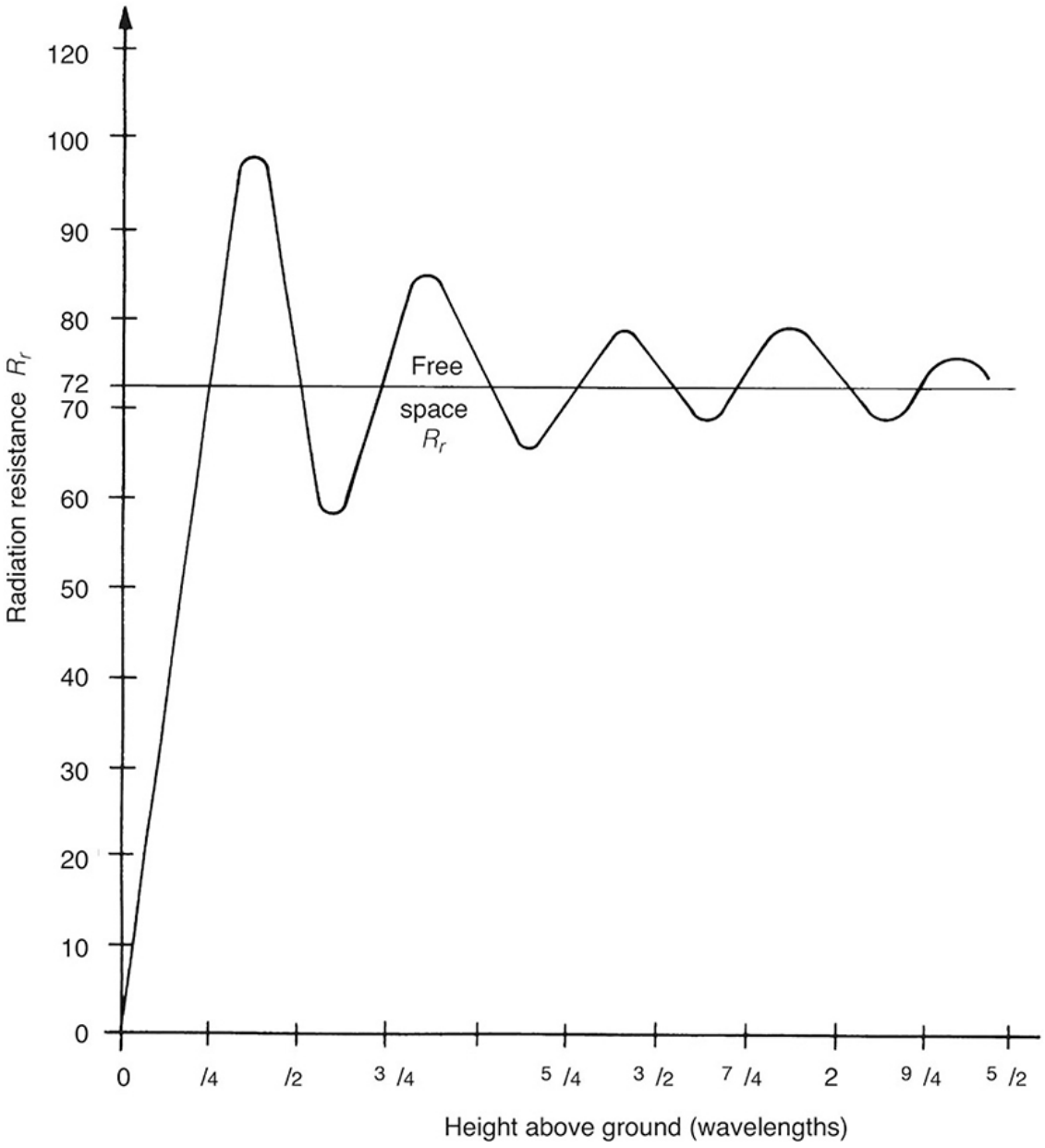


FIGURE 8.3 Radiation resistance versus height above ground.

At heights of many wavelengths, the oscillation of the radiation resistance curve settles down to the free-space impedance (73 Ω) as the effect of the earth becomes negligible. At higher radio frequencies (VHF and above), it becomes practical to install a dipole at a height of many wavelengths. In the 2-m amateur radio band (144–148 MHz),

$\lambda \approx 6.5\ \text{ft}$

(i.e.,
$2\ \mathrm{m} \times 3.28\ \mathrm{ft/m}$
), so a height of many wavelengths is relatively easy to achieve. In the 80-m band (3.5–4.0 MHz),
$\lambda \approx 262\ \mathrm{ft}$
, so the antenna will almost always be electrically close to the ground at lower frequencies. To deal with the variation of radiation resistance caused by the environment around the antenna, several strategies can be used:

- Ignore the problem altogether. In many installations, the height above ground will be such that the radiation resistance will be close enough to represent only a slight impedance mismatch to a standard coaxial cable. For $Z_0 > R_{\mathrm{rad}}$, the voltage standing wave ratio (VSWR) is

  $$\mathrm{VSWR} = \frac{\mathrm{Z_0}}{R_{\mathrm{rad}}} \quad (8.5)$$

  For
  $Z_0 < R_{\mathrm{rad}}$
  , it is the ratio

  $$\mathrm{VSWR} = \frac{R_{\mathrm{rad}}}{Z_0} \quad (8.6)$$

  where
  Z_0
  is the feedline coaxial cable characteristic impedance and
  R_{rad}
  is the radiation resistance of the antenna.
- Consider mounting the antenna at a height near a quarter-wavelength, such that the radiation resistance is around 60 Ω. The resulting VSWR when feeding a 60-Ω antenna with either 52- or 75-Ω standard coaxial cable is

  For 75-Ω coaxial cable:

$$\begin{aligned} \mathrm{VSWR} &= \frac{\mathrm{Z}_0}{R_{\mathrm{rad}}} \\ &= \frac{75}{60} \\ &= 1.25:1 \end{aligned}$$

For 52-Ω coaxial cable:

$$\begin{aligned} \mathrm{VSWR} &= \frac{R_{\mathrm{rad}}}{Z_0} \\ &= \frac{60}{52} \\ &= 1.15:1 \end{aligned}$$

The VSWR created by the mismatch is reasonably good. The same low VSWR results for an antenna fed by 52-Ω coaxial cable when it is slightly less than *λ*/4 above ground and
$R_{\mathrm{rad}} = 45\ \Omega$
. An 80-m dipole strung at a height of 50 ft is a rather good antenna.

- Mount the antenna at whatever height is convenient and use an impedance-matching scheme at or near the feedpoint to reduce the VSWR. Chapter 6 contains information on various suitable broadband impedance-matching methods including Q-sections, coaxial impedance transformers, and broadband RF transformers. Homebrew and commercially available transformers are available to cover most impedance transformation tasks.
- Increase R_{rad} by lengthening the two sides of the dipole equally and use a stub or other matching method to tune out the reactance introduced by the longer elements.
- Use open-wire line between the antenna and an antenna tuning unit (ATU). The ATU converts the complex impedance seen at the transmitter end of the feed-line to

50 Ω (usually the preferred impedance to present to the transmitter). For a given length, open-wire line generally exhibits lower loss than coaxial cable or other solid-dielectric transmission lines—especially in the presence of high VSWR on the line. True open-wire line (as contrasted to receiving type television twin-lead or similar lines) has a higher breakdown voltage specification than most of the commonly available coaxial cables. One of the penalties of a high VSWR on a transmission line is that for a given transmitter power the maximum voltage on the line increases with increasing VSWR. Furthermore, feeding a balanced antenna (the horizontal center-fed dipole) with balanced transmission line helps maintain balance in the antenna, thus minimizing pattern skew and unwanted radiation from (or signal pickup by) the feedline. Open-wire line is usually lighter than coaxial cable, and the center of a dipole supported only at its ends will sag less. Lastly, an ATU located at the transmitter end of the feedline is usually more easily accessed if adjustment for a good match is necessary.

8.1.3 The Dipole Radiation Pattern

In Chap. 5, we developed the pattern of a very short (relative to a wavelength) dipole in free space and compared it to that of a fictitious reference antenna called the isotropic radiator. Compared to the isotropic radiator, the short dipole generates a greater radiated signal broadside to the axis of the wire and has little or no radiation off its ends. The short dipole has 1.5-dB gain relative to an isotropic radiator in its favored direction. The gain came at the expense of radiated signal in other directions. Gain in an antenna is a consequence of the directivity of its radiation pattern (see Sec. 5.8).

For dipoles having lengths of $\lambda/4$ or longer, the radiation efficiency is usually close to 100 percent ($\eta_{\text{rad}} = 1.0$), and gain and directivity are close in value. For some antenna types that we cover later in this book,

η_{rad}
is much smaller than unity. A Beverage wire, for instance, has very low efficiency and terrible gain, yet can have excellent directivity.

Figure 8.4 shows the radiation pattern of a horizontal dipole antenna in free space as seen from different perspectives. In the horizontal plane (Fig. 8.4*a*), when viewed from above, the pattern is a figure eight that exhibits bidirectional gain broadside to the axis of the antenna. The bulk of the RF energy radiates in the direction of the pattern main lobe, with little or no power radiating off the ends of the antenna axis. In the vertical plane, the radiation pattern of the dipole is omnidirectional (Fig. 8.4*b*). When sliced this way the radiation pattern appears circular. When the two patterns are combined in a 3-D rendering, we see the donut-shaped radiation pattern (Fig. 8.4*c*) that helps us visualize the true pattern of an unobstructed dipole in free space.

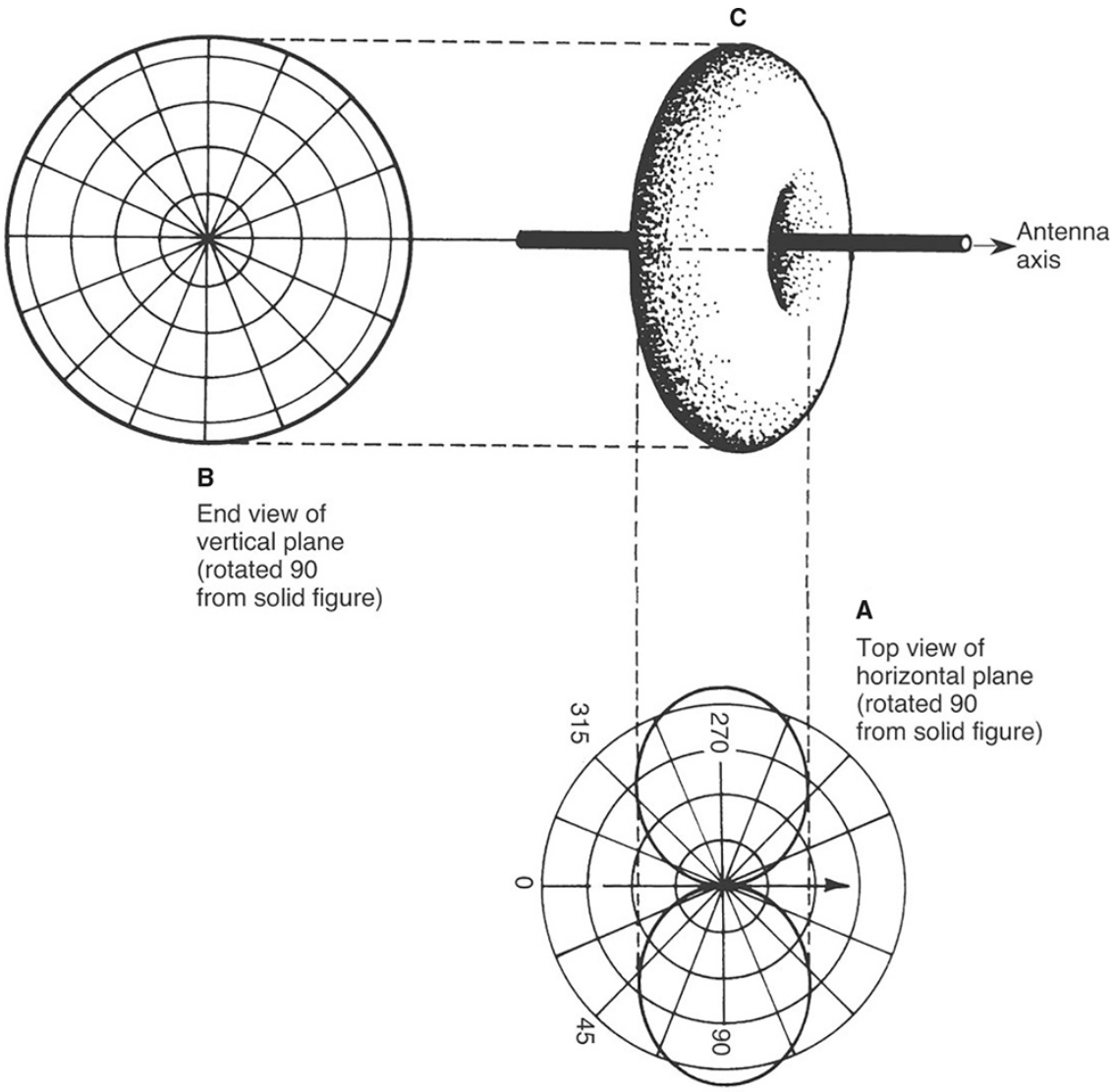


FIGURE 8.4 Idealized dipole radiation pattern.

When a dipole is installed near to the earth's surface or other ground, as is the case for most HF antennas, the antenna pattern is distorted from that of Fig. 8.4. Two effects must be taken into consideration:

- The ground below the antenna causes scattered fields which interact with the antenna and cause small changes in the gain of the antenna as its height is varied.
- Energy radiated downward from the antenna at all angles below horizontal is reflected from the earth's surface back

into the atmosphere. As discussed in Chap. 7, the ground acts like an image antenna and the resulting array effect alters the radiation pattern of the dipole in the vertical plane as measured at distant receiving points. Figures 7.22 and 7.23 at the end of the previous chapter plot typical radiation pattern versus elevation for horizontal and vertical dipole antennas installed within a few wavelengths of the earth's surface.

8.1.4 Tuning the Dipole Antenna

There are generally two tasks to address when building a dipole antenna: achieving resonance and impedance matching the feedline to the antenna. These two issues are closely related. Impedance matching means eliminating reflections from the matching network and antenna that carry unwanted power back to the transmitter. When the antenna is resonant, the reactive part of the input impedance is zero, which makes it easier to impedance match the transmitter and connecting transmission line to the antenna.

Resonance

At an antenna's resonant frequency, the input impedance is purely resistive. The antenna exhibits neither inductive nor capacitive reactance at its feedpoint. To tune a dipole to resonance, we most commonly adjust the length of the antenna. In practice, for wire dipoles, one can initially cut the wire to a longer length than needed, and the wire can be shortened until the antenna is resonant.

Resonance is indicated by a minimum in the VSWR curve. Figure 8.5 shows a graph of VSWR versus frequency for several different cases. Curve A represents a poor VSWR all across the band. The actual value of VSWR can be anything from about 3.5:1 to 40:1 or more, but the cause is nonetheless the same: The antenna is either open or shorted or it is so far off resonance as to appear to be open or shorted to an SWR meter.

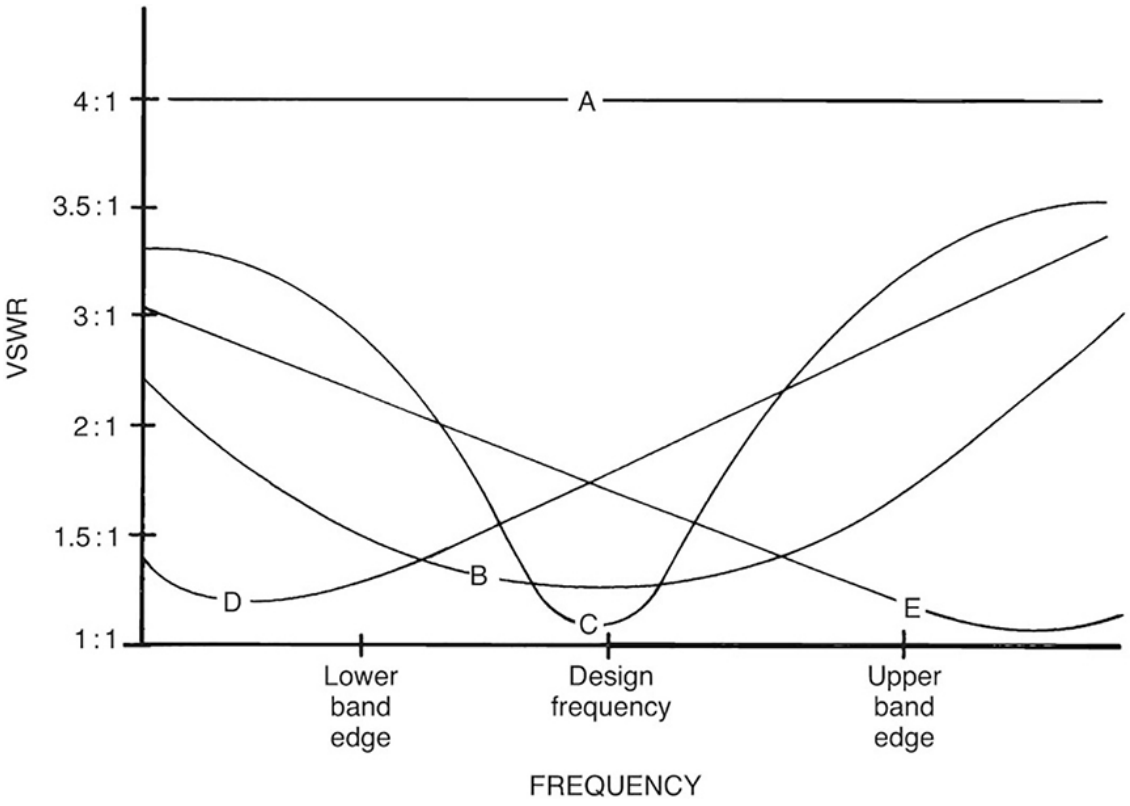


FIGURE 8.5 VSWR versus frequency for several cases.

Curves B and C represent antennas that are resonant within the band of interest. Curve B represents a broadband antenna with a relatively flat curve across the band and does not exhibit excessive VSWR until the frequency is outside of the band. Curve C is also resonant within the band, but this antenna has a higher quality factor (*Q*) than curve B. We might think the broadbanded antenna is always better, but that's true only if the radiation efficiency is reasonably good. Resistive losses tend to broaden an antenna's response curve (the best response curve of all is obtained from a dummy load!), but losses reduce the radiation efficiency and gain of the antenna. When the dipole is used for transmitting purposes, giving up efficiency to obtain a wider SWR bandwidth is usually a poor trade.

Curves D and E are resonant outside the band of interest. In D, the antenna is resonant below the band, implying that the dipole is too long. Shortening the antenna should raise the resonant point to the desired center frequency inside the band. In E, the antenna is resonant above the upper limit of the band, so this antenna is too

short and must be lengthened (hence, our earlier suggestion to leave a little extra wire folded back at each end insulator).

How much to cut depends on how far from the desired frequency the resonant point is found and which band is being used. The latter requirement comes from the fact that the resonant frequency change per unit length varies from one band to another. Here is a simple procedure for estimating the frequency change per unit length:

1. Calculate the length required for the upper band edge.
2. Calculate the length required for the lower band edge.
3. Subtract the first length from the second.
4. Determine the width of the band by subtracting the lower band edge from the upper band edge.
5. Divide the bandwidth obtained in step 4 by the difference in length obtained in step 3; the result is the variation in frequency per unit length.

Example 8.4 Calculate the frequency change per unit of length for 80 m and for 15 m.

Solution For the 80 m band (3.5 to 4.0 MHz in the United States):

1. $L_{\text{UPPER}} = \frac{468}{4} = 117 \text{ ft}$
2. $L_{\text{LOWER}} = \frac{468}{3.5} = 133.7 \text{ ft}$
3. Difference in length: $133.7 \text{ ft } - 117 \text{ ft } = 16.7 \text{ ft}$
4. Frequency difference: $4000 \text{ kHz} - 3500 \text{ kHz } = 500 \text{ kHz}$
5. The frequency change per unit length is $500 \text{kHz}/(16.7 \text{ ft}) = 30 \text{ kHz}/f\text{t or } 2.5 \text{ kHz/in}$

For 15 m ($21.0 – 21.45 \text{ MHz}$ in the United States):

1. $L_{\text{UPPER}} = \frac{468}{21.45} = 21.82 \text{ ft}$
2. $L_{\text{LOWER}} = \frac{468}{21} = 22.29 \text{ ft}$

3. Difference in length: $22.29\ \text{ft} - 21.82\ \text{ft}\ = 0.47\ \text{ft}$
4. Frequency difference: $21,450\ \text{kHz} - 21,000\ \text{kHz}\ = 450\ \text{kHz}$
5. The frequency change per unit length is
   $450/0.47 = 960\ \text{kHz}/ft\ \text{or}\ 80\ \text{kHz/in}$

At 80 m, the frequency change per foot is small, but at 15 m, small changes in a dipole's length can result in large shifts in its resonant frequency.

Using the preceding steps, you can determine approximately how much to add (or subtract) from an antenna under test. If, for example, you design an antenna for 21,390 kHz but find the actual resonant point is 21,150 kHz, the frequency shift required is 21,390 to 21,150 kHz, or 240 kHz. To determine how much to add or subtract (as a first guess):

1. The factor for 15 m is 80 kHz/in.
2. The required frequency shift is 240 kHz.
3. The needed change in length is

$$\Delta L = 240\ \text{kHz}\ \times \frac{1\ \text{in}}{80\text{kHz}} = 3\ \text{in}$$

Each side of the antenna must be changed by half of the length calculated, or 1.5 in. Because we need to raise the resonant frequency, each side should be shortened by 1.5 in. Once the length is correct, as determined by new VSWR measurements, the connections at the center insulator are soldered and made permanent, and the antenna is hoisted back up to its original height.

Impedance Matching

Once the antenna is tuned to be resonant at the desired operating frequency, we must match the antenna's input impedance to the characteristic impedance of the feeding transmission line. The tuned antenna has an SWR curve with a minimum at the desired frequency, but impedance matching brings the minimum value of the SWR close to unity (or 0 dB).

For many applications, it is adequate to keep the VSWR below 2.0:1 across the transmitting frequency range of interest. At frequencies below a few tens of MHz, where sky noise is strong, it is unlikely that VSWRs modestly greater than 2.0:1 are of any consequence when using the antenna for receiving purposes. Most transceivers and power amplifiers are capable of delivering close to rated output power to loads with a VSWR of 2.0:1 or less without incurring any damage.

If the SWR realized after tuning the antenna is inadequate, further improvement can be achieved by:

- Adding an ATU between the antenna and the transmitter.
- Using a stub matching or lumped component matching network.

Matching an antenna to a strict SWR requirement often requires an antenna analyzer or vector network analyzer (VNA; Chap. 30) to provide knowledge of the resistive and reactive parts of the antenna impedance before and after adjustments to the antenna matching network.

8.2 Modified Dipoles

Thus far, the dipoles covered in this chapter have been classical half-wave dipoles, in which a thin half-wavelength conductor is fed at its center and each arm of the dipole is connected to one side of a two-conductor transmission line. This antenna is typically installed horizontally at a half-wavelength or more above the earth's surface. If the antenna is made of wire, it is usually suspended under tension from a high support at each end. If the dipole is made from rigid metal tubing or rod stock—most commonly, aluminum— it is often supported at its center, as we shall see in Chap. 14. On the lower HF bands, where available space and suitably tall supports are often in short supply, ideal dipole installations are often hard to implement. In such cases, the modified dipole-type antennas described in this section can be used. Some of these antennas are comparable to the

horizontal dipole, while others sacrifice performance for other advantages.

8.2.1 Inverted-V Dipole

A popular substitute for a dipole on lower HF bands is the inverted-V or inverted vee dipole. This is a half-wave dipole whose ends are significantly lower than its center. Like a conventional dipole, it is usually fed in the center with a balanced or unbalanced transmission line. The inverted vee is generally better matched to the 50 Ω impedance of most transceivers and common coax lines than the nominal 72 Ω of a linear dipole.

The inverted vee is often the antenna of choice when only a single high support is available. Because the inverted vee is supported at its center, heavy feedlines, such as those made of RG-8 or RG-213 coaxial cable, can be suspended from the same center support, eliminating downward pull on the center of the antenna. Another mechanical advantage of the inverted vee is that the antenna wires do not support the weight of the feedline or the center insulator and there is no need for substantial tension in the antenna wires themselves. The inverted vee can in many cases be constructed from #18 or smaller gauge wire.

Figure 8.6 shows a typical inverted vee installation. The operating frequency and the height of the center support determine the approximate height above ground of the ends. The ends can come almost to the ground if provision for preventing people and animals from touching the antenna is included. The angle *a* can be anything convenient, provided that

$$\alpha > 90 \text{ degrees}$$

; typically, *a* is between 90 and 120 degrees. On the 80 m band, the minimum practical height for the center support is about 35 ft for

$$\alpha = 120 \text{ degrees}$$

and the overall length of the antenna footprint is about 110 ft. If space is at a premium, reducing the included angle to 90 degrees leads to a minimum center support height of nearly 50 ft with a footprint just over 90 ft. An inverted vee for 40 m could have

minimum dimensions in height and footprint of approximately 50 percent to 60 percent of these figures.

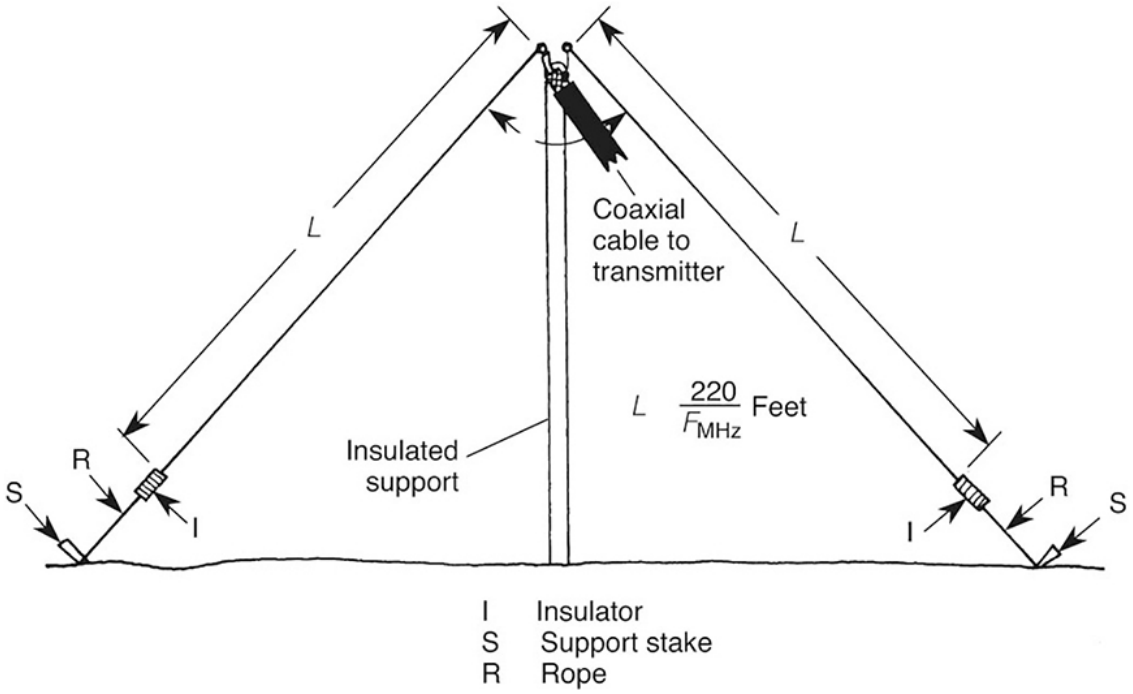


FIGURE 8.6 Inverted-vee dipole.

Sloping the antenna elements down from the horizontal to an angle, as shown in Fig. 8.6, lowers the resonant frequency for a given length of wire. The wire lengths may need to be cut slightly shorter for any given frequency than for a horizontal dipole. A rule of thumb is to make the antenna about 6 percent shorter than a dipole for the same frequency. The initial cut of the antenna element lengths (each quarter-wavelength) is then

$$L(\text{feet}) = \frac{220}{f(\text{MHz})} \quad (8.7)$$

After the length *L* is calculated, the actual length is found from the cut-and-try method used to tune the dipole in the previous section. Bending the elements downward changes the feedpoint impedance of the antenna and narrows its bandwidth, so impedance matching or an antenna tuner (ATU) may be needed.

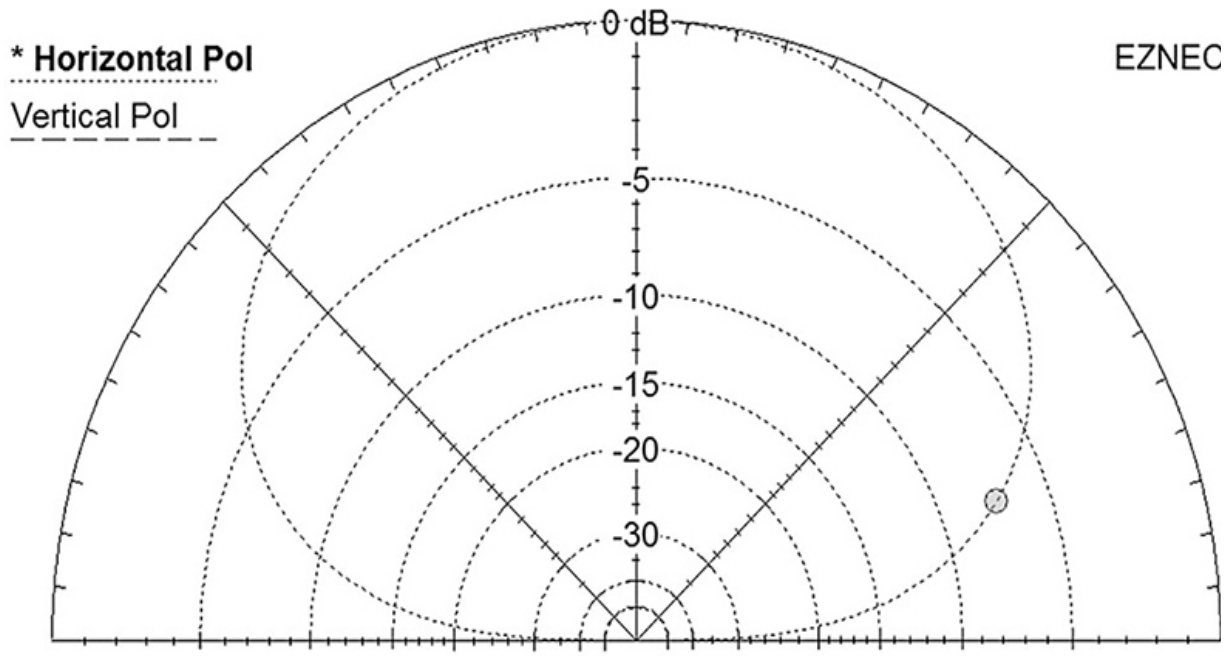
The inverted vee is a compromise antenna. We gain mechanical advantages (previously described) while incurring electromagnetic disadvantages:

- Part of the radiation from each leg of the inverted vee is horizontally polarized and part is vertically polarized. For $\alpha = 90$ degrees, these two components are roughly equal; simple trigonometry then tells us each is about 70 percent of the magnitude of the original horizontally polarized E-field from a horizontal dipole.
- Remembering that the radiation field far from the antenna is the sum of the radiation from many small segments of wire making up the antenna, the average height of the high-current portions of the antenna is less than that of a comparable dipole whose entire length is at the same height as the inverted vee's center. Thus, the horizontally polarized radiation from the inverted vee tends to favor higher takeoff angles, compared to a horizontal dipole at the same height as the center of the inverted vee.

Since some of the radiation is now vertically polarized, one might think there will be an increased amount of low takeoff angle radiation, increasing potential range. However, just as the currents on opposite sides of a balanced two-wire transmission line feeding a dipole are 180 degrees out of phase with each other (thus minimizing radiation from the feedline at distant points by cancellation of the fields from the two sides), the vertical component of the inverted vee's radiation on one side of the center insulator is out of phase with the vertical component on the other side. Since the average spacing of the high-current portions of the two sides of the inverted-vee is wider than that of a parallel-wire transmission line—perhaps $\lambda/8$—there is incomplete cancellation and some vertically oriented energy is radiated. But the direction of maximum radiation is not broadside to the wire as it is for the horizontal component. Rather, the vertically polarized pattern is typically maximum along the wire axis, at right angles to the optimum direction of horizontal radiation, and over

average ground is about 4 dB lower in amplitude at a 20-degree elevation angle than a simple $\lambda/4$ vertical would be.

Figure 8.7 shows the modeled patterns for the horizontally and vertically polarized E-field components of an 80-m inverted vee with a 90-degree included angle and a center support height of 50 ft over average ground. Note the 90-degree difference in the compass headings of the horizontal and vertical E-field lobes.

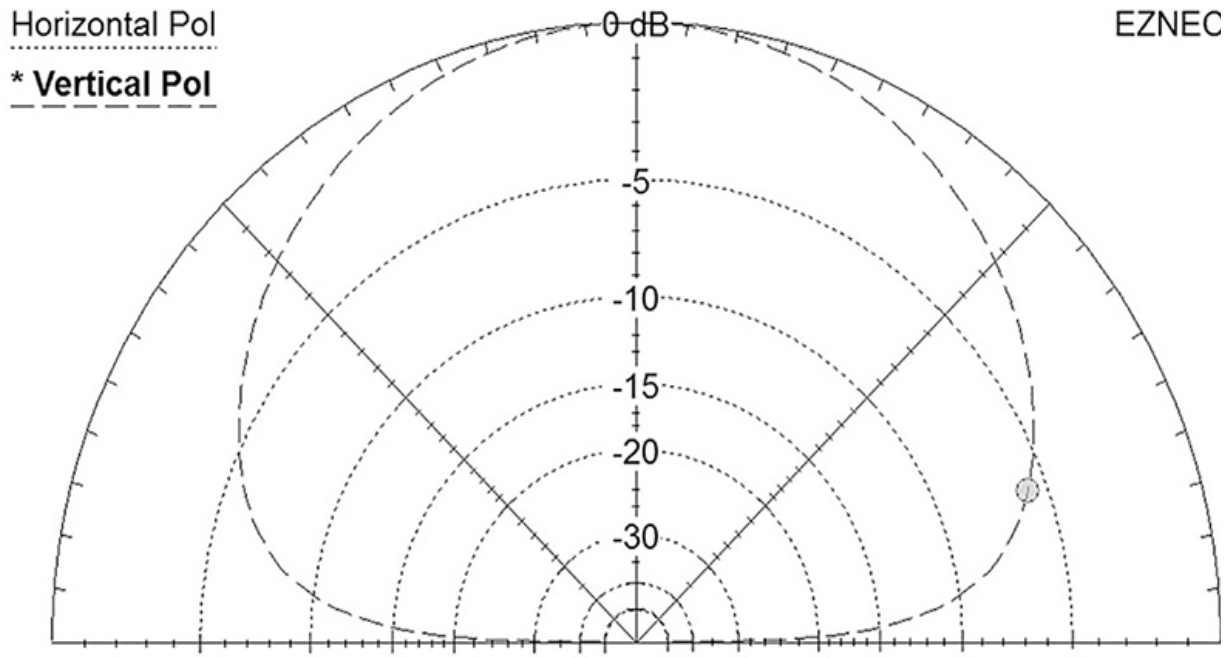


Elevation Plot
Azimuth Angle 0.0 deg.
Outer Ring 3.66 dBi

Slice Max Gain 3.66 dBi @ Elev Angle = 90.0 deg.
Beamwidth 105.8 deg.; -3dB @ 37.1, 142.9 deg.
Sidelobe Gain < -100 dBi
Front/Sidelobe > 100 dB

3.55 MH

Cursor Elev 20.0 deg.
Gain -3.59 dBi
-7.25 dBn



Elevation Plot
Azimuth Angle 90.0 deg.
Outer Ring 3.66 dBi

Slice Max Gain 3.66 dBi @ Elev Angle = 90.0 deg.
Beamwidth 100.8 deg.; -3dB @ 39.6, 140.4 deg.
Sidelobe Gain < -100 dBi
Front/Sidelobe > 100 dB

3.55 MH

Cursor Elev 20.0 deg.
Gain -2.15 dBi
-5.81 dBn

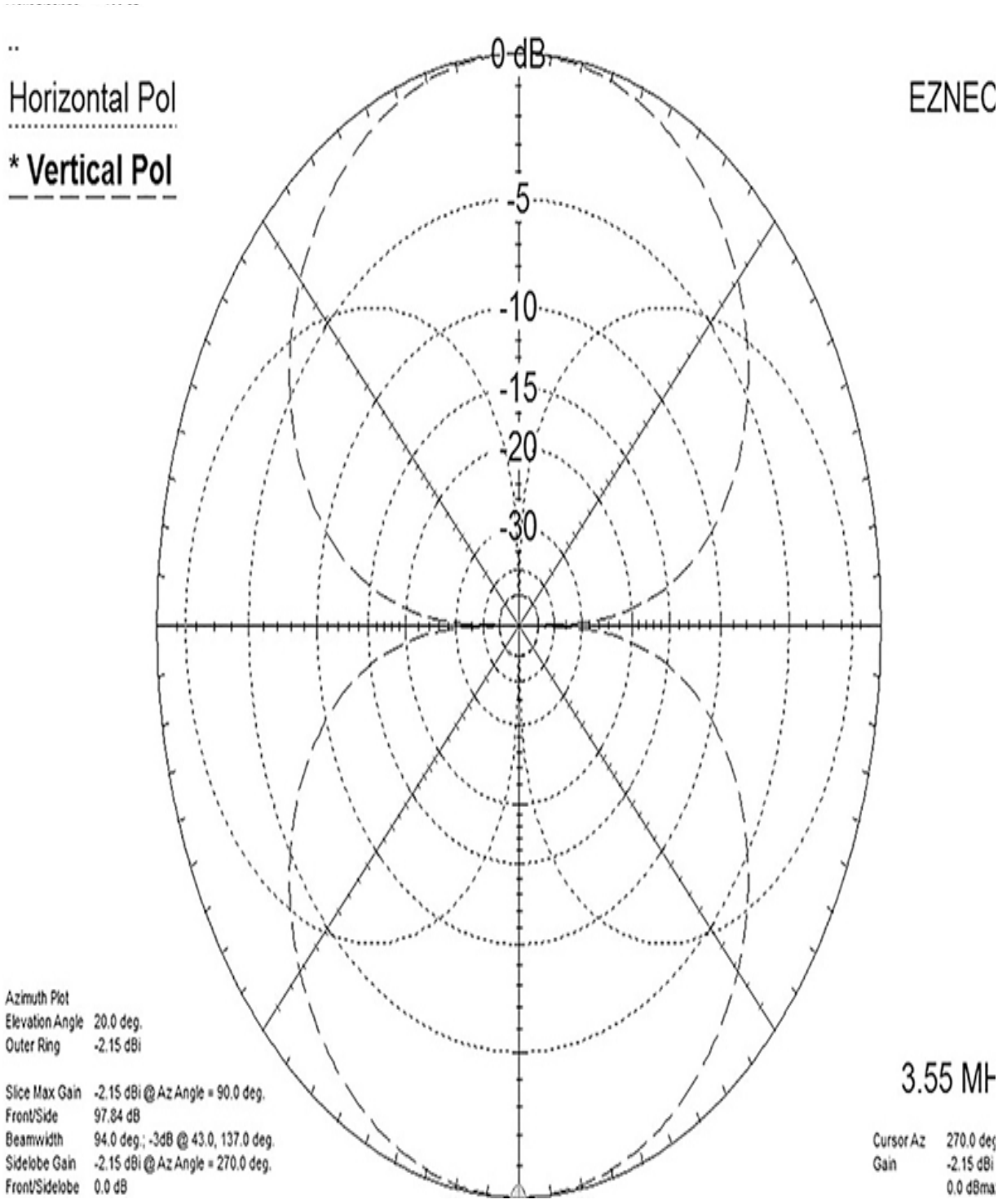


FIGURE 8.7 Inverted vee radiation patterns.

8.2.2 Sloping Dipole ("Sloper")

The sloping dipole (Fig. 8.8) is popular with operators who need a low angle of radiation and have at least one tall support. This antenna, informally referred to as a sloper, is a half-wavelength dipole that is erected with one end at the top of a support and the other end close to the ground so as to make a 45-degree (more or less) angle with the support. Like a horizontal dipole, it is usually fed in the center by coaxial cable or two-wire balanced transmission line. When fed with a coaxial (unbalanced) line, common practice is to connect the outer conductor or shield of the line to the lower half of the antenna, but there is little provable basis for doing so. The antenna is so unbalanced with respect to nearby ground that how the coaxial cable is connected is essentially immaterial.

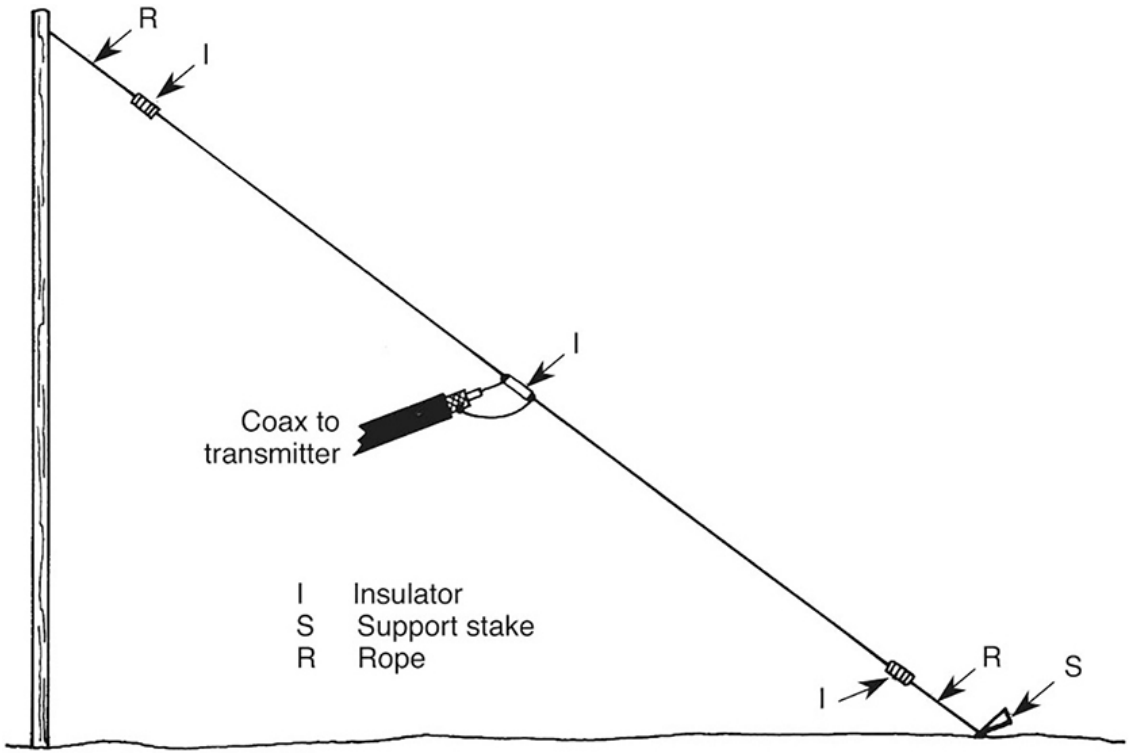
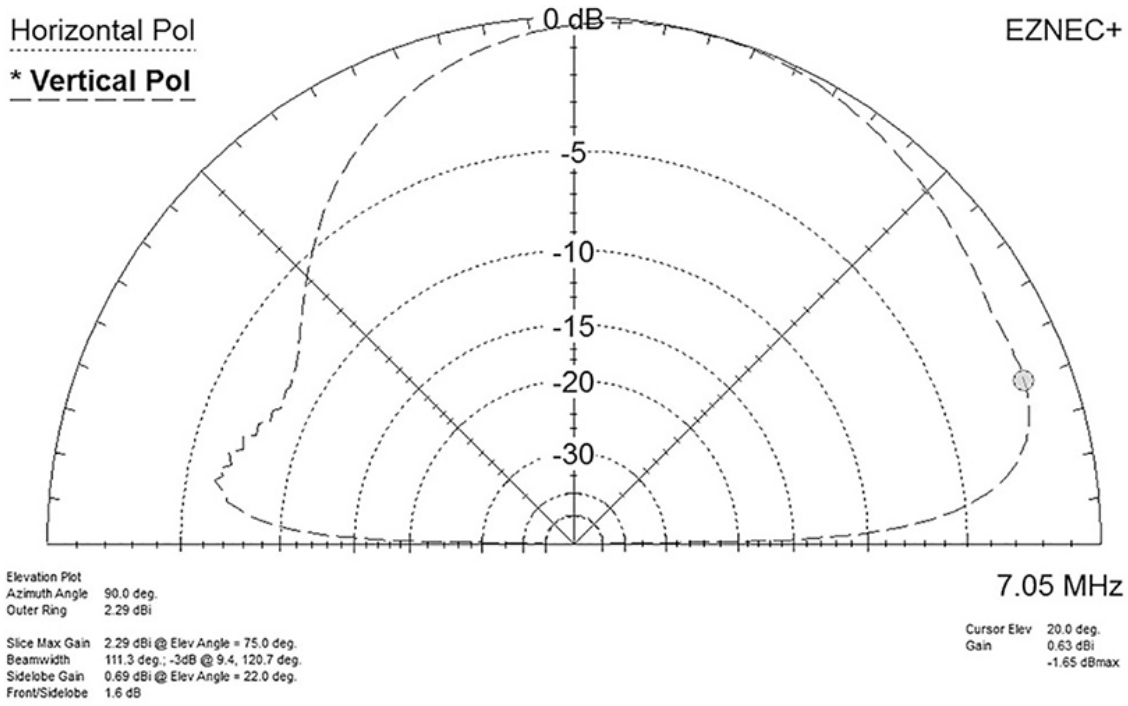


FIGURE 8.8 Sloping dipole ("sloper").

Like the inverted-vee, a sloping dipole sacrifices the high-current portions of the antenna to lower heights, but unlike the inverted-vee, the vertically polarized contributions to the radiated field from the two halves of the antenna are in phase because the two halves of the antenna are not folded back toward each other. The vertical

components of the E-field are offset spatially, with the result that the antenna exhibits a small amount of gain in the direction of the lower end. This forward gain is accompanied by a modest reduction at the rear (i.e., in the direction of the higher end), so the antenna exhibits a small but noticeable front-to-back ratio, as shown in Fig. 8.9.



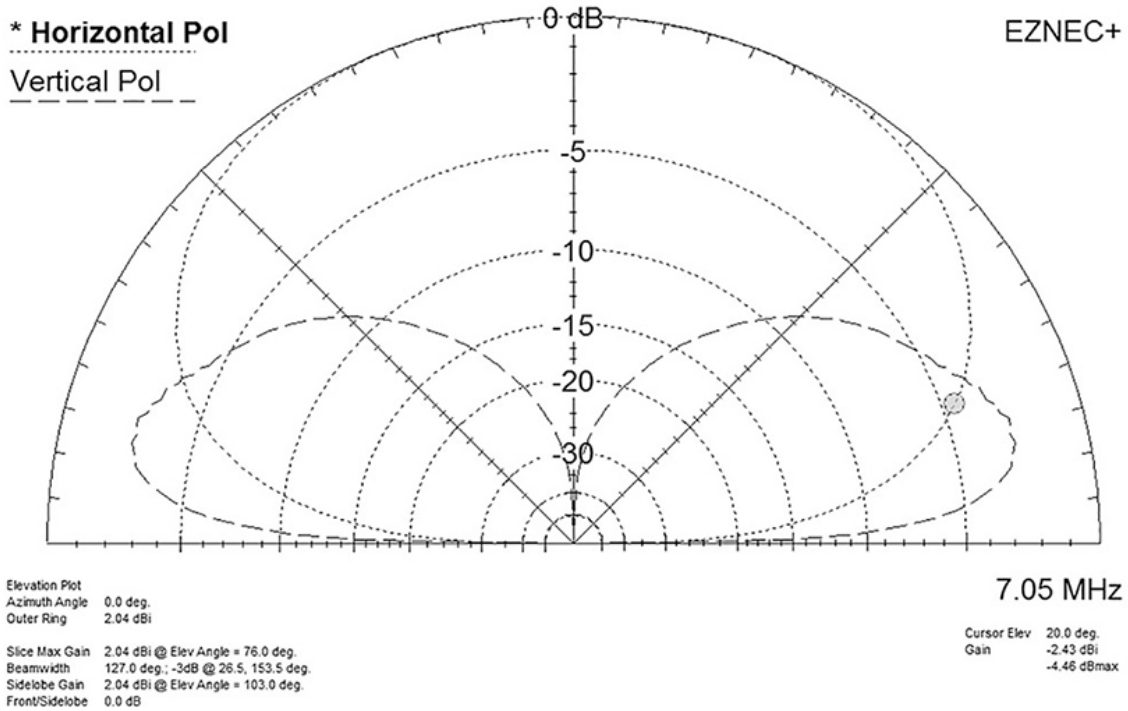


FIGURE 8.9 Sloping dipole radiation patterns.

Space permitting, some operators arrange four sloping dipoles from the same mast in different directions around the compass (Fig. 8.10). A single four-position coaxial switch midway up the center support facilitates switching a directional beam around the compass to favor four compass headings. If the coaxial cables to the four-position switch are of identical lengths chosen based on whether the remote switch short-circuits the unused cables or leaves them open, additional front-to-back rejection on the order of 10 dB from the three idle slopers is possible.

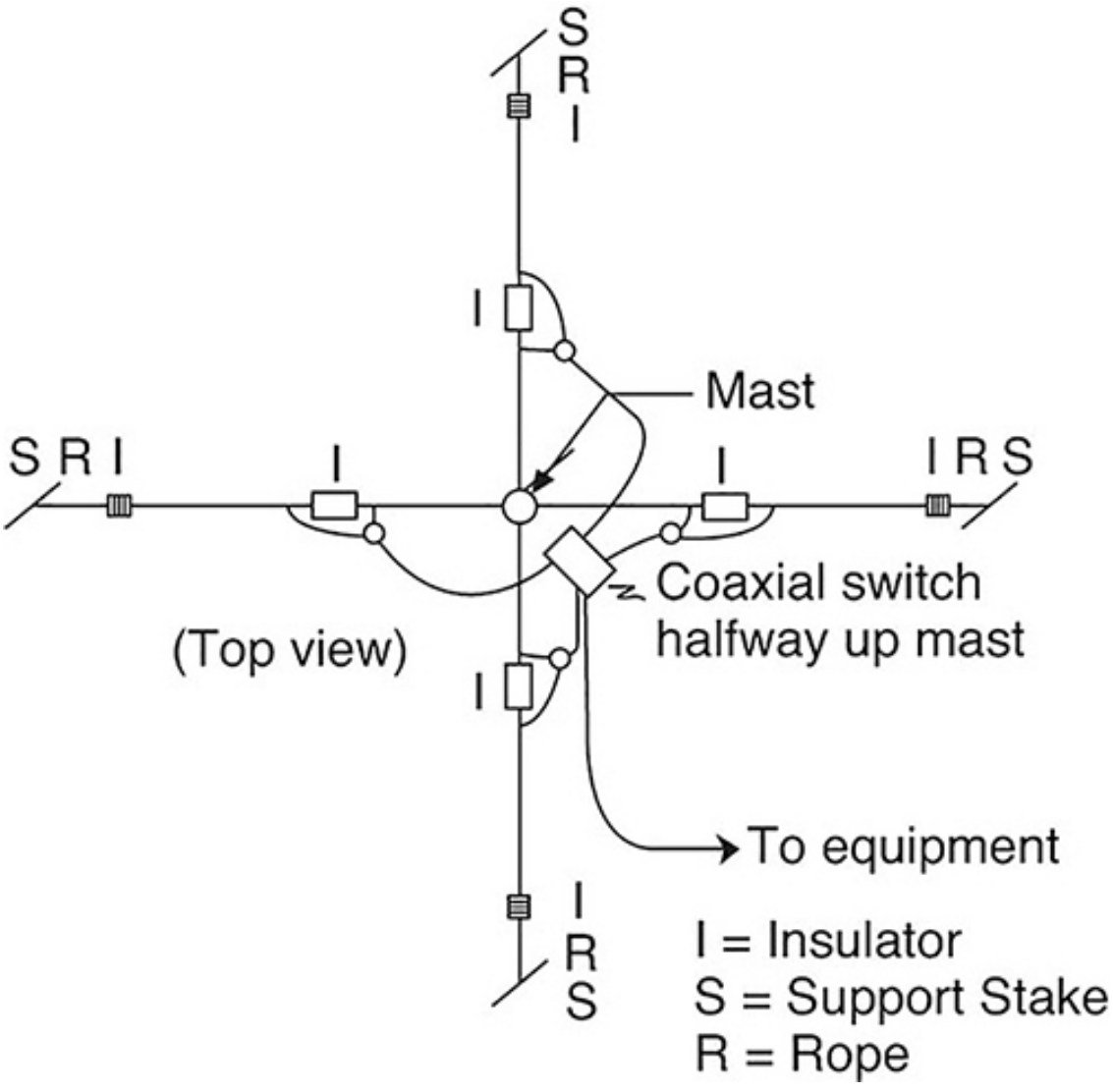


FIGURE 8.10 Directional antenna made of four slopers (top view).

8.2.3 Broadband Dipoles

A basic dipole may have a fractional bandwidth relative to the center operating frequency of around 5 to 10 percent. If this is insufficient for the application, by making various modifications to the geometry the bandwidth of the dipole antenna can be increased.

One simple approach is to increase the cross-sectional area of the dipole arms. When mechanical considerations permit, this suggests the use of aluminum tubing instead of copper wire for the antenna radiator. On 14 MHz and above, that is often a preferred solution. At those frequencies, dipoles constructed from aluminum tubing typically require support only at their centers, so a dipole made of aluminum tubing may be relatively inexpensive when compared to the cost of wire and supporting structures. As frequency decreases, the weight becomes greater because the tubing is both longer and (for structural strength) of greater diameter. On 40 m and below, dipoles of aluminum tubing require special supporting mechanisms and are generally not cost-effective.

Coverage of the entire 80-m band is a significant problem, especially for newer transmitting equipment, because the band is 500 kHz wide, which corresponds to a bandwidth of 13 percent if the antenna is tuned for 3.75 MHz. Solid-state exciters and transceivers can include fold-back circuitry that shuts down the amplifier stages or greatly reduces the output power level in the presence of high VSWR. High-power amplifiers, whether employing vacuum tubes or solid-state devices, include VSWR-driven shutdown protection circuits that often require the operator to resolve the problem and manually reset the amplifier controls before they can be used again. To overcome mismatch and improve bandwidth, an ATU can be inserted between the output of the transmitter and the feedline.

When increasing the radius of the dipole arms or adding an ATU are insufficient or impractical, bandwidth can be increased by modifying the dipole in other ways. Here are four approaches to widening the inherent bandwidth of a center-fed dipole antenna: folded dipole, bowtie dipole, cage dipole, and fan dipole.

Folded Dipole

Figure 8.11 shows the folded dipole antenna. This antenna consists of two or more half-wavelength conductors of identical diameter shorted together at both ends and fed in the middle of one of the conductors. Except at the ends, the two conductors are held a few inches apart (at HF, closer for microwave frequencies) by insulated spacers. The two-wire folded dipole can be made with 300-Ω television antenna twin-lead transmission line. Because the free-space feedpoint impedance is nearly 300 Ω, or four times that of a conventional dipole, the antenna is a reasonably good match (depending on the height of the antenna above ground) for the same twin-lead used as the transmission line. Such a dipole will exhibit roughly a 50 percent improvement in 2:1 VWSR bandwidth compared to a single-wire dipole. On 80 m a folded dipole resonant at 3.750 MHz will cover from 3.6 to 3.9 MHz with an SWR of 2:1 or lower.

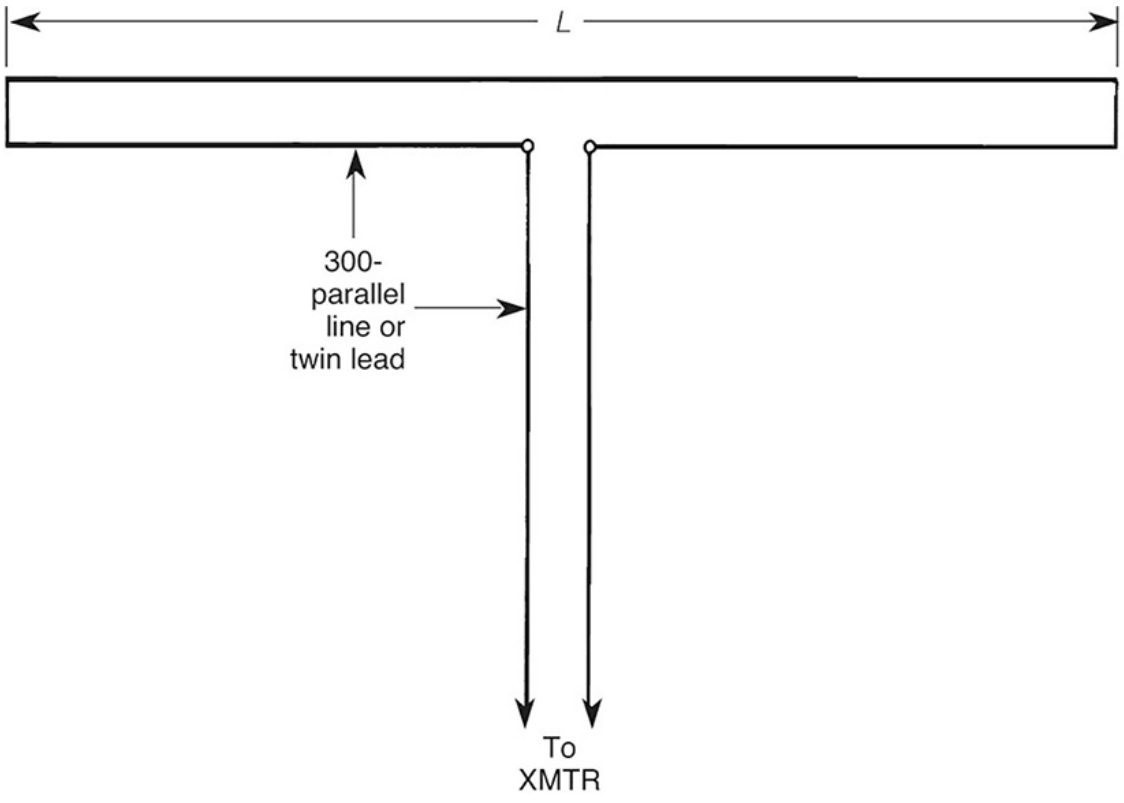


FIGURE 8.11 Folded dipole fed with 300-Ω line.

Why is the input impedance four times the input impedance of a conventional dipole when one of the two wires is fed? Radiation from

the current on the unfed wire adds to the radiation from the fed wire, so that for a given input current the radiated field for the folded dipole is double that of a dipole. Since the radiated power goes as the square of the field, the radiated power is four times higher. This leads to

$$\frac{1}{2}|I|^2 R_{\text{dipole}} = 4\frac{1}{2}|I|^2 R_{\text{folded}} \;\Rightarrow\; R_{\text{folded}} = 4R_{\text{dipole}} \simeq 290\ \Omega \quad (8.8)$$

The impedance step-up ratio for a two-wire folded dipole is 4:1 if the two conductors are of equal diameter. When the diameters are different, the impedance step-up increases by the ratio of the unfed wire diameter to the fed wire diameter. A folded dipole can be designed to provide a specific feedpoint impedance to the transmitter and transmission line, within limits, by making one of the two wires larger than the other.

In many installations, the best feedline for the folded dipole will be 300-Ω twin-lead or, better yet, open-wire line connected to a balanced wire ATU at the transmitter end. Although coaxial cable can be connected to the antenna feedpoint directly, it will experience a roughly 6:1 VSWR over its entire length. Another popular alternative is to use a 4:1 balun transformer at the feedpoint (Fig. 8.12). This arrangement provides a reasonable match between the folded dipole and either 52- or 75-Ω coaxial transmission line.

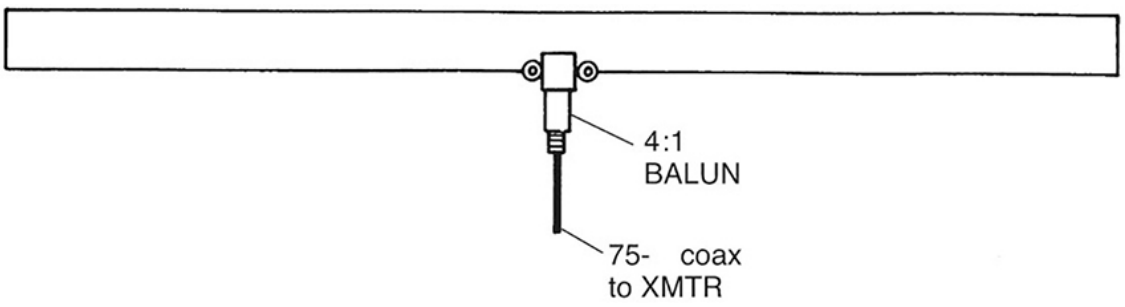


FIGURE 8.12 Folded dipole fed with coaxial cable.

Adding a third unfed wire on the opposite side of the fed wire from the first unfed wire creates the three-wire dipole. A three-wire center-fed dipole offers twice the 2:1 SWR bandwidth of a conventional

dipole. The impedance step-up ratio for conductors of equal diameter is 9:1, so this antenna makes an excellent mate to 600-Ω open-wire line. As before, all three wires are tied together electrically at both ends of the antenna. For both the folded dipole and the three-wire dipole, the tradeoffs when compared to a single-wire dipole are increased bandwidth on the design band versus increased weight and loss of multiband capability.

Bow Tie Dipole

Another method for broadbanding a half-wave dipole is to use two dipoles fed from the same transmission line and arranged to form a bow tie shape, as shown in Fig. 8.13. The use of two identical dipole elements on each side of the transmission line increases the apparent conductor cross-sectional area so that the antenna has a slightly improved length/diameter ratio.

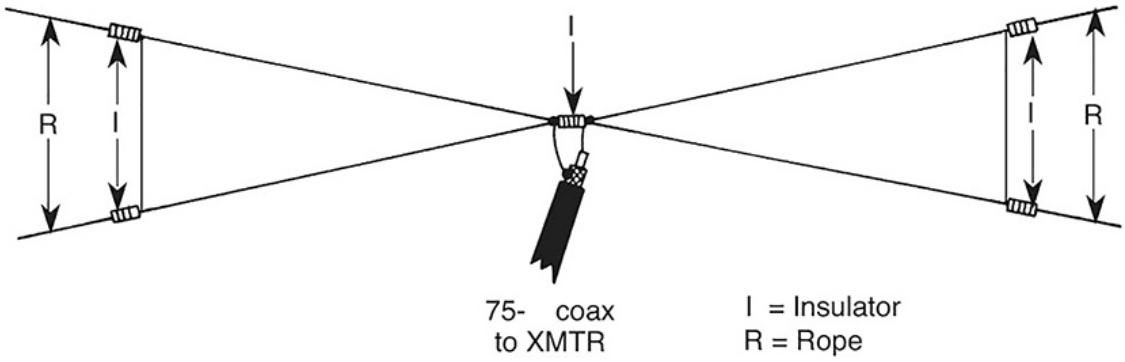


FIGURE 8.13 Bow tie dipole.

The bow tie dipole became the basis for the earliest television receiver antennas. The bow tie attained popularity in amateur circles during the 1950s as the Wonder Bar antenna for 10 m. It still finds use, but its popularity has faded in the intervening period. At microwave frequencies and higher, printed bow tie antennas are used on PCBs for wideband applications.

Cage Dipole

The cage dipole (Fig. 8.14) is similar in concept, if not construction, to both the bow tie and the three-wire dipole. The idea is to connect

several parallel dipoles to the same transmission line in an effort to increase the apparent cross-sectional area. Insulated spreaders to keep the wires separated are typically made from plexiglass or ceramic. They can also be made from materials such as wood treated with varnish or polyurethane or from any other lightweight insulating material that has been rendered impervious to moisture. The spreader disks are held in place with wire jumpers (see inset to Fig. 8.14) that are soldered to or wrapped around the main element wires.

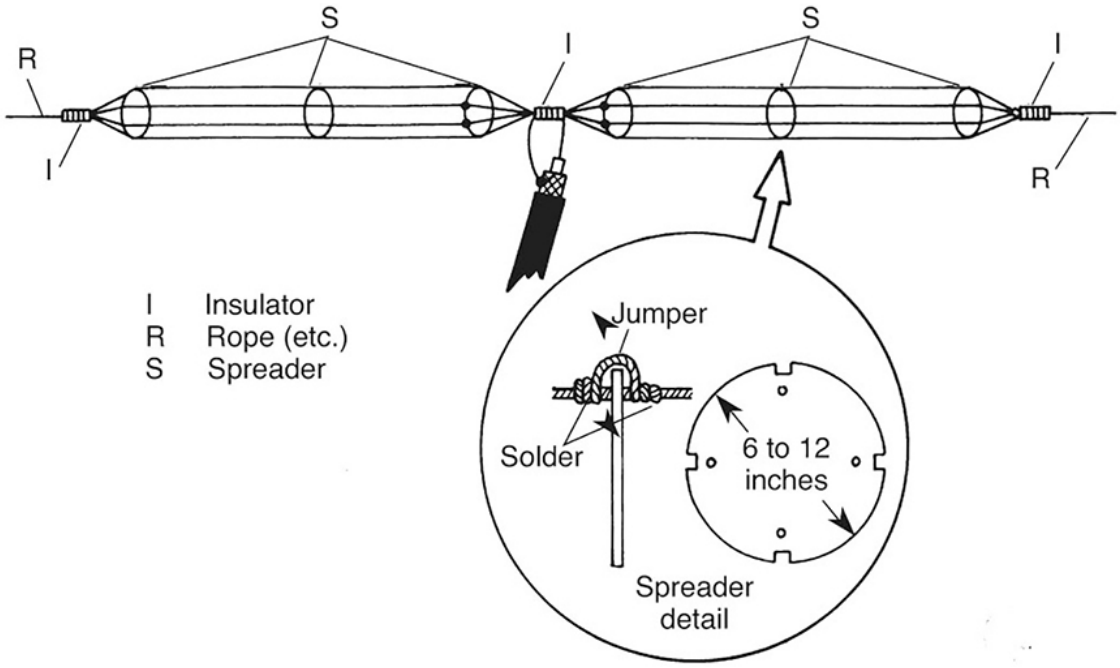


FIGURE 8.14 Cage dipole.

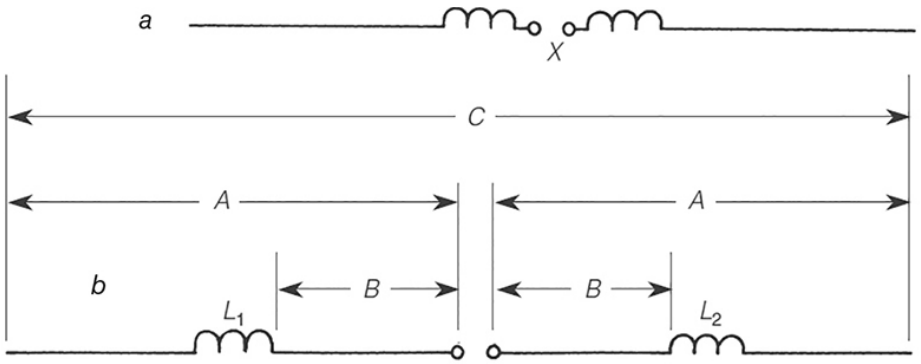
Fan Dipole

An antenna used by 80-m and 10-m amateurs is to parallel two or more dipoles cut for different parts of the same band, fanning the individual wires on each side of the common center insulator out with distance from the center. This "stagger tuning" method sends most of the RF to the shorter dipole at the upper end of the band and to the longer dipole at the lower end of the band. The result is to somewhat flatten variations in VSWR across the entire band. Modeling can be used to optimize the wire lengths. If three or four separate half-

wavelength elements are employed, it is possible to overlap even narrower sections of the band in order to create an even flatter VSWR characteristic. Unlike the broadbanding schemes previously described, the wires at each end are not connected to each other.

8.2.4 Shortened Coil-Loaded Dipoles

At lower frequencies, the half-wavelength dipole is often too long for some installations. The solution for many operators is to use a coil-loaded shortened dipole such as one of those shown in Fig. 8.15. In the absence of loading, a shortened dipole (i.e., one that is appreciably less than a half-wavelength) exhibits capacitive reactance at its feedpoint. This reactance can be canceled with an inductance placed at almost any point along the radiator; in Fig. 8.15*a* and *b*, coils are placed at the feedpoint and at 50 percent of the element length.



$L_1 = L_2 = L$

$C = 2A$

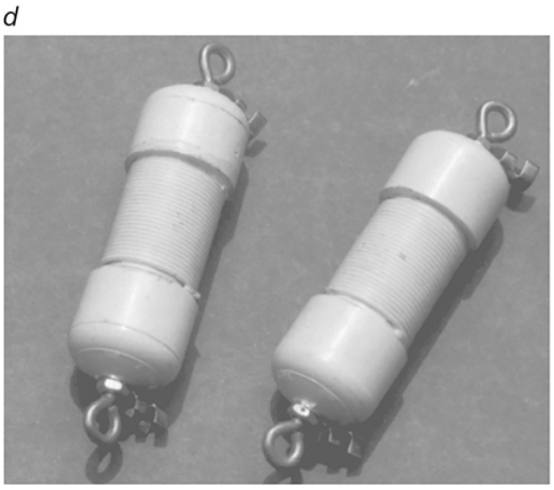
$A = 1/2C$

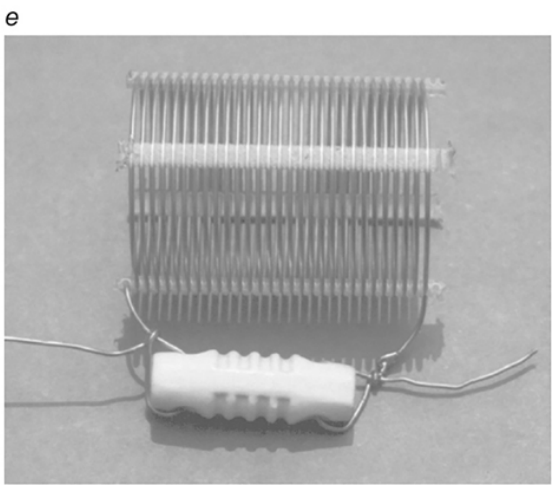
c

$$C = \frac{468M}{F_{MHz}}$$

$0 \leq M \leq 1$

Percent of half-wavelength	Coils at feedpoint (Ω)	Coils at middle of radiators (Ω)
20	1800	2800
30	950	1800
40	700	1800
50	500	1300
60	360	950
70	260	700
80	160	500
90	75	160
95	38	80
98	15	30





F IGURE 8.15 Shortened or loaded dipole. (*a*) Inductors at feedpoint. (*b*) Inductors midway along elements. (*c*) Chart of reactances for coils. (*d*) Commercially available coils. (*e*) Homemade coil based on commercial coil stock.

Figure 8.15*c* is a table of inductive reactances (in ohms) as a function of the shortened radiator's length, expressed as a

percentage of a half-wavelength. The maximum allowable percentage is dictated by the specific installation, but the general rule is to pick the largest percentage that will fit within the available space.

Loading coils are most effective (in terms of how much inductance is necessary to "replace" a specified length of the dipole) at the center (or feedpoint) of the dipole. But for overall antenna efficiency, coils in the middle of both sides are preferable to coils at the feedpoint. Coils near the feedpoint carry more current, increasing loss, and they replace more of the higher radiating portions of the antenna. The shorter the total wire length of the dipole is, the smaller the overall efficiency of the antenna structure.

Example 8.5 Suppose you have about 40 ft of backyard available for a 40-m antenna that normally needs about 65 ft for a half-wavelength. What value of inductor do you need?

Solution Because 39 ft is 60 percent of 65 ft, you could use this value as the design point for this antenna. From the table, a shortened dipole that is 60 percent of a full
$\lambda/2$
dipole's length requires an additional inductive reactance of 950 Ω with the loading coils at the midpoint of each radiator element to allow the feedline to see a purely resistive feedpoint impedance. Rearrange the standard inductive reactance equation (
$X_L = 6.28\,FL$
) to the form

$$L(\mu\text{H}) = \frac{X_L}{6.28 f(\text{MHz})} \quad (8.9)$$

where *L* is the required inductance in microhenrys, *f* is the frequency in MHz, and
X_L
is the inductive reactance obtained from the table in Fig. 8.15*c*. If the antenna is cut for 7.150 MHz, the calculation is as follows:

$$L(\mu\text{H}) = \frac{X_L}{6.28F(\text{MHz})} = \frac{950}{6.28 \cdot 7.15} \simeq 20.7\ \mu\text{H}$$

The inductance calculated here is approximate; it might have to be altered by cut-and-try methods.

The shorter the dipole is, the greater the amount of loading required and the narrower the bandwidth of the combined assembly. The loaded dipole is a very sharply tuned antenna. Because of this, one must either confine operation to one segment of the band or provide an antenna tuner to compensate for the sharpness of the bandwidth characteristic. Even with a tuner, efficiency drops markedly far from resonance.

Figure 8.15*d* and *e* shows two methods for constructing a coil-loaded dipole antenna. Figure 8.15*d* shows a pair of commercially available loading coils especially designed for this purpose. The ones shown are for 40 m, but other models are also available. The inductor shown in Fig. 8.15*e* is a section of commercial coil stock connected to a conventional end insulator or center insulator. No structural stress is absorbed by the coil—all forces are applied to the insulator, which is designed for normal wire dipole tensions.

Inductance values for other lengths of antennas can be approximated from the graph in Fig. 8.16. This graph contains three curves for coil-loaded short dipoles that are 10, 50, and 90 percent of the normal half-wavelength size. Find the proposed location of the coil, as a percentage of the wire element length, along the horizontal axis. The intersection of a vertical line from that point and one of the three curves yields the required inductive reactance (see along vertical axis). Inductances for other overall lengths can be rough-guessed by interpolating between the three available curves and validated by cut and try.

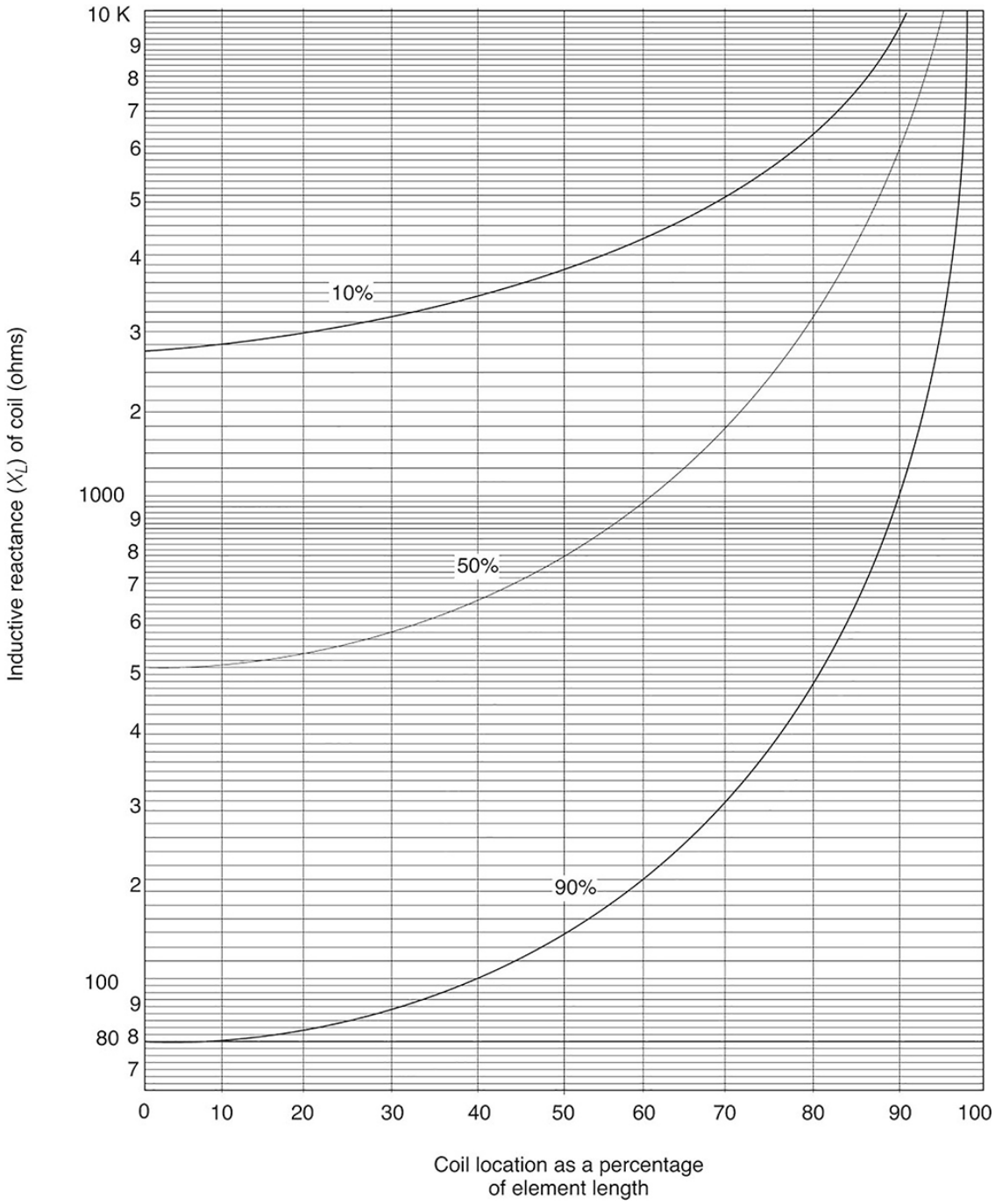


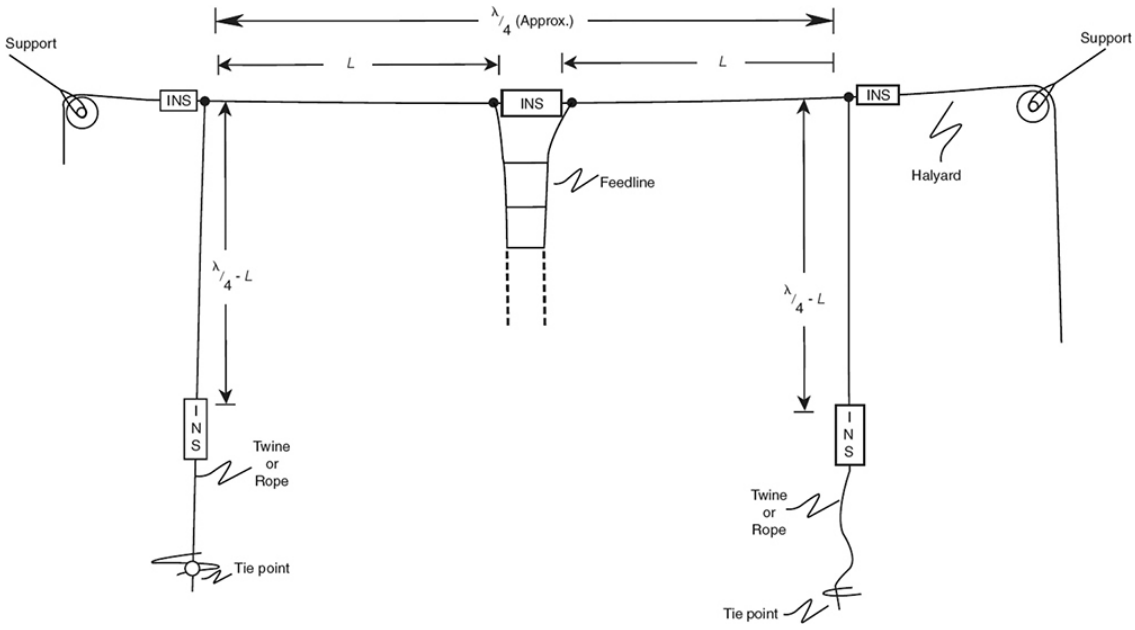
FIGURE 8.16 Inductive reactance versus coil location for shortened inductance-loaded dipole.

8.2.5 Bent Dipole

For a dipole antenna, most of the radiated field originates from the center half of its total length. The outer $\lambda/8$ section on each end of

the dipole is important for establishing resonance but contributes little to the total signal received at a distant point. The bent dipole allows the use of a full
$\lambda/2$
dipole in limited space by bending the outer portions of the dipole. The simplest approach is to let them dangle, perhaps with a slight weight or pull-down cord at each end to keep them from swinging in strong winds, but the ends can also be bent horizontally. In practice, the ends can be bent just about any way you want. Just keep them away from the horizontal center portion of the dipole and avoid doubling back at any points along the wires.

Figure 8.17 shows a bent dipole with dangling ends. The main lobe of the horizontally polarized field is down about 0.5 dB in free space and 0.6 dB when the center portion is 0.3 λ above typical ground. The feedpoint impedance will be lower than for the conventional horizontal
$\lambda/2$
dipole but still a decent match to 52-Ω coaxial cable.



F IGURE 8.17 Bent dipole.

8.2.6 Extended Double Zepp Antenna

As the length of each side of a half-wave dipole is lengthened beyond $\lambda/4$, the gain (or directivity) of the main lobe increases. This continues until each side is $5\lambda/8$ long, at which point current reversal in adjacent $\lambda/2$ sections begins to progressively subtract from the main lobe. The maximum gain of a standard dipole "donut" or figure eight pattern occurs when the total length of the antenna is $2 \times 5/8\lambda$, or $5/4\lambda$. This configuration (Fig. 8.18) is known as the extended double Zepp (EDZ) antenna, and it provides a gain of about 2 dB over a $\lambda/2$ dipole at right angles to the antenna wire plane. It consists of two sections of wire, each one of a length

$$L_1\,(\text{feet}) = \frac{600}{f(\text{MHz})} \quad (8.10)$$

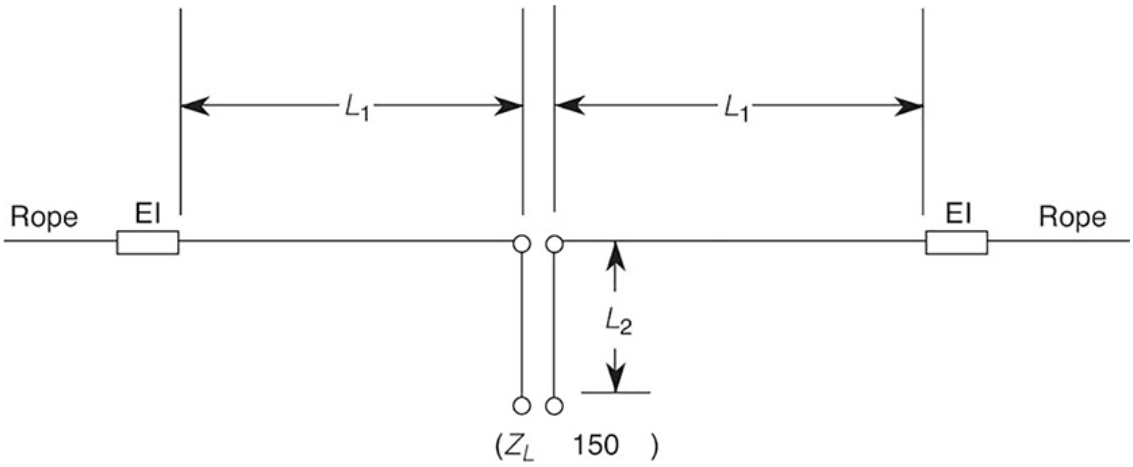


FIGURE 8.18 Extended double Zepp (EDZ) antenna.

Typical lengths for L_1 (or half the total wire length) are 20.7 ft on 10 m, 28 ft on 15 m, 42 ft on 20 m, 84 ft on 40 m, and 168 ft on 80 m.

Because the feedpoint of the extended double Zepp is midway between a current node and a voltage node, the input impedance is also midway between high and low, and contains reactance as well as resistance. The antenna can be fed directly with 450-Ω open-wire line, especially if a balanced antenna tuner is available at the equipment end of the line. Alternatively, it can be fed from a quarter-wavelength matching section (made of 450-Ω twin-lead or equivalent open air parallel line)

L_2

, as shown, or a balun if coaxial cable is preferred. The length of the matching section is

$$L_2 \text{ (feet)} = \frac{103}{f(\text{MHz})} \quad (8.11)$$

The extended double Zepp will work on several different bands. For example, a 20-m EDZ will work as an EDZ on the design band, as a slightly overlength

$\lambda/2$

dipole on 40 m, as a slightly short

$\lambda/2$

dipole on 80 m, and as a four-lobe cloverleaf-patterned antenna on frequencies above the design band. If multiband use is intended, the best feed system is open-wire line back to an ATU.

8.2.7 Multiband Fan Dipole

A standard half-wavelength dipole exhibits a current maximum at its feedpoint not only on its fundamental frequency but also on its third harmonic (where it is 3⁄2 *λ* in length). A 40-m half-wavelength dipole fed with 52-Ω coaxial cable on 7 MHz can also be used on 21 MHz with reasonably comparable VSWR. Its pattern changes from the classic bidirectional figure eight pattern on its fundamental frequency to a four-lobe cloverleaf pattern at its third harmonic.

Dipoles for multiple bands can be hung in parallel and fed from the same feed-line and center insulator. Because a dipole operated far from its resonant frequency exhibits a much higher impedance, most

of the transmitter power will go into the dipole that is resonant on the band being used. Figure 8.19 shows three dipoles (A1–A2, B1–B2, and C1–C2) cut for different bands, operating from a common feedline and balun transformer. Each of these antennas is a half-wavelength on a different band.

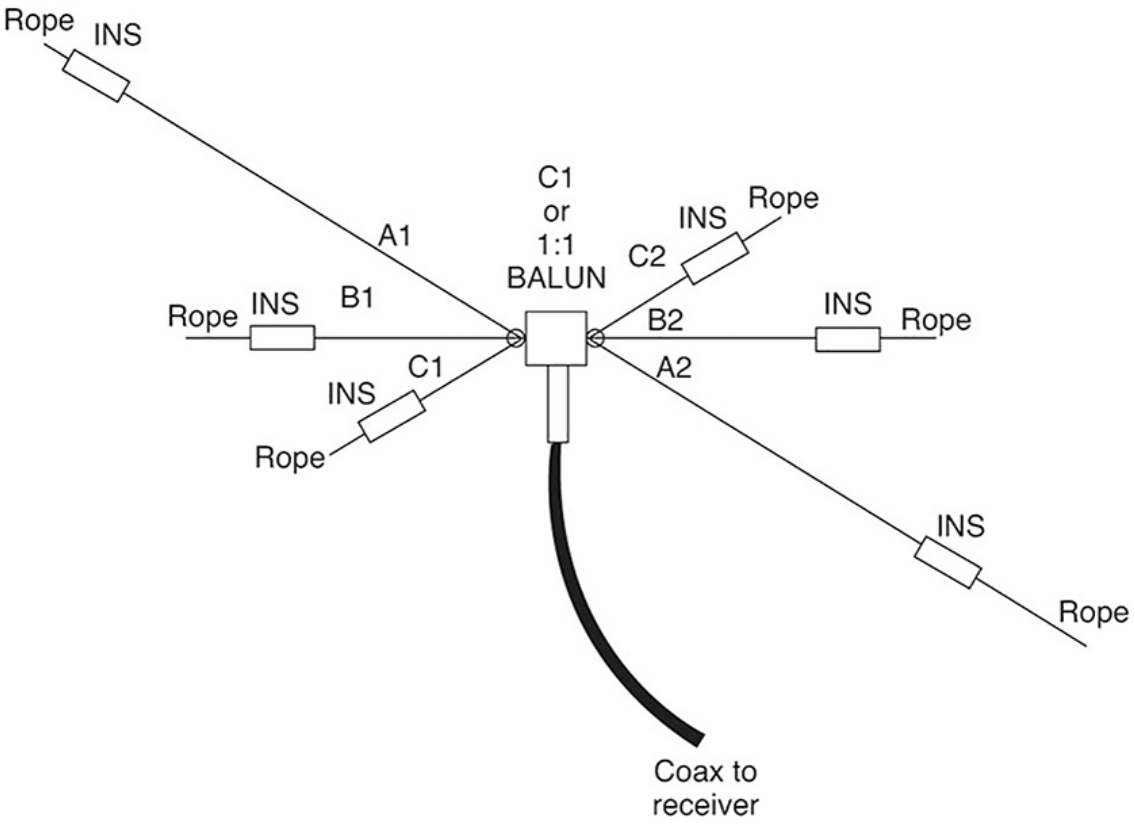


FIGURE 8.19 Multiband dipole antenna.

When building this antenna, we keep the dipoles separated from each other. It will not be necessary to include a separate dipole for any band having a wavelength one-third that of a lower-frequency band that already has a dipole. For example, if you cut A1–A2 for 7 MHz, don't bother including a separate dipole for 21 MHz. If you do, the feedpoint impedance will not be what was expected.

8.3 Summary

From HF to the UHF region, the

$\lambda/2$
dipole is the cornerstone of many antenna installations around the world. The dipole and other derivatives that we will discuss in subsequent chapters provide easy and inexpensive access to the airwaves. For each variation on the basic dipole presented in this chapter, there are many more in use.

CHAPTER 9
Large Wire Loop Antennas

The loop antenna is complementary to the dipole antenna. The dipole can be viewed as a capacitor from which electric field lines detach and radiate into space as electromagnetic waves. Similarly, the loop antenna (when it is small compared to the wavelength) is an inductor from which magnetic field lines detach and radiate into space. Many other antennas with more complex geometries trace their basic mode of producing radiation to either the dipole or the loop antenna.

Loop antennas are characterized as small or large depending on the dimensions of the loop relative to the wavelength of the operating frequency. Small loops have a wire length less than about 0.2λ, and the current flowing in the wire has approximately the same phase and amplitude at every point in the loop. Such loops respond to the magnetic field component of the electromagnetic radio wave. A large loop antenna has a wire length greater than 0.2λ, with most being $\lambda/2$, 1λ, or 2λ. The direction of current flow reverses in each $\lambda/2$ section—a fact that is useful in the simplified analyses of loops we will use in the following sections.

9.1 Half-Wavelength Large Loops

The radiation pattern of a large wire loop antenna is dependent on its size. Figure 9.1 shows a half-wavelength loop (i.e., one in which each of the four sides is $\lambda/8$ long). Assuming S_1 is in its closed position, the real part of the feedpoint (

$X_1 - X_2$
) impedance is on the order of 3 kΩ because it occurs at a voltage maximum. The imaginary or reactive part of the impedance is close to 20,000 Ω. Matching to this impedance is difficult. The antenna has little to recommend it compared to a bent dipole in the same physical configuration with an insulator inserted in the midpoint of the wire segment opposite the feedpoint (corresponding to
S_1
being open).

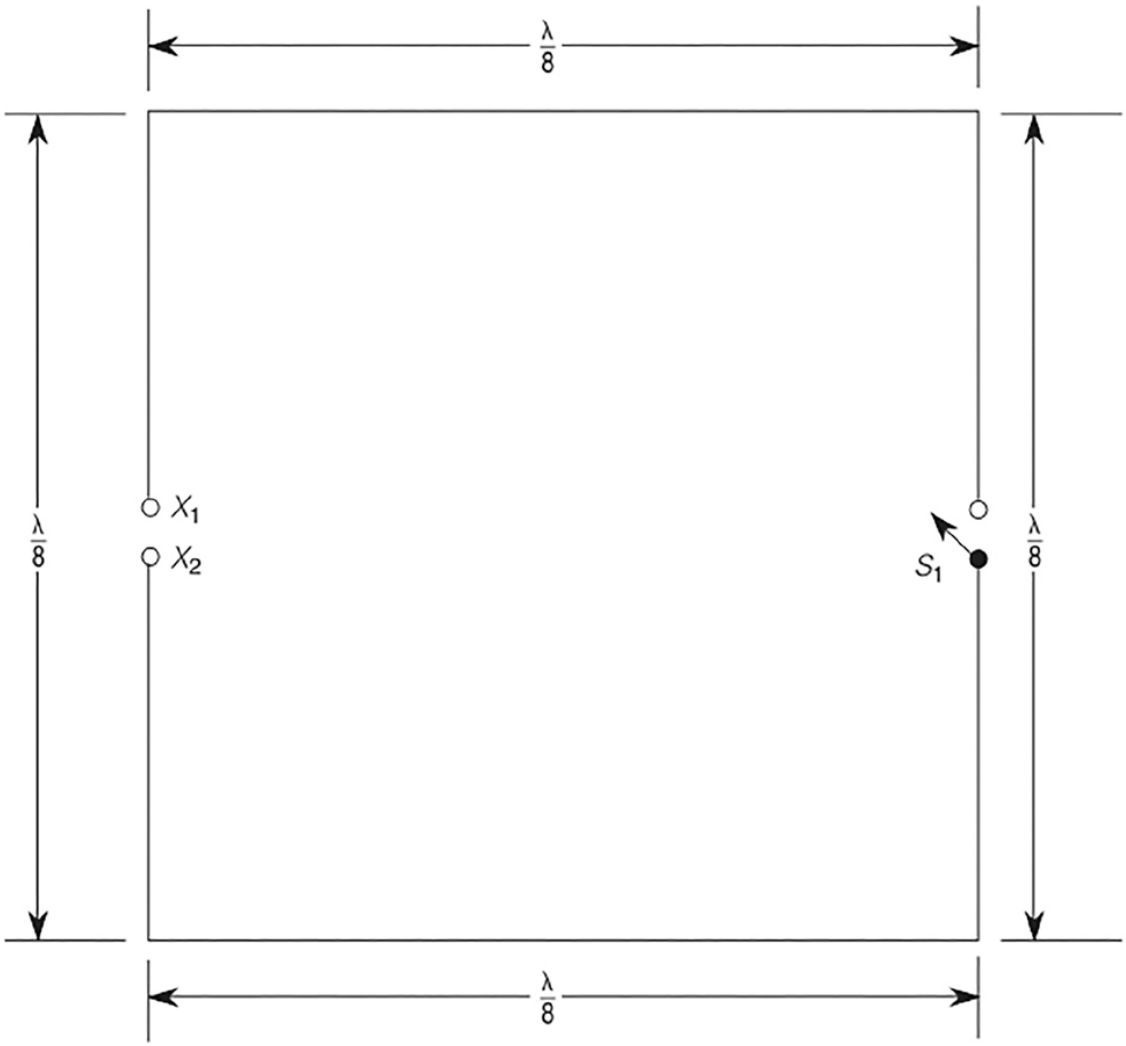


FIGURE 9.1 Half-wave square loop antenna.

For the

$\lambda/2$
loop, the two ends of the wire appearing at the terminals of
S_1
, each *λ*/4 from the feedpoint, carry high voltages that are out of phase with each other in normal operation of a half-wave dipole. Tying the ends together by closing
S_1
results in high resistive and reactive components of feedpoint impedance. A simple trick can tame the input impedance of the
$\lambda/2$
loop. In Fig. 9.2, an inductor (
L_1
or
L_2
) is inserted into the circuit at the midpoint of each wire segment adjacent to the driven segment. These inductors should have an inductive reactance
X_L
of about 370 Ω at the center of the chosen operating band. The inductance of each coil is

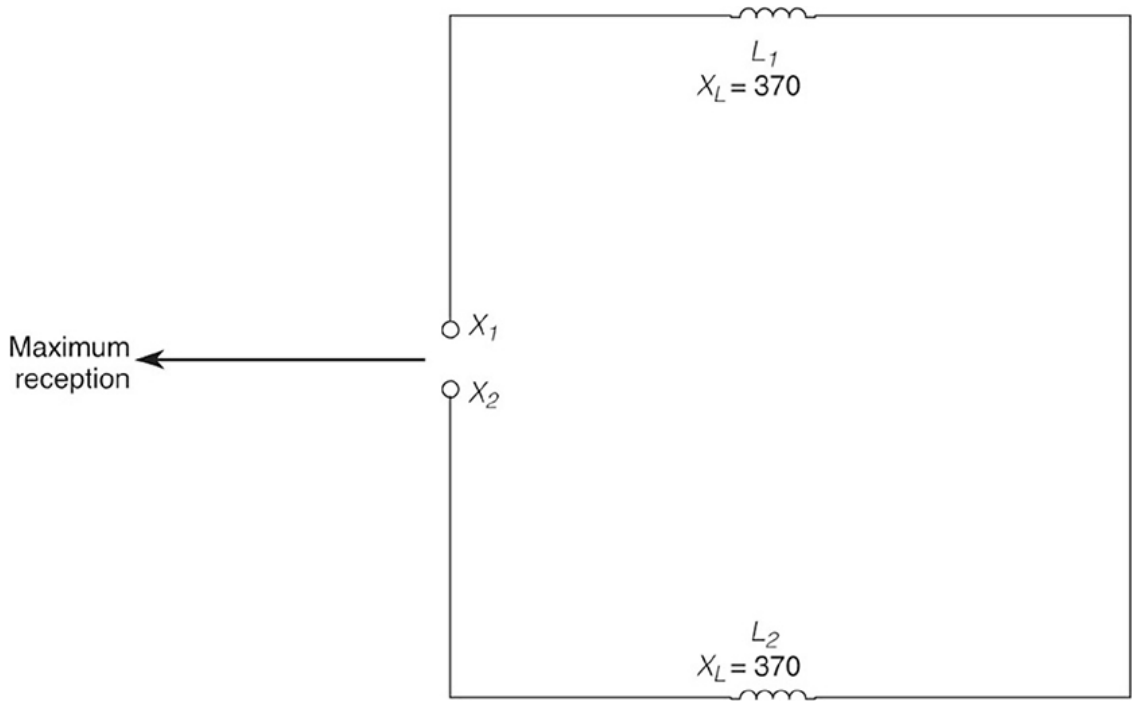


FIGURE 9.2 Inductive loading improves $\lambda/2$ loop.

$$L_1 = L_2 = \frac{3.6 \times 10^8}{2\pi f} \quad (9.1)$$

where *L* is the coil inductance in microhenrys (μH) and *f* is the midband frequency in hertz (Hz).

The coils add lumped circuit inductive loading to the loop, giving it a greater effective length in a small space. The currents flowing in the antenna can be quite high, so, when making the coils, be sure to use a size that is sufficient for the power and current levels anticipated. 2- to 3-in-diameter B&W Air-Dux style coils are sufficient for most amateur radio use. Smaller coils are available on the market, but their use should be limited to low-power situations.

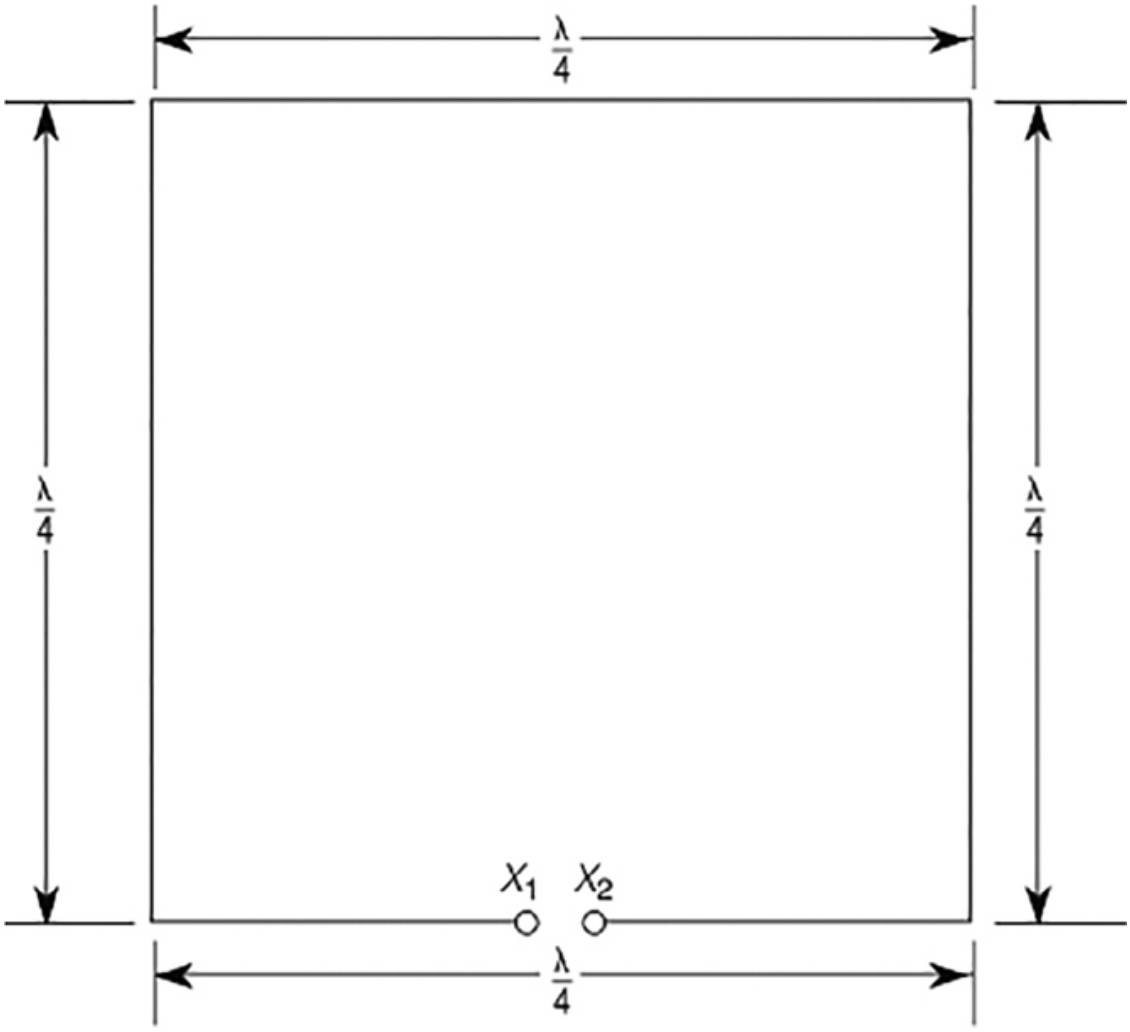
Example 9.1 Find the inductance for each coil in a loaded half-wavelength closed-loop antenna that must operate in a band centered on 10.125 MHz.

Solution Use Eq. (9.1).

$$L = \frac{3.6 \times 10^8}{2\pi \cdot 10,125,000} = 5.7\ \mu\text{H}$$

9.2 1-Wavelength Large Loops

If space constraints allow, a 1λ loop has many desirable features, including a manageable feedpoint impedance, more gain than a dipole in favored directions, and ease of analysis. The 1λ loop of Fig. 9.3 can be treated as two horizontal half-wave bent dipoles whose outer halves are angled toward each other and connected together electrically. As we have seen in an earlier chapter, the center half of a
$\lambda/2$
dipole is responsible for most of the radiated field strength. The ends establish resonance, minimize feedpoint reactance, and help raise the input or feedpoint impedance to a reasonable value. Bending the outer halves of the dipoles 90 degrees has limited effect on the operation of the elements.



F IGURE 9.3 One-wavelength square loop (single-element quad).

Only one of the two dipoles is fed directly from the transmission line. A short circuit is placed across the feedpoint of the other. Current reverses direction in adjacent half-wave sections of a collinear array of dipoles, so after an even number of half-wave segments, the drive current naturally is in phase with the current coming around the full periphery of the loop. At the resonant operating frequency, the current goes to zero at what would be each end of the driven dipole and then reverses direction in the second dipole. This results in the currents in the horizontal sections of the two dipoles being in phase spatially—that is, they both point

in the same direction (left or right in Fig. 9.3). In free space the 1λ loop can be thought of as a two-element driven array, whereby the $\lambda/2$ section opposite the dipole connected to the feedline is driven at its high-impedance ends, rather than at its center, from voltages set up by the fed dipole.

The 1λ loop of Fig. 9.3 produces a bidirectional gain of about 2 dB over a dipole in directions perpendicular to the plane of the loop (into and out of the page). Because the top and bottom horizontal sections are in quadrature phasing straight up and down, the elevation pattern formed by this loop is somewhat flattened.

The loop of Fig. 9.4 also contains one wavelength of wire, but the square has been put into a diamond orientation so that only one tall support is needed. In effect, we have replaced two bent dipoles with two inverted-vees facing each other, one above the other. The connection to the feedline is no longer in the center of a leg but at a vertex joining two legs. Broadside to the loop (in and out of the page), the horizontal components of their radiation are in phase. Since the effective spacing of the two antennas is somewhat less than $\lambda/2$, they interact a bit differently from the square loop of Fig. 9.3, and the net gain compared to a single inverted-vee is compromised a bit. The vertically polarized component of radiation cancels out completely broadside to the loop and creates a weak cloverleaf in the plane of the loop (to the left and right on the page).

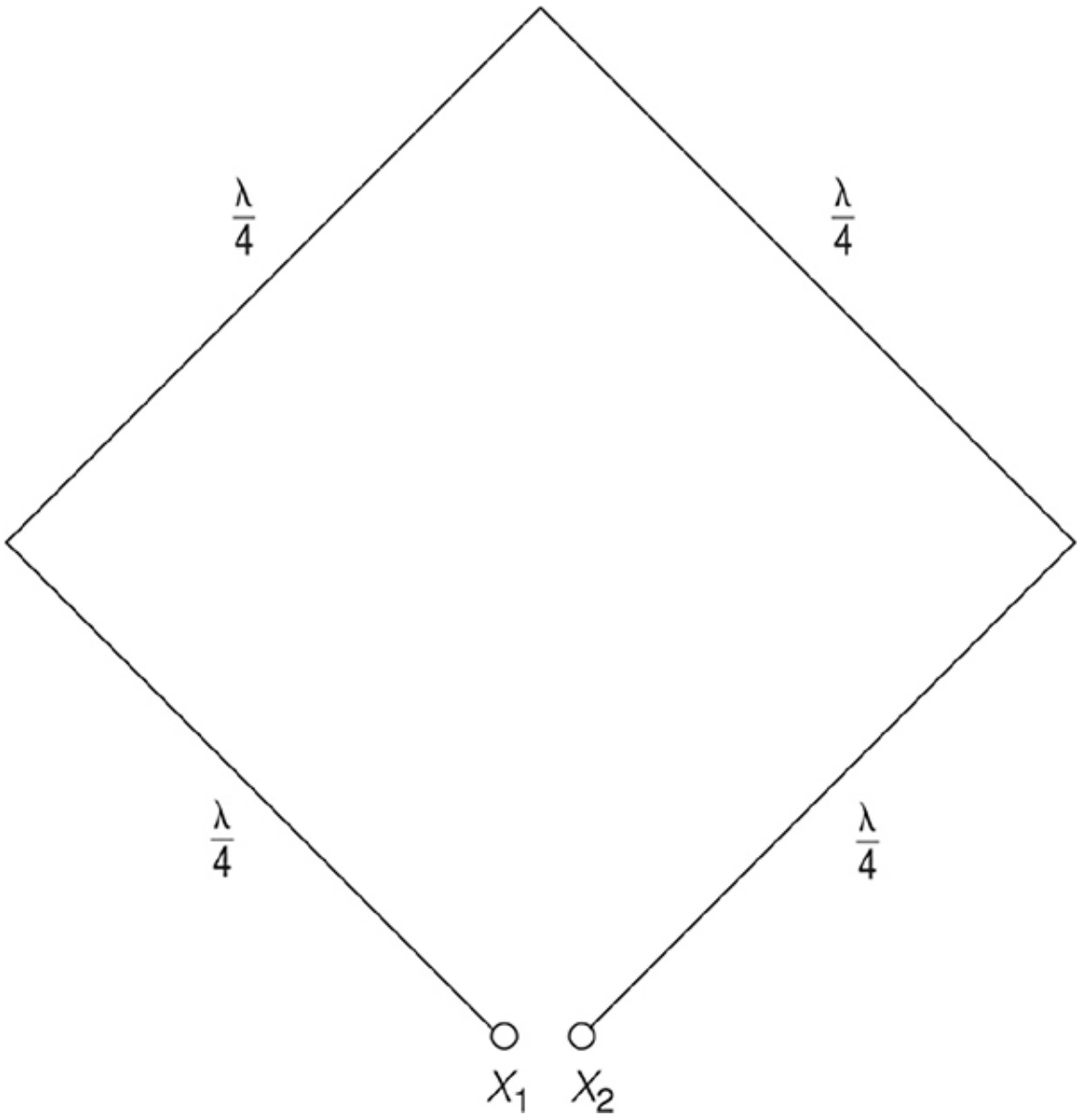


FIGURE 9.4 Bottom-fed diamond loop.

Broadside patterns of both the square and diamond loops are relatively directive in free space, showing 2- to 3-dB gain over a single dipole or inverted-vee. The presence of ground beneath them compromises their gain, especially on the lower HF bands, where few users will be able to raise the bottom of the loop much higher than $\lambda/8$ above ground, and gains of 1 to 2 dB are more likely.

9.3 Delta Loop

The delta loop antenna, like the Greek uppercase letter Δ from which it draws its name, is triangular (Fig. 9.5). Delta loops for VHF frequencies are sometimes oriented point-down (as shown in the figure) to put the horizontally polarized radiating portion of the loop as high as possible. HF delta loops are erected point-up so as to require only one high support. The wire required to form a delta loop is a full wavelength, with elements approximately 2 percent longer than the natural wavelength (like the quad). The actual length will be a function of the proximity and nature of the underlying ground, so some experimentation may be necessary. For the isosceles triangle of Fig. 9.5, the approximate preadjustment lengths of the sides are found from:

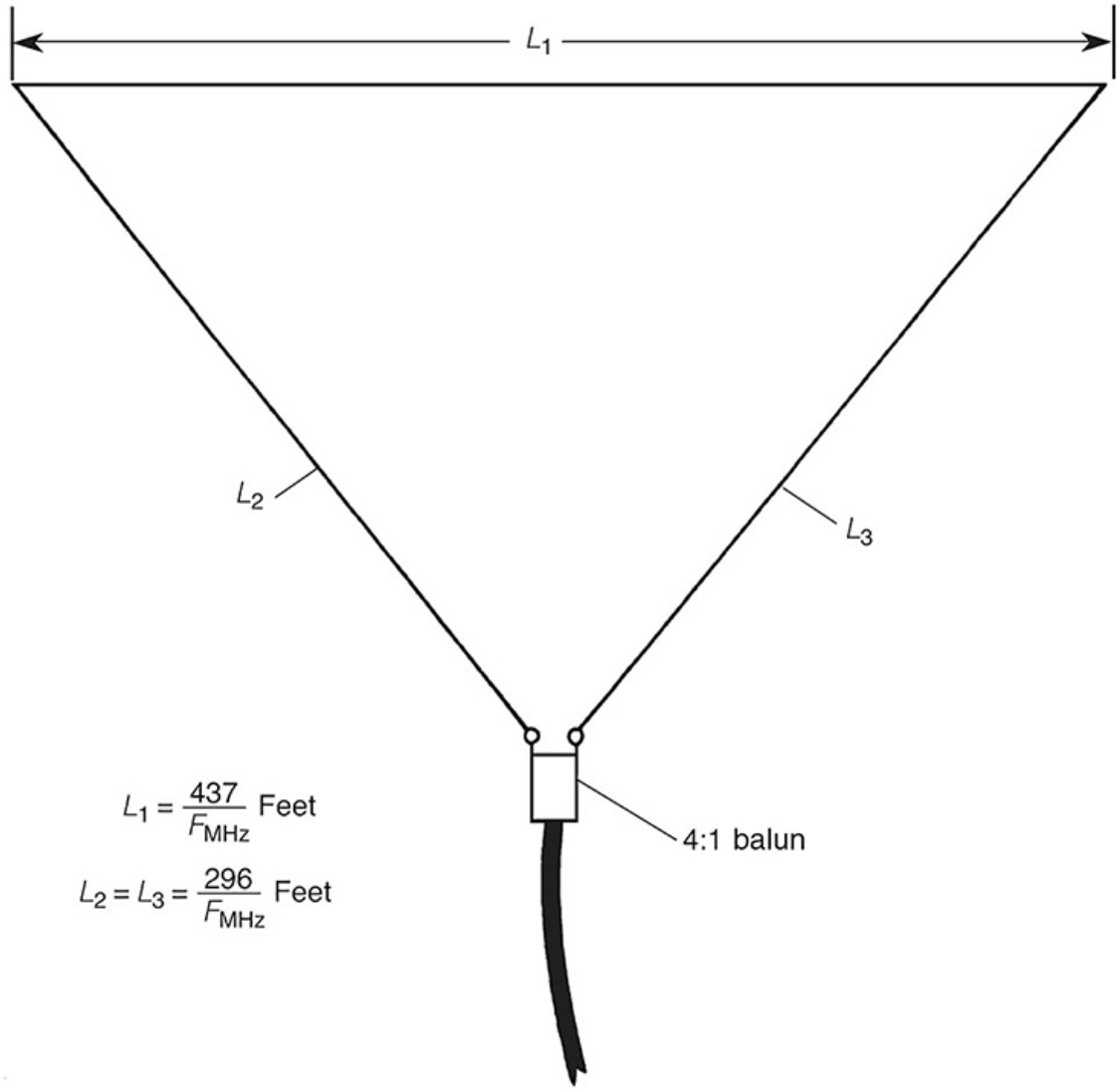


FIGURE 9.5 Delta loop antenna.

$$L_1 = \frac{437}{f(\text{MHz})} \ (\text{ft}) \quad (9.2)$$

$$L_2 = L_3 = \frac{296}{f(\text{MHz})} \ (\text{ft}) \quad (9.3)$$

There is no reason why the delta loop can't be an equilateral triangle, in which case the starting point for each side is

$$L_1 = L_2 = L_3 = \frac{343}{f(\text{MHz})} \text{ (ft)} \quad (9.4)$$

The overall length of wire needed to build any of the 1λ loop antennas is

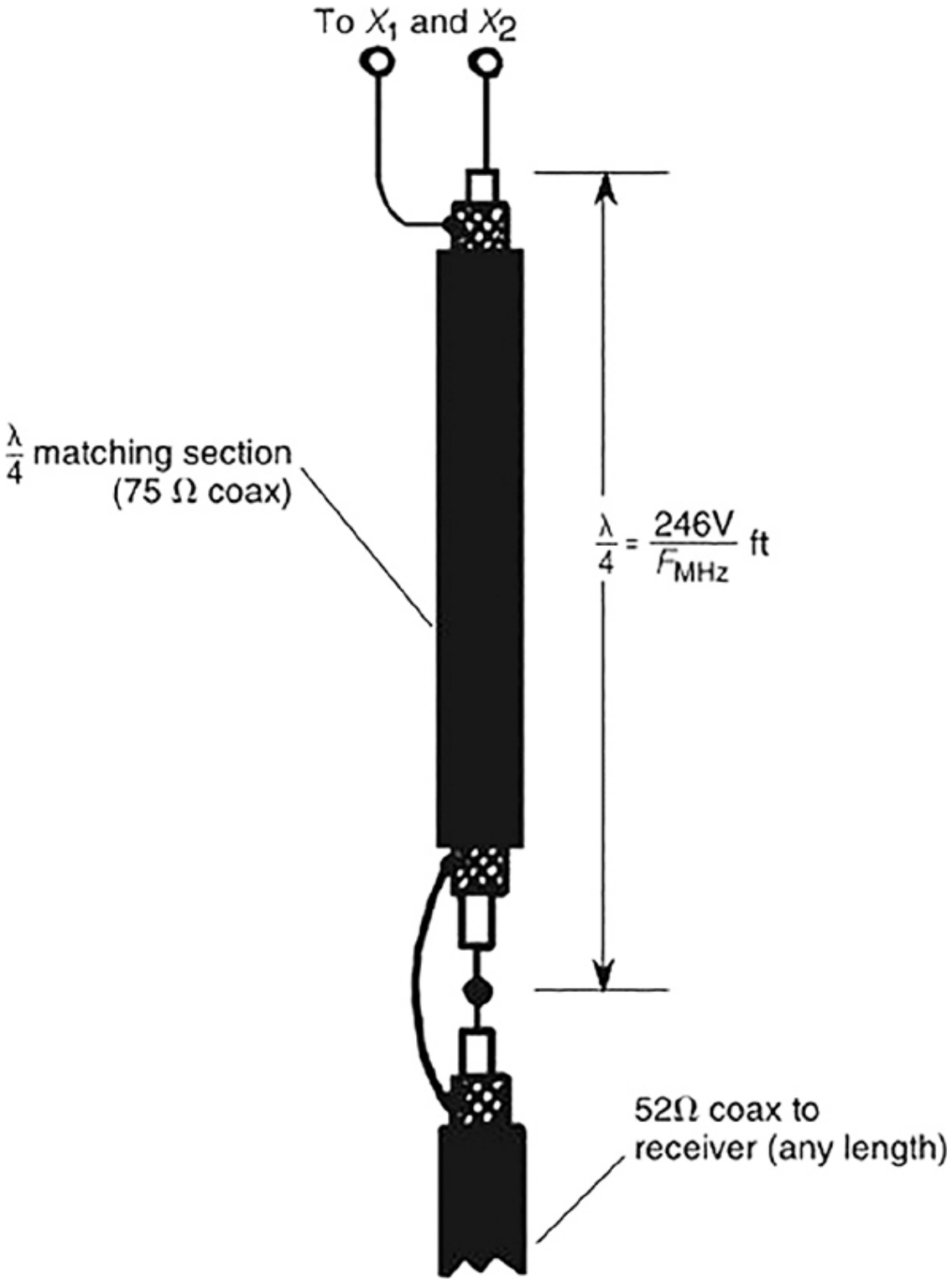
$$L = \frac{1005}{f(\text{MHz})} \text{ (ft)} \quad (9.5)$$

Delta loops for 14 MHz and above can be constructed from either copper wire or aluminum tubing, and are usually mounted on rotatable masts. Delta loops for 7 MHz and below are too large to be easily rotated, but fixed wire versions are popular with many amateurs.

The sharpness of the included angles on the delta loop is not particularly desirable and results in the antenna having less broadside gain than the square (quad) loop, but the triangular shape allows it to be suspended from a single high support. In all cases, maximum gain is in the two directions perpendicular to the plane of the loop.

The polarization of the loop antennas fed, as shown in Figs. 9.3 through 9.5, is horizontal because of the location of the feedpoint. The highest-current sections of a 1λ loop are the two λ/8 sections that straddle the feedpoint and the corresponding continuous λ/4 section halfway around the loop from the feedline. There is nothing magic about that particular feedpoint location. On the square loop, moving the feedpoint to the middle of either vertical side will change the polarization to vertical. On the diamond loop, vertical polarization is realized by moving the feedpoint to either of the two side vertices. On the delta loop, placing the feedpoint at either of the two other vertices produces a diagonal polarization that offers approximately equal vertical and horizontal polarization components, but vertical polarization is easily attained by moving the feedpoint one-fourth of the way up (or down, as the case may be) either side.

The feedpoint impedance of a 1λ loop is influenced by several effects: the baseline feedpoint impedance of a simple $\lambda/2$ "bent" dipole or inverted-vee in free space, the radiated fields from the second $\lambda/2$ section on the fed section, and ground reflection effects, if too close to ground to use free-space assumptions. The resulting input impedance for the array is around 110 Ω for reasonable heights above ground, so it provides a slight mismatch to 75-Ω coax and a 2:1 mismatch to 52-Ω coax. Good matching to 52-Ω coax can be produced using the scheme of Fig. 9.6. Here, a quarter-wavelength coaxial cable matching section (see Chap. 6) is made of 75-Ω coaxial cable. The length of this cable should be



F IGURE 9.6 Quarter-wavelength coaxial matching section.

$$L = \frac{246 v_F}{f\,(\mathrm{MHz})}\ (\mathrm{ft}) \quad (9.6)$$

where
v_F
is the coaxial cable velocity factor. The impedance
Z_0
of the cable used for the matching section should be

$$Z_0 = \sqrt{Z_L Z_s} \quad (9.7)$$

where
Z_0
is the characteristic impedance of coax used in matching section in ohms,
Z_L
is the feedpoint impedance of the antenna, and
Z_S
is the source impedance (i.e., 52-Ω characteristic impedance of line to receiver in standard systems). When Eq. (9.4) is applied to a system having
$Z_s = 52\ \Omega$
and
$Z_L = 110\ \Omega$
, we find for the matching section impedance

$$Z_0 = \sqrt{110 \cdot 52} = 76\ \Omega \quad (9.8)$$

This is a good application of 75-Ω coaxial cable or CATV hardline.

9.4 Half-Delta Sloper (HDS)

The half-delta sloper (HDS) antenna (Fig. 9.7) is similar to the full delta loop, except that (like the quarter-wavelength vertical) half of the antenna is in the form of an image below the ground. Gains of 1.5 to 2 dB are achievable. The HDS antenna consists of a *λ*/3-

wavelength sloping wire and a $\lambda/6$ vertical wire on an insulated mast (or a $\lambda/6$ metal mast). Because the ground currents are important, much like the vertical antenna, either an extensive radial system at both ends is needed or a base ground return wire must be provided.

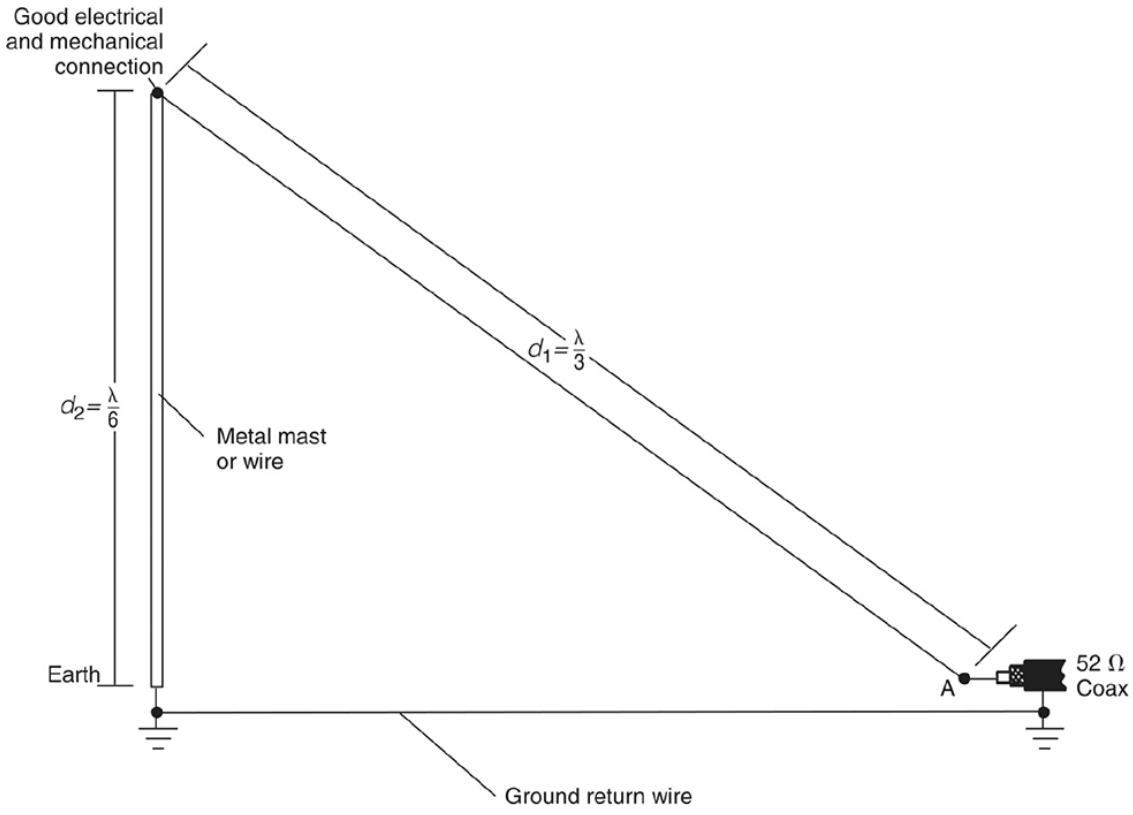


FIGURE 9.7 Half-delta sloper (HDS) antenna.

The HDS will work on its design frequency and harmonics of the design frequency. For a fundamental frequency of 5 MHz, a vertical segment of 33 ft and a sloping section of 66 ft are needed. The approximate lengths for any frequency are found from

$$d_1 \text{ (ft)} = \frac{\lambda}{3} = \frac{328}{f(\text{MHz})} \quad (9.9)$$

and

$$d_2\ (\text{ft}) = \frac{\lambda}{6} = \frac{164}{f(\text{MHz})} \quad (9.10)$$

The HDS is fed at one corner, close to the ground. If only the fundamental frequency is desired, it can be fed with 52-Ω coaxial cable. At harmonics, the feedpoint impedance is as high as 1000 Ω. If harmonic operation is intended, then an antenna tuning unit (ATU) is needed at point A to match these impedances.

As with the delta loop, the sharp interior angle (i.e., <90 degrees) of the HDS is not ideal because components of the radiated fields from the different sections of the antenna will cancel in directions where other components are reinforcing.

9.5 2λ Bisquare Loop Antenna

When the size of the basic large loop is again doubled to 2λ, loop analysis and patterns are more complex. Loops larger than 1λ normally have most gain in the main lobe (relative to a simple dipole) and reasonable feedpoint impedances only when the number and arrangement of
$\lambda/2$
sections are such that the fields from multiple sections consistently reinforce each other in a few specific directions. One way to ensure this is to make sure the physical configuration of the loop results in currents on opposing sides of the loop being in phase with each other. Unfortunately, this becomes much more complicated for 2λ loops.

As an example, let's return to the 1λ loop of Fig. 9.3. When fed at the center of the bottom horizontal λ/4 element, the horizontally polarized radiation of the top element reinforces that of the bottom element, and the primary gain is broadside to the plane of the loop, and horizontally polarized. Radiation from the upper half of each side element is canceled by out-of-phase radiation from the bottom half. So the predominant mode of operation for this antenna with

the feedpoint located as shown is as a two-element horizontally polarized array.

If we increase the length of each side to $\lambda/2$, we can consider this larger loop as comprised of four $\lambda/2$ dipoles, one on each side of the square, connected end to end and fed, as before, at the center of the bottom element. Noting that the current reverses in adjacent dipoles, we see that the current in the top element is now exactly out of phase spatially with the current in the bottom element. The same situation is true for the vertical currents in the two side elements. Because the top and bottom are $\lambda/2$ apart, as are the two sides, maximum horizontally polarized radiation is straight up and down (a cloud burner, in other words) and maximum vertically polarized radiation is to the left and right on the page. Thus, this loop antenna radiates best as a vertically polarized two-element array in the plane of the loop.

Unfortunately, the peak gain of the main vertically polarized lobe is only about 0.6 dB greater than for a single $\lambda/2$ vertical dipole. In the presence of a nearby ground, some small horizontally polarized lobes appear, but over ground up to a height of 1.25λ for the top wire, the antenna does best as an end-fire vertically polarized array. If its overall length is adjusted for minimum reactance at the feedpoint, a 4:1 balun will make it a decent match to either 52-Ω or 75-Ω transmission line.

The 2λ bisquare antenna, as shown in Fig. 9.8, is built like the diamond loop shown earlier (i.e., it is a large square loop fed at a vertex located at the bottom of the assembly) except that each side is now $\lambda/2$ in length. If the reactance is canceled out through slight shortening or lengthening of the total loop, the feedpoint impedance is a good match for either 52-Ω or 75-Ω cable directly. The difference in

feedpoint impedances between the 2λ diamond loop and the 2λ quad loop is a direct result of changing the location of the feedpoint from the midpoint of a side to the corner between two sides.

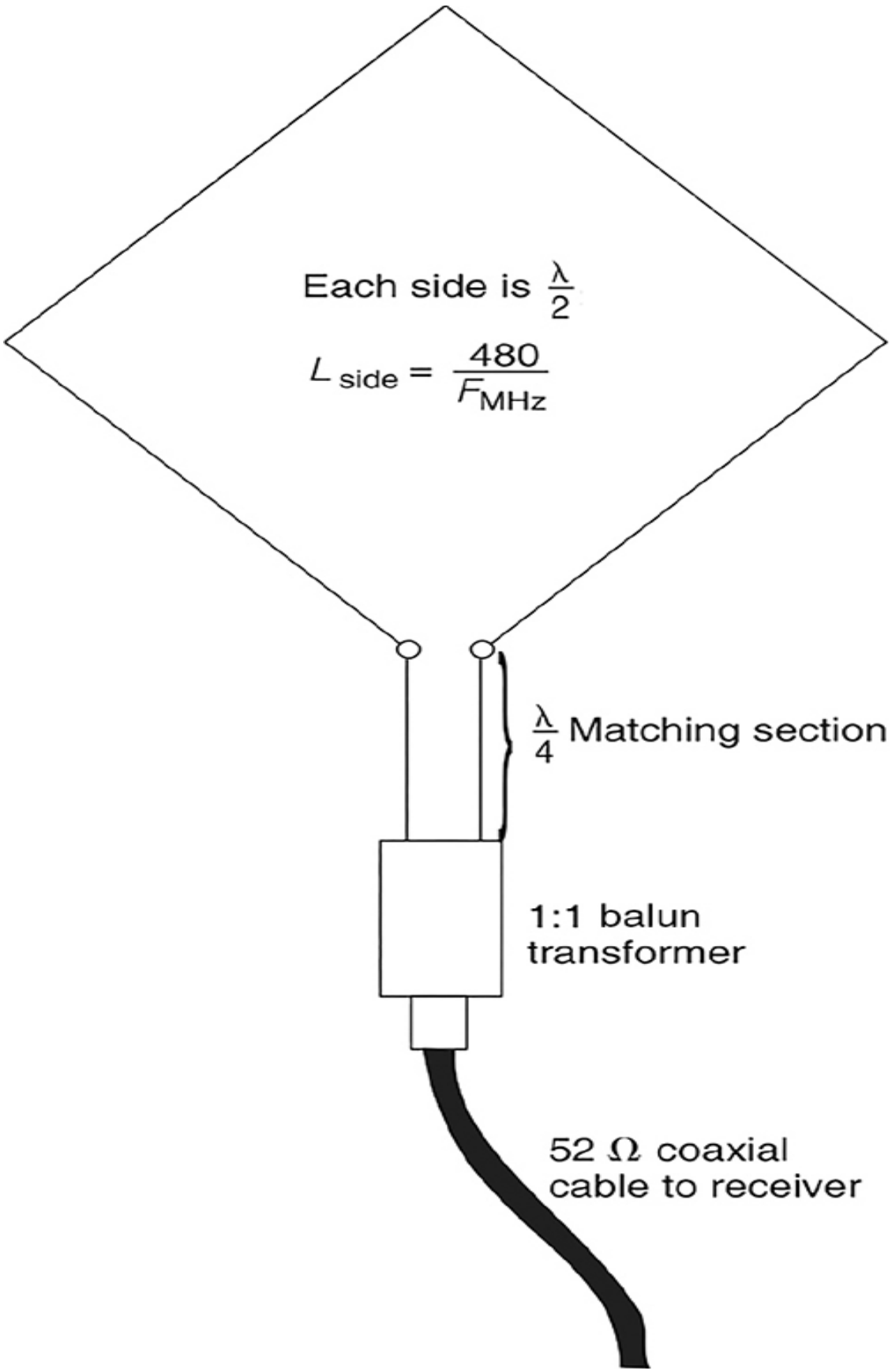

Each side is $\frac{\lambda}{2}$
$L_{side} = \frac{480}{F_{MHz}}$
$\frac{\lambda}{4}$ Matching section
1:1 balun transformer
52 Ω coaxial cable to receiver

FIGURE 9.8 Bisquare 2λ square loop antenna.

When operated at its design frequency and fed, as shown in Fig. 9.8, the bisquare antenna offers maximum horizontally polarized radiation in free space at a 45-degree elevation angle broadside to the plane of the antenna (42 degrees when the feedpoint is $\lambda/4$ above ground). The peaks of the vertically polarized radiation pattern are as before, in the plane of the wire (to the left and right on the page), but the gain is 2 dB or so less than that of a single vertical

$\lambda/2$

dipole. There is usually little to recommend the added mechanical complexity of a 2λ loop unless the phase relationship between opposing sections is corrected with the use of phasing lines or inductors.

9.6 Near-Vertical Incidence Skywave (NVIS) Antenna

The preceding descriptions of large loop operation in the presence of ground are based on hanging each loop from one or two high supports. These loop antennas can instead be suspended so that the plane of the loop of wire is parallel to the ground beneath it. When that is done, at practical heights above ground these loops will send most of their radiation straight up, since the innate array pattern of the 1λ loops generates maximum radiation at right angles to the plane of the loop.

Given the high-angle horizontal polarization of the bisquare loop, which favors short-haul communications links, comparable results can be obtained by laying the same amount of wire on its side and supporting at a more practical height. A 2λ loop of wire arranged in a diamond of uniform height $\lambda/8$ or more can provide nearly the same gain as the bisquare, albeit in a more omnidirectional pattern. The pattern of the 2λ NVIS antenna consists of two separate four-

lobed cloverleafs, one with horizontally polarized lobes at right angles to the four wires, and the other with vertically polarized lobes midway in between.

CHAPTER 10
Multiband and Tunable Wire Antennas

Most communication operators require multiple antennas for different bands or a multiband antenna that operates on any number of different bands. This situation is especially likely on the high-frequency (HF) bands from 3.5 to 29.7 MHz. Even for single-band operation, the feedpoint impedance of many antennas can be highly variable from one end of the band to the other. It is typical for amateurs to design an antenna for the portion of the band that they use most often and then tolerate a high voltage standing wave ratio (VSWR) at the other frequencies. It is possible to tune an antenna over a wide bandwidth using an adjustable antenna tuning unit (ATU) similar to those described in Chap. 28. Resistive losses in an antenna can increase the bandwidth at the expense of efficiency. When you see an antenna that seems to offer a low VSWR over such a wide range, it is possible that the broad response is the result of resistive losses. Here, we cover antennas that are designed to operate in or be tuned to multiple bands.

10.1 Multiband Antennas

For users whose interest (or license) is limited to narrow frequency ranges at a number of different places in the frequency spectrum, multiband antennas provide a reasonable approach to putting a good signal out on each of the desired bands. The sections that follow provide examples of some time-tested ways to do this.

10.1.1 Trap Dipoles

A common form of multiband wire antenna is the trap dipole, as shown in Fig. 10.1. In this antenna category, parallel resonant traps are used in combination with shortened segments of wire or

aluminum tubing to provide the equivalent of quarter-wavelength half-elements on each band of interest. The technique can be applied to monopoles or to both sides of a dipole or other balanced antenna. Typically, each trap is parallel resonant at one of the desired operating frequency ranges. The high impedance associated with a parallel-resonant circuit allows very little radio frequency (RF) energy at the trap frequency to pass from one side of the trap to the other. At frequencies other than the trap design frequency, the RF excitation of the element passes through the trap relatively unattenuated. Below the trap design frequency, the trap appears as a net inductive component and above the design frequency it appears as a net capacitive component.

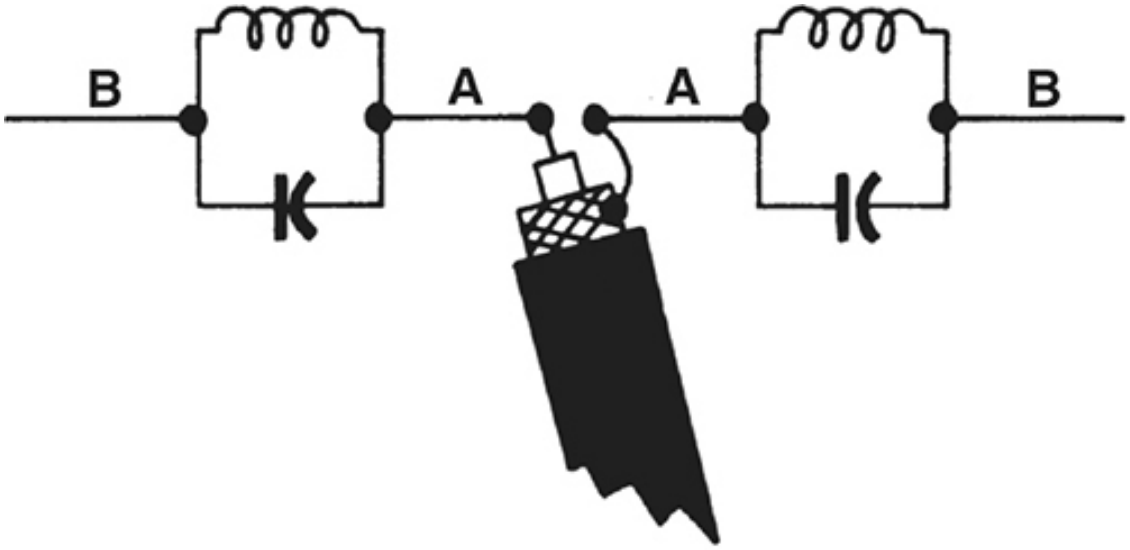


FIGURE 10.1 Trap dipole for multiband operation.

Let's consider how one pair of traps can provide two-band operation with a low SWR for a simple dipole. Suppose that the traps on each side of the center insulator in Fig. 10.1 are parallel and resonant on 10 m. Because of the high impedance of the trap on that band, very little 10-m energy gets beyond the traps, and

only the sections of wire labeled "A" have any appreciable RF in them. If the length of each section A is approximately one quarter-wavelength (or about 8 ft long on the 10-m band), the antenna will function as a resonant half-wave dipole on that band. If a transmitted signal on a lower frequency (say, 15 m) is applied across the center of this antenna, the reactance of the trap capacitor will increase but the reactance of the inductor will decrease. Since the capacitor and the inductor are in parallel, the capacitor will not have a major impact on the operation of the antenna on 15 m, but the inductor will provide a modest amount of lumped-element loading. As a result, the sum of the lengths of A and B will be less than a full

$\lambda/4$

(the natural nontrap length) on 15 m.

In general, trap dipoles are shorter than nontrap dipoles cut for the same band. The actual amount of shortening depends upon the values of the components in the traps, so consult the manufacturer's data for each model of trap purchased.

Some trap antennas employ multiple traps on each side of the center insulator to cover three or more bands with a single feedline. Popular combinations are 20-15-10 and 80-40-20, employing two separate parallel-resonant traps on each side of the insulator. The principle of operation is as previously described, except that a third wire segment is usually found between the two sets of traps on a side. Where more than one pair of traps is used in the antenna, make sure they are intended to work together.

On all bands except the highest one, the radiation efficiency of the trap antenna will be somewhat less than that of a full

$\lambda/4$

monopole or

$\lambda/2$

dipole for the same band. That is because a small portion of the radiating element has been converted to a nonradiating lumped inductance. The effect is small, usually resulting in a net reduction in radiated field strength of 0.5 dB or so.

A disadvantage of trap antennas is that they provide less harmonic rejection than an antenna designed for a single band. The antenna has no idea what band the transmitter "thinks" it's transmitting on, so any harmonic energy that falls within any of the design bands of the trap antenna will be radiated with the same efficiency. As a result, users of multiband antennas need to ensure that harmonics and other out-of-band spurious emissions from their transmitters and associated amplifiers are as low as possible.

10.1.2 Multiple Dipoles

Another approach to multiband operation of an antenna consists of two or more half-wave dipoles fed from a single transmission line, as shown in Fig. 10.2. There is a practical limit on how many dipoles can be accommodated based on coupling effects and the total weight of the antennas and their spacers.

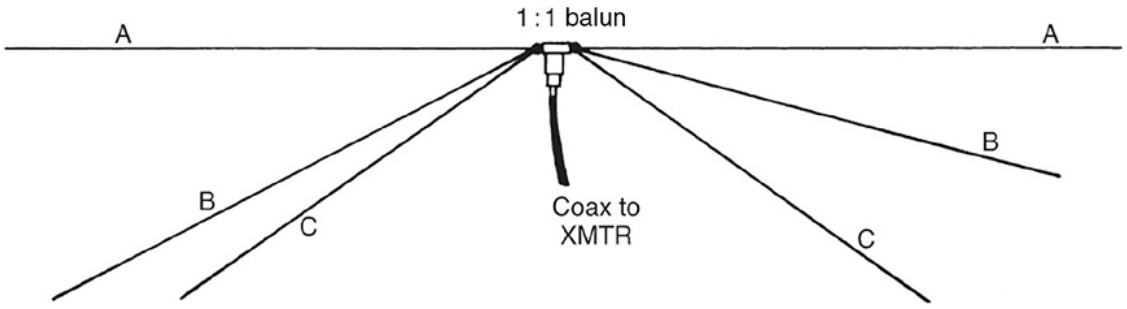


FIGURE 10.2 Multiband dipole consists of several dipoles fed from a common feedline.

Assuming the dipoles are at least
$\lambda/2$
above ground at the lowest frequency, a reasonably good match to the feedpoint impedance is provided by either 50-Ω or 75-Ω coaxial cable. The two sides of the coax (center conductor and shield) can be connected to the center of the multi-dipole directly or through a 1:1 balun transformer, as shown in Fig. 10.2. Each antenna (A-A, B-B, or C-C) is cut to
$\lambda/2$

at its design frequency, so the length of each dipole can be found from the standard expression for dipole length. The overall length ($\mathrm{A}+\mathrm{A}$, $\mathrm{B}+\mathrm{B}$, or $\mathrm{C}+\mathrm{C}$) for each dipole can be found from

$$L = \frac{468}{f(\mathrm{MHz})} \text{ ft} \quad (10.1)$$

If the antenna is close to the earth's surface, these equations are approximations. Some experimentation may be necessary to optimize resonance on each band. Drooping dipoles (B and C in this case) may act more like an inverted-vee antenna (see Chap. 8) than a straight dipole, so the equation length will be a few percent too short. Some amateurs build the multiple-band dipole from four- or five-wire flat TV rotator cable. Starting with the highest frequency, cut each wire to the required length and strip off any unused portions.

Another possibility is the jumper-tuned inverted-vee, as shown in Fig. 10.3. In this situation, a single conductor, broken into segments A, B, and C (or more, if desired), is used for each leg of the antenna. Adjacent segments are separated by inline insulators (standard end insulators are suitable). Segment A is a quarter-wavelength on the highest-frequency band of operation, $\mathrm{A}+\mathrm{B}$ is a quarter-wavelength on the next highest band of operation, and $\mathrm{A}+\mathrm{B}+\mathrm{C}$ is a quarter-wavelength on the lowest-frequency band of operation. The antenna is tuned to a specific band by either connecting or disconnecting a wire jumper (see inset) across the insulator between the segments. A switch or an alligator clip jumper will short out the insulator to effectively lengthen the antenna for a lower band.

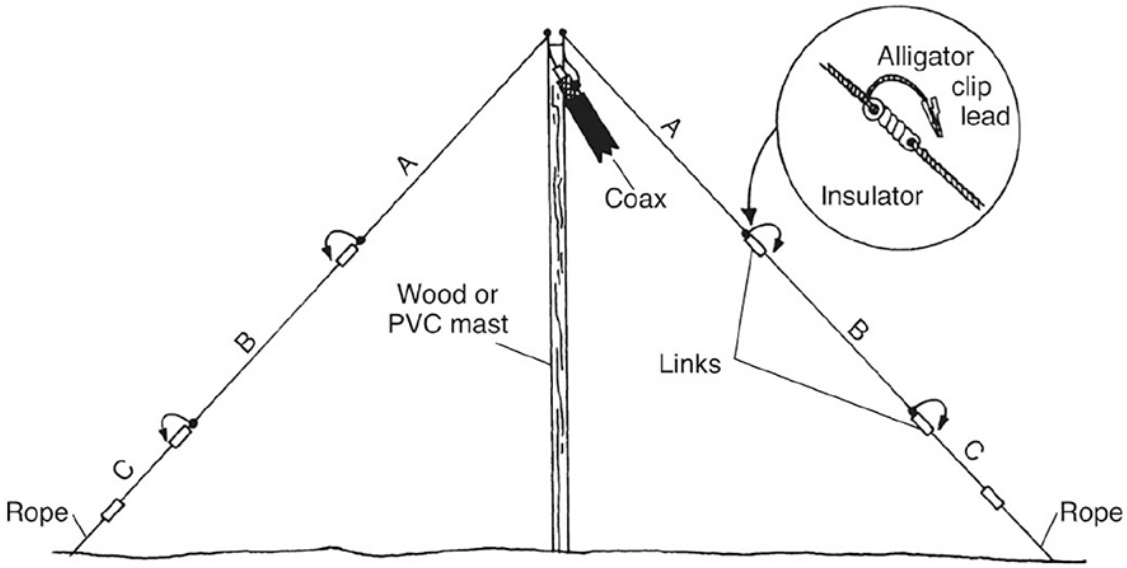


FIGURE 10.3 Multiband inverted-vee uses shorting links to change bands.

A disadvantage of this type of multiband antenna is that one must manually switch the jumpers to change bands, which probably explains why other antennas are more popular. One way to get at the higher jumpers is to make the center support a tilt-over mechanism (as described in Chap. 32).

10.1.3 Log-Periodic Antenna

The log-periodic (LP) antenna resembles a long boom Yagi (Chap. 14) parasitic array. Unlike the Yagi, it is an all-driven array that derives its name from the fact that each element length and the spacing from that element to the next one is a constant percentage of the previous element's length and spacing. An open-wire parallel transmission line that runs the length of the boom feeds the centers of all elements, which are split and insulated at their midpoints. The transmission line is flipped as it passes from element to element so that each element is fed 180 degrees out of phase with respect to the ones on either side.

Log-periodic antennas have been around for many years and a wide variety of commercial products are available, as well as design guides and tutorials for those who wish to build their own. An

example of log-periodics is the old multichannel VHF TV antenna. Larger HF LPs can be found at many military bases around the world.

A log-periodic antenna can attain over a 2:1 or greater frequency span the same performance characteristics (forward gain, front-to-back ratio, etc.) that a typical three-element Yagi exhibits over a very narrow band of frequencies. To accomplish this, the LP requires a boom length that is on the order of a wavelength or greater at the lowest operating frequency. LPs can make sense for wideband applications, but they are mechanically more cumbersome than the available alternatives for amateurs and others with widely separated operating bands. Compared to a trap triband Yagi for 20-15-10 m, an LP of equivalent electrical performance requires a substantially stronger supporting tower and rotator and has a far larger turning radius.

10.2 Tuned Feeder Antennas

A popular approach to multiband operation with a single antenna is the tuned feeder antenna, as shown in Fig. 10.4. When fed this way, an 80-m dipole can be used from 160 through 10 m, but it requires a good ATU and a length of balanced (parallel wire) transmission line.

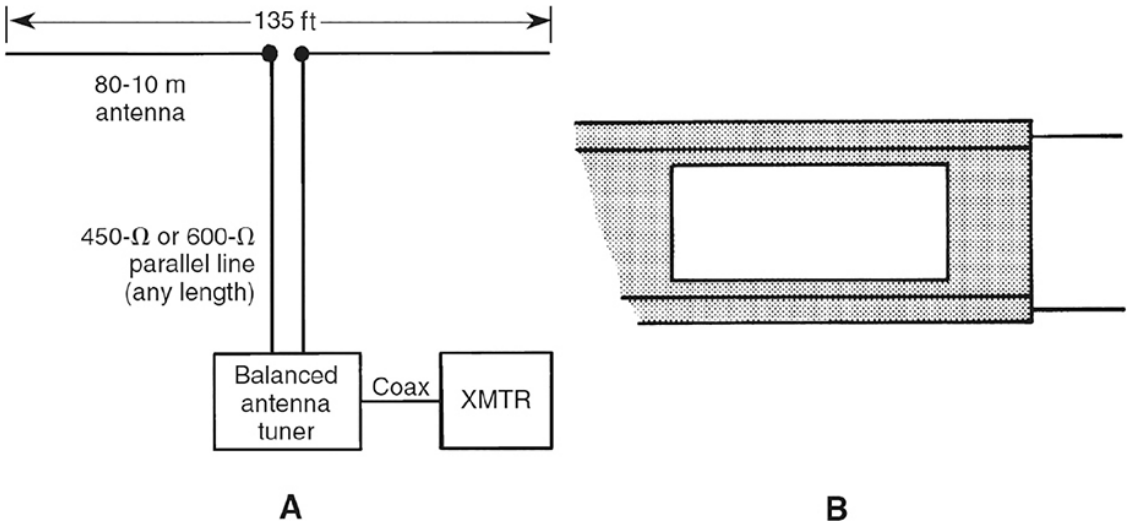
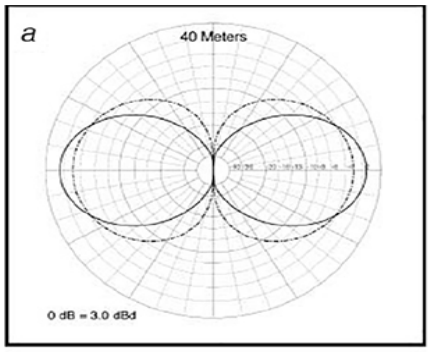


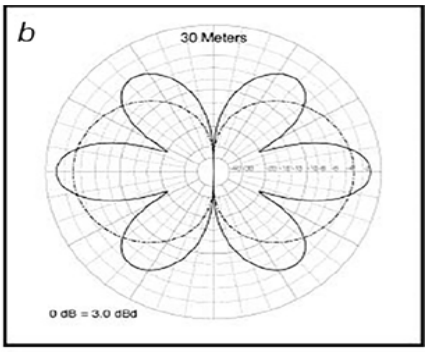
FIGURE 10.4 Tuned feeder antenna can be used on several bands.

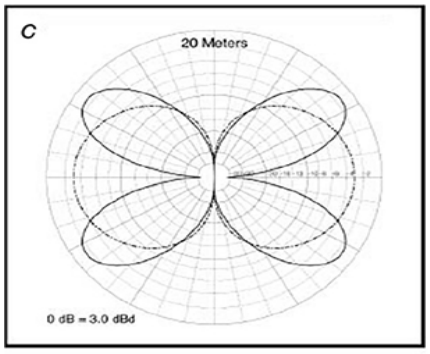
A suitable ATU for this antenna should be capable of matching to a wide range of impedances and that has heavy-duty components to minimize internal losses caused by high circulating currents in the coupler components and the possibility of arc-over caused by high voltages. Even at power levels of 150 W or less, an ATU designed for 1500 W into a matched load may be needed. An ATU can also accommodate balanced and unbalanced loads ($\text{antenna} + \text{feedline}$).

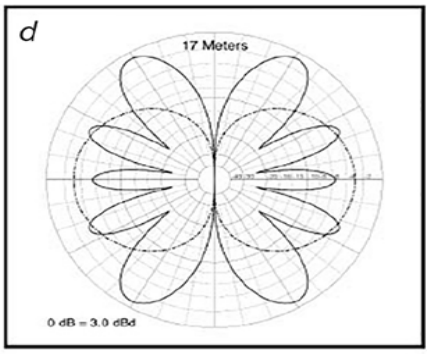
The antenna of Fig. 10.4 is a center-fed dipole that exhibits the familiar figure eight or donut-shaped radiation pattern near its fundamental frequency, when each side of the antenna is approximately $\lambda/4$ in length. As the operating frequency is increased, the dipole legs become longer in terms of wavelength. This causes the peak of the main to grow even larger than it is for a $\lambda/2$ dipole and eventually to decline and break into additional lobes that begin to pop up at other angles relative to the axis of the antenna. Figure 10.5 shows the radiation patterns for an 80-m dipole

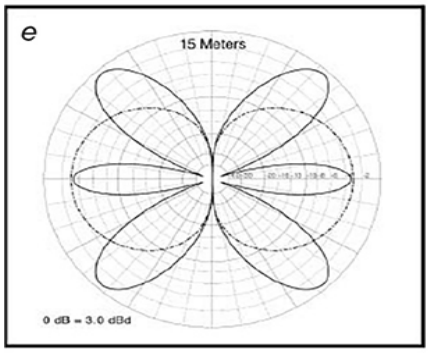
operated in free space at a few selected higher frequencies. When used closer to earth, the nulls in the radiation pattern are less deep than shown here. As a result, this antenna will actually be useable over a wide range of signal angles. At higher frequencies, a dipole in free space that is resonant at 3.6 MHz will exhibit resistive input impedances ranging from 40 or 50 Ω up to 5K Ω or so (the second harmonic is often the worst) and reactances up to ±2000 Ω. Adding a 40-m dipole to the same feedpoint can reduce the range of impedances that must be matched on the higher bands.

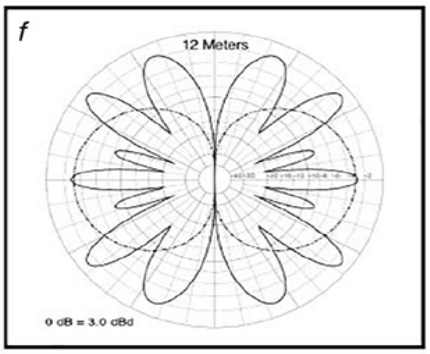

a
40 Meters
0 dB = 3.0 dBd

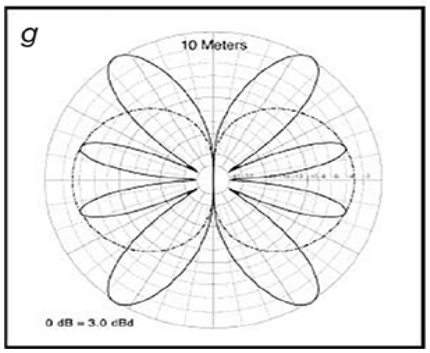

b
30 Meters
0 dB = 3.0 dBd


c
20 Meters
0 dB = 3.0 dBd


d
17 Meters
0 dB = 3.0 dBd


e
15 Meters
0 dB = 3.0 dBd


f
12 Meters
0 dB = 3.0 dBd


g
10 Meters
0 dB = 3.0 dBd

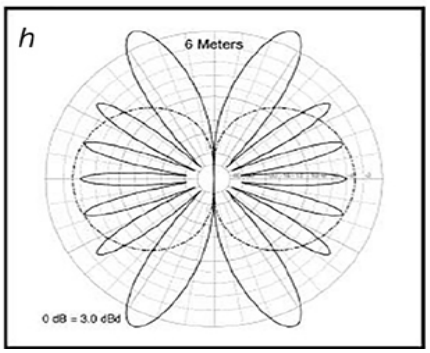

h
6 Meters
0 dB = 3.0 dBd

FIGURE 10.5 Patterns for an 80-m center-fed dipole on higher frequencies (solid line) versus its 80-m pattern (broken line).

A dipole can be fed at other locations than the center. Feeding a $\lambda/2$ dipole at one end means the feedpoint impedance will be very high (a few thousand ohms), but a good ATU should be able to handle this. Figure 10.6 shows the end-fed Zepp antenna. This antenna uses a half-wavelength radiator but it is fed at a voltage node rather than a current node (i.e., the end of the antenna rather than the center). Typically, 450- or 600-Ω parallel-conductor air-dielectric open-wire transmission line is used to feed the Zepp because of the high voltages on the line as a result of the extreme impedance mismatch between the line and the antenna. In theory, the line can be any length, but the task of the ATU is simplified if the length of the feed line is an odd number of quarter-wavelengths for those bands where the antenna length is a multiple of $\lambda/2$ at the operating frequency. When that condition is met, the transmission line transforms the high feedpoint impedance to a lower value that is more apt to fall within the ATU's range. For example, a $\lambda/4$ section of 600-Ω open-wire line will transform a 3000-Ω feedpoint impedance at one end of a $\lambda/2$ dipole down to an easily matched 120 Ω.

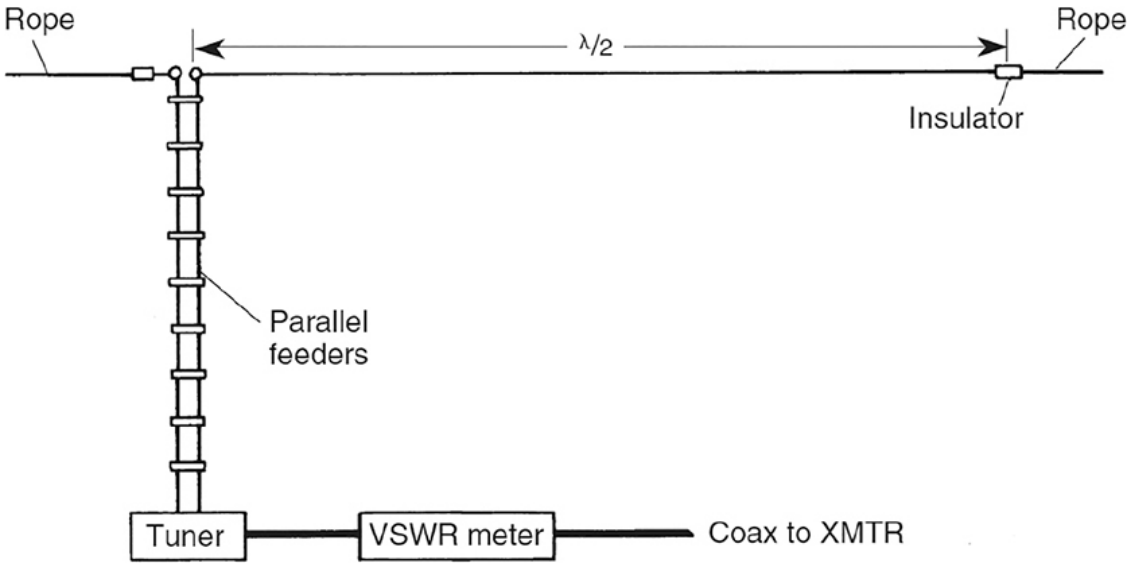


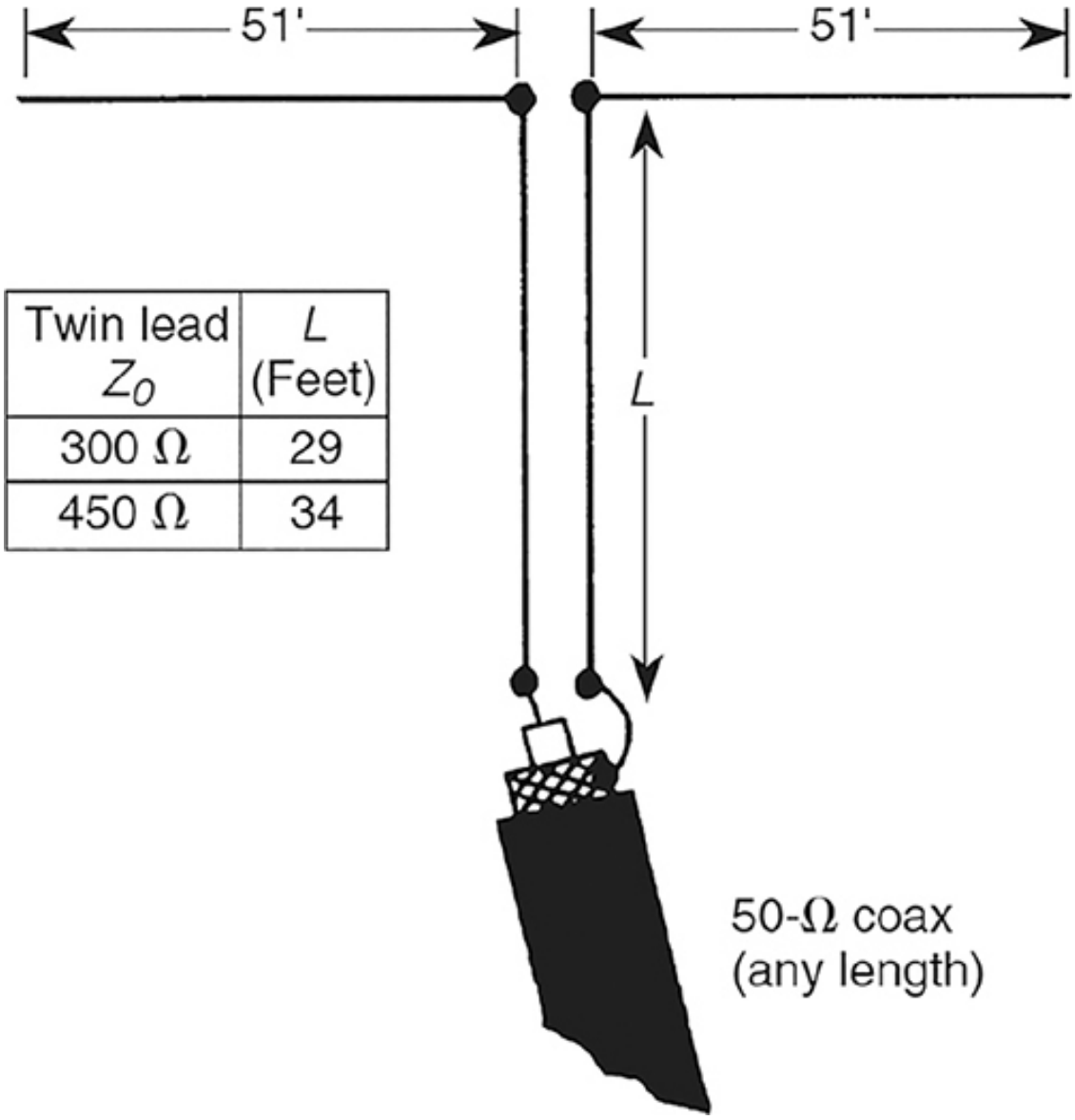
FIGURE 10.6 End-fed Zepp antenna.

As the operating frequency is raised above the point where the wire is
$\lambda/2$
in length, the radiation pattern begins to depend on the location of the feedpoint. At the fundamental operating frequency (80 m in our example), there is no difference in radiation pattern between the end-fed Zepp and the center-fed dipole of Fig. 10.4. On 40 m, the center-fed dipole still has a figure eight pattern similar to the one on 80 but with about 2 dB more gain in the main lobe. The end-fed Zepp's pattern resembles a four-leaf clover with the pattern peaks at 45 degrees to the axis of the wire. The end-fed Zepp's 40-m pattern is actually similar to the center-fed dipole's 20-m pattern.

10.2.1 G5RV Multiband Dipole

Figure 10.7 shows the basic dimensions of an antenna that has enjoyed much popularity over the years, the *G5RV* antenna. It can be used as a horizontal dipole, a sloper, or an inverted-vee antenna. In its standard configuration, each side of the dipole is 51 ft long and it is center-fed with a matching section of either 29 ft of 300-Ω line or 34 ft of 450-Ω line. The antenna is shorter than a
$\lambda/2$

80-m dipole but longer than a 40-m
$\lambda/2$
dipole. The dipole itself is not a particularly good match on 80, 40, 20, or 15 m, but the effect of the short matching section is to bring the feedpoint impedance of the combination closer to a reasonable match for either 50-Ω or 75-Ω coaxial cable on those bands.



Twin lead Z_O	L (Feet)
300 Ω	29
450 Ω	34

F IGURE 10.7 G5RV antenna.

The VSWR of the assembly varies significantly across the HF spectrum, making it difficult to use for amateur bands at 12, 17, and 30 m. With a good ATU, the antenna can be matched throughout the HF spectrum without a specific length of 300-Ω or 450-Ω line.

10.2.2 Longwire Antenna

If, instead of using a feedline, a single wire is brought from the end-fed antenna directly into the radio room or to the ATU, the antenna is a longwire antenna. In this case, the feedline is part of the antenna radiating system. Such an antenna configuration can work but requires an RF ground system (see Chap. 33) attached to the chassis of the ATU or transmitter where the wire connects. Bringing the radiating portion of the antenna into the radio room, coupled with an inadequate RF ground, may have some unintended consequences: greater exposure to RF fields, interference with other electronics, or RF burns to the fingers caused by touching the metal chassis of the transmitter.

By longstanding convention, a longwire is a full wavelength or longer at the lowest frequency of operation. In Fig. 10.8, we see a random-length longwire antenna fed from a tuning unit. As the operating frequency is varied, the feedpoint impedance of the long wire will vary significantly, with a reactive component that can be quite substantial. An ATU is important for matching the feedpoint impedance of the antenna to the typical transmitter output impedance of 50 Ω.

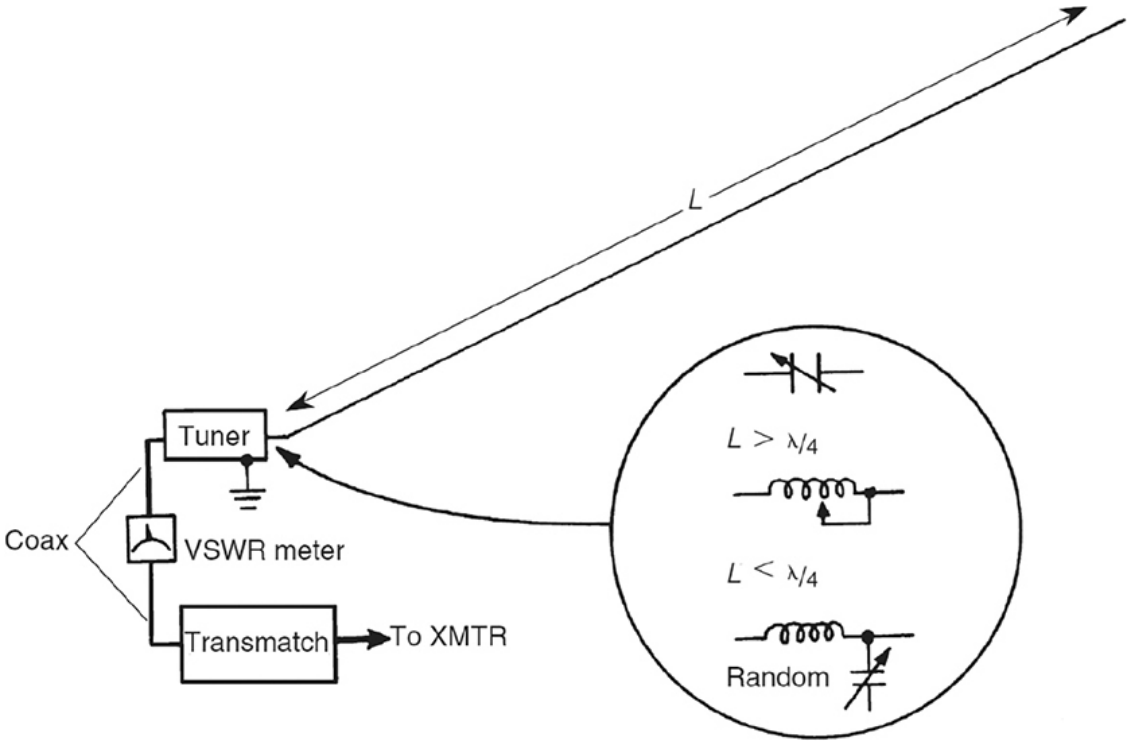


FIGURE 10.8 Random-length longwire antenna.

10.2.3 Off-Center-Fed Dipole

The radiated field from a half-wave dipole operated at its fundamental resonant frequency is the result of a standing wave of current and voltage along the length of the dipole. The center of the half-wave dipole is a point of maximum current and minimum voltage. If we feed the
$\lambda/2$
dipole at its center, the feedpoint impedance is 73 Ω in free space, and it oscillates between a few ohms and 100 Ω as the antenna is brought closer to ground. If we feed the dipole at one end instead of in the center, the feedpoint impedance is quite high, perhaps between 3000 Ω and 6000 Ω, depending on the effects of insulators and other objects near the ends of the wire. By feeding the dipole at an intermediate location between the center and one end it is often possible to find a more attractive match to a specific feedline's characteristic impedance. Off-center feeds suffer from a lack of balance between the two sides of the feedline. As a result, it

is harder to keep the feedline from radiating and becoming an unintended part of the antenna.

10.2.4 Windom

The Windom antenna (Fig. 10.9) has been popular since the 1920s. Although Loren Windom is credited with the design, there were actually multiple contributors, including John Byrne, E. F. Brooke, and W. L. Everett. Additional (later) contributions were rendered by G2BI and GM1IAA. The Windom is a roughly half-wavelength antenna that also works on even harmonics of the fundamental frequency. Like the off-center-fed dipole, the idea is that a dipole's feedpoint resistance varies from about 50 Ω at the center to about 5000 Ω at either end, depending upon the location of the feedpoint. In the Windom antenna of Fig. 10.9, the feedpoint is placed about one-third the way from one end, where the impedance is usually about 600 Ω.

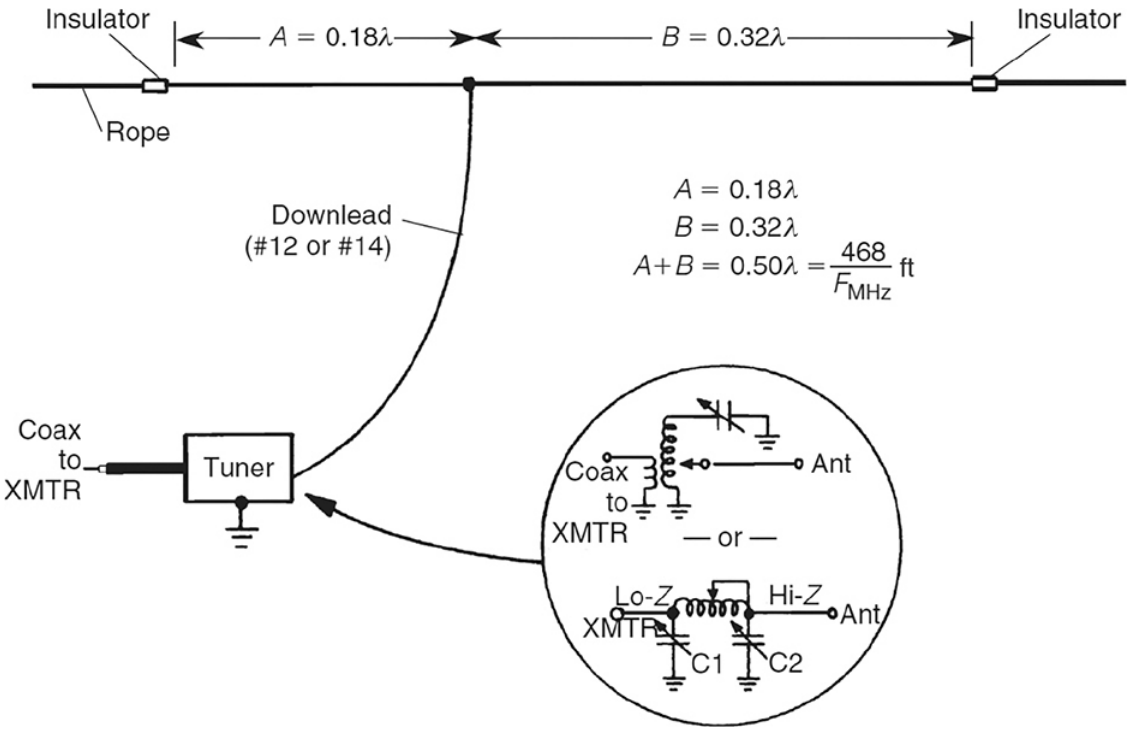


FIGURE 10.9 Windom antenna.

The Windom antenna works moderately well with some caveats. The return path for a single-conductor feedline is the ground system underneath the antenna and feedline. In contrast to center-fed horizontal dipoles, the quality of the ground beneath the antenna is a major factor in the overall radiation efficiency of the Windom. Further, this is an inherently unbalanced radiating system, which can lead to interference and unwanted induced RF currents near the transmitter. One could view the Windom as a lopsided "T" antenna or as an inverted-L with a secondary section of top loading attached with the feedline itself as the primary radiator. The choice of tuning unit for the Windom will depend on the frequencies it is to be used on, but a good tuner capable of removing large amounts of reactance will be required for some of the HF amateur bands.

An alternative to the single-conductor feedline Windom is shown in Fig. 10.10. A 4:1 balun transformer is placed at the feedpoint and is connected to 75-Ω coaxial transmission line to the transmitter. A transmatch or similar antenna tuner may also be needed at the transmitter end of the transmission line.

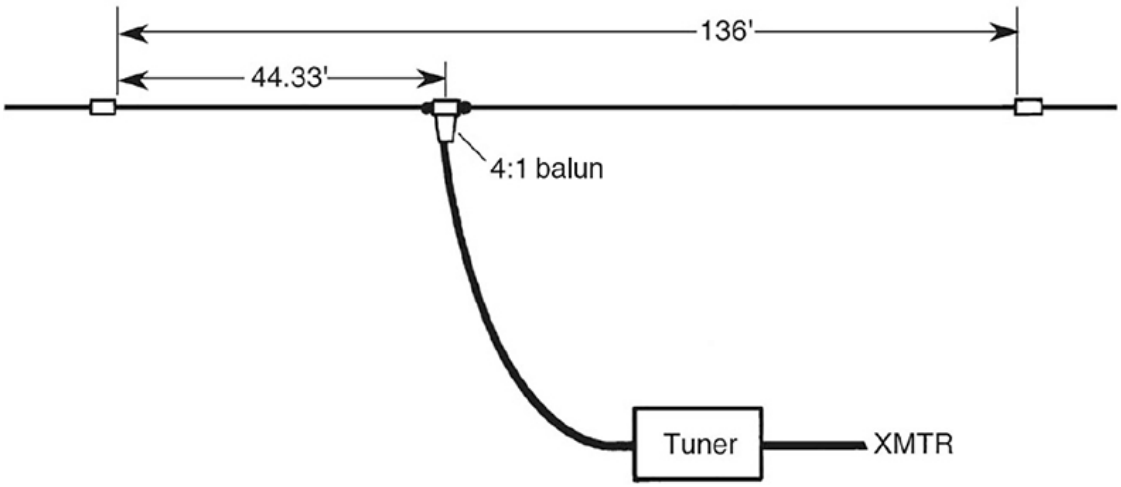


FIGURE 10.10 Coaxial-fed Windom.

CHAPTER 11
Vertically Polarized Antennas

In previous chapters, we defined the polarization of an antenna as being the same as the orientation of the electrical field radiated by the antenna. The direction of the electric field for a specific antenna design is a function of the geometry of the radiating structure. For a complex structure with different portions creating fields having different orientations and phasing, the radiated field is the vector sum of all the individual E-fields. The fields radiated by an antenna are affected by any metallic or conducting bodies in close proximity to it. The earth's surface, the roof of an automobile, or buildings and other structures can change the polarization of the field radiated by the antenna.

In simple cases, it is easy to determine the polarization of antenna. For an antenna consisting of a single wire or thin rod oriented vertically, the E-field points in the same direction as the long dimension of the antenna, so the antenna polarization is also vertical, by definition. Throughout this book, we will generally assume that the antenna orientation is relative to the earth below the antenna, so that "up," "down," and "vertical" all coincide with the direction of gravity's pull directly beneath the antenna. Similarly, horizontal is normally 90 degrees away, parallel to the surface of the earth.

11.1 Vertical Dipole

To develop our understanding of vertical antennas, we start with a half-wave dipole, oriented vertically relative to earth, as in Fig. 11.1. We will assume that the antenna is many wavelengths above the earth's surface. The vertical dipole antenna is used in many locations where it is impossible to properly mount a horizontal

dipole or where a roof-or mast-mounted antenna is impossible to install because of logistics, a hostile landlord, or a homeowners' association.

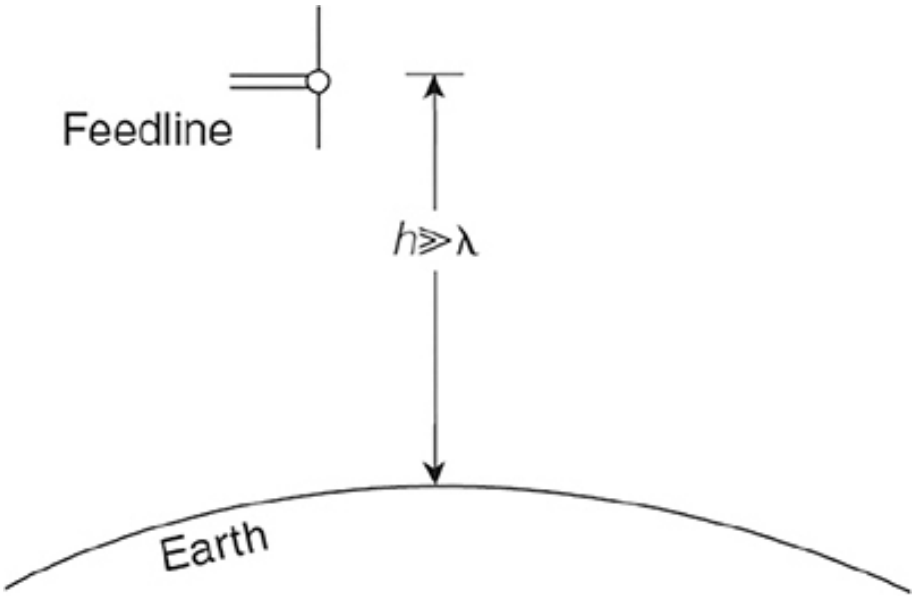


FIGURE 11.1 Vertical dipole far from the earth's surface.

Like the horizontal dipole, the approximate length of the vertical dipole is

$$L\ (\text{ft}) = \frac{468}{f(\text{MHz})} \quad (11.1)$$

where *f* is the operating frequency in megahertz.

Example 11.1 Calculate the length of a half-wavelength vertical dipole for operation on a frequency of 14.250 MHz in the 20-m amateur radio band.

Solution

$$L = \frac{468}{14.250} = 32.8\ \text{ft}$$

Each arm of the vertical dipole in this example is half of the calculated length, or 16.4 ft.

To avoid unbalancing the pattern of the dipole as well as the currents in the two sides of the transmission line, the feedline should extend away from the dipole at right angles for some distance, although HF band dipoles are likely close enough to earth ground or other structural grounds that the feedline geometry is unimportant.The lower side of the dipole is normally connected to the shield or outer conductor of an unbalanced feedline, such as coaxial cable or CATV hard line, and the upper side of the dipole to the center conductor, but this generally is not a requirement for proper operation.

A coaxial vertical is similar to the vertical dipole in that it uses a pair of vertical radiator elements. Such an antenna is an extreme case of imbalance because the feedline comes away from the dipole completely in line with it. In this antenna the radiating element that is closer to the ground is coaxial with the transmission line as shown in Fig. 11.2. The use of a transmission line choke immediately beneath the coaxial sleeve will minimize the tendency of the feedline to act as if it were part of the antenna. One form of choke consists of an appropriate number of turns of the coaxial cable looped through one or more ferrite cores.

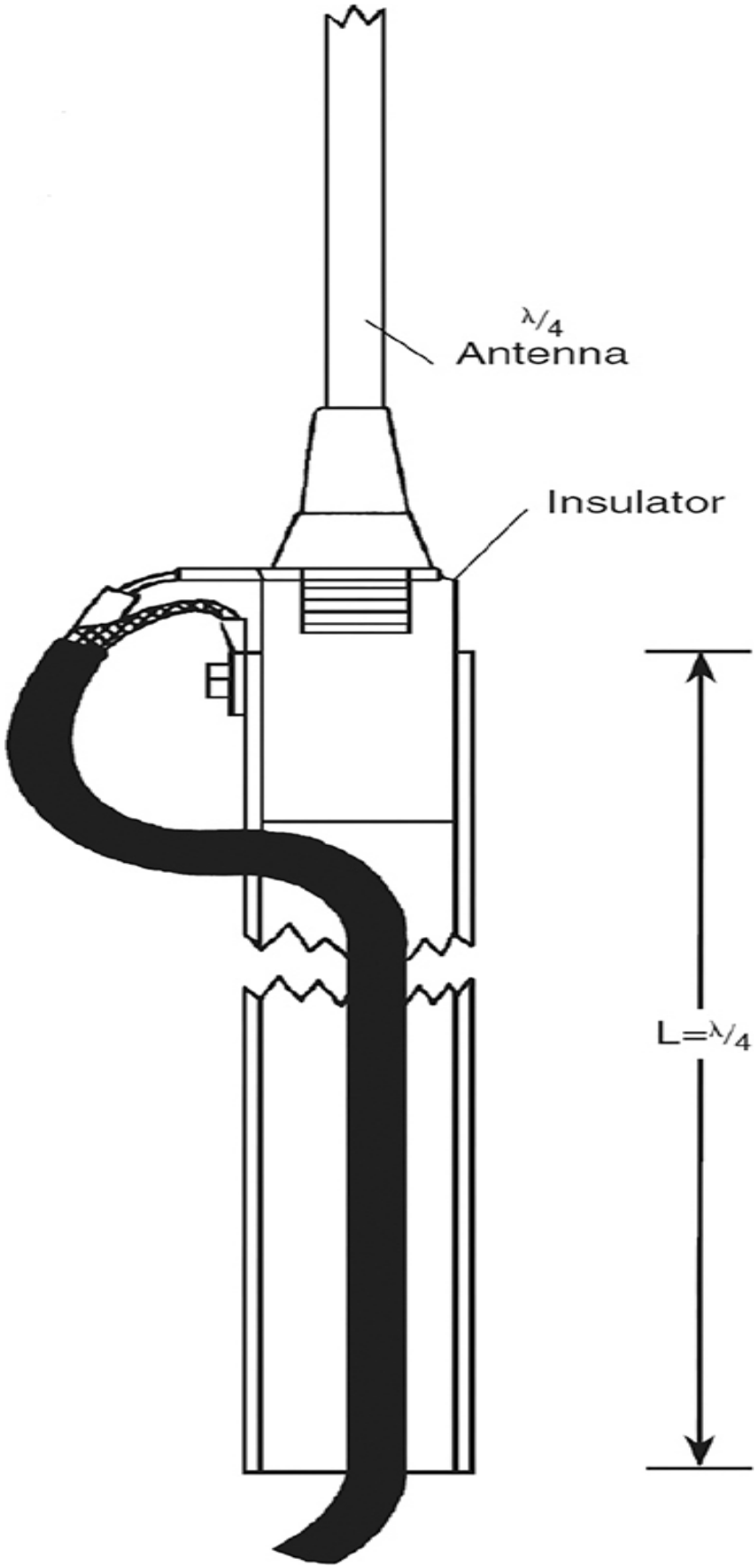

λ/4
Antenna
Insulator
L=λ/4

FIGURE 11.2 Coaxial vertical dipole.

The coaxial vertical antenna was once popular with CB operators. A CB antenna for 11 m can be cut for 10-m operation at a slightly higher frequency. Conversion for bands below 27 MHz is a little more difficult; most likely only the insulator and mounting assembly are salvageable. Adjacent sizes of 0.058-in-thick aluminum tubing are designed such that the inside diameter (ID) of the larger piece is a slip-fit for the outside diameter (OD) of the smaller piece. Adjacent sizes of aluminum tubing can be connected without the need for special couplers. Salvaged insulator assemblies can be cut off with 6 to 10 in of the former radiator and sleeve and new radiators can be created from telescoped tubing sections.

11.2 Grounded Vertical Monopole

The vertical dipole is a good antenna and has formed the basis for many VHF and UHF broadcast antennas for decades. It can be built using exactly the same dimensions as required for the horizontal dipole of Chap. 8. The radiation pattern is as shown in Fig. 11.3 as long as it is kept far (many wavelengths) from conducting objects with dimensions comparable to or larger than the antenna. That turns out to be fairly easy (for fixed stations, at least) at VHF and UHF frequencies and not so easy at MF, HF, and microwave frequencies.

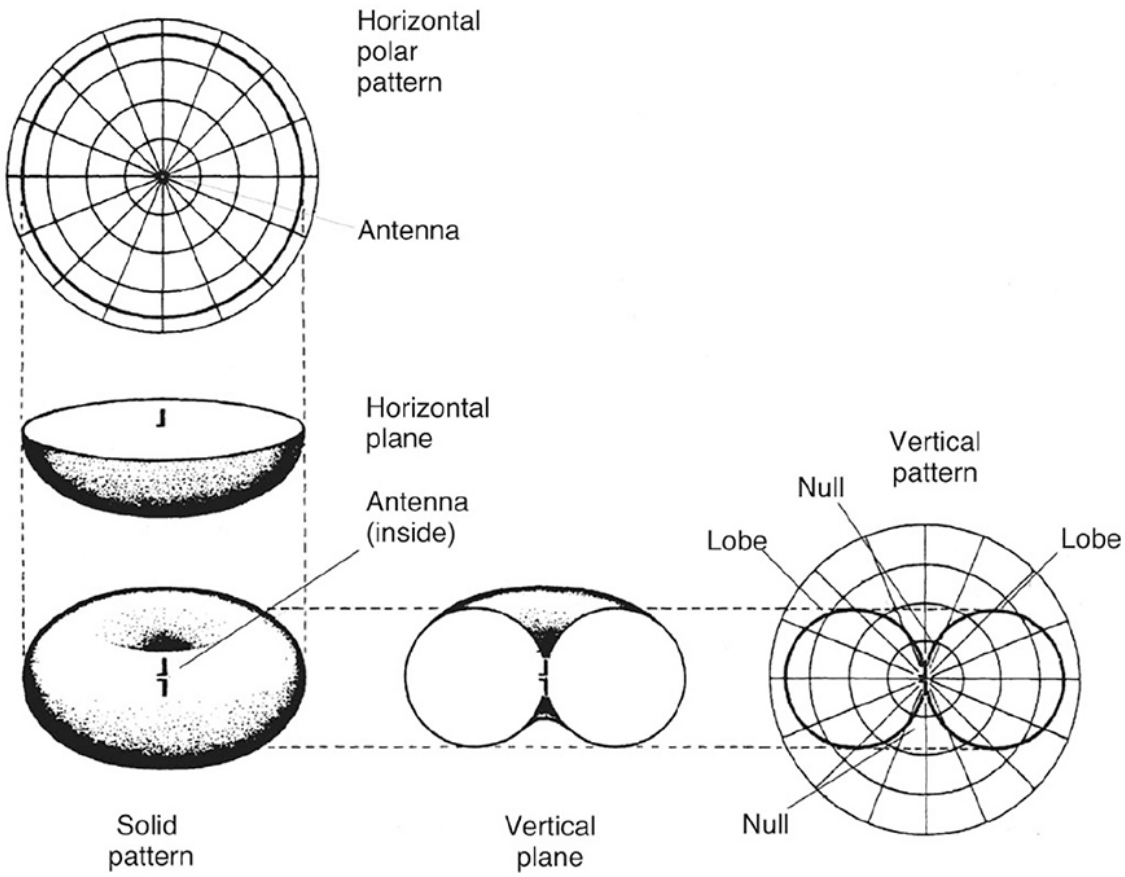


FIGURE 11.3 Vertical antenna radiation pattern.

The mechanical difficulty of keeping vertical dipoles far from conducting surfaces makes the grounded quarter-wave vertical useful. To suspend a vertical dipole 80 m above the earth requires a support more than 150 ft high and bringing the transmission line away from the center of the dipole at right angles would present a completely separate challenge.

We saw in Chap. 7 that we can use the earth beneath us as a substitute for one side of a dipole. As the saying goes, "If you can't beat 'em, join 'em!" Over the range of frequencies where we would have to incur tremendous cost and effort to lift a vertical dipole into the air high enough to avoid unbalanced currents in the two sides, we typically choose to use a quarter-wave monopole operated against the ground beneath.

For the MF and lower HF bands (up through perhaps 5 MHz or so), the simplest approach is to mount the vertical on the earth's

surface. For still higher frequencies, lifting the vertical into the air (to get above any nearby obstructions) works well with an artificial ground plane consisting of a finite number of radials below the antenna.

Figure 11.4 shows the geometry of the vertical monopole antenna. An RF source is applied at the base of a radiator of length *L*. Most verticals are a quarter-wavelength ($L = \lambda/4$) long, but other lengths are possible.

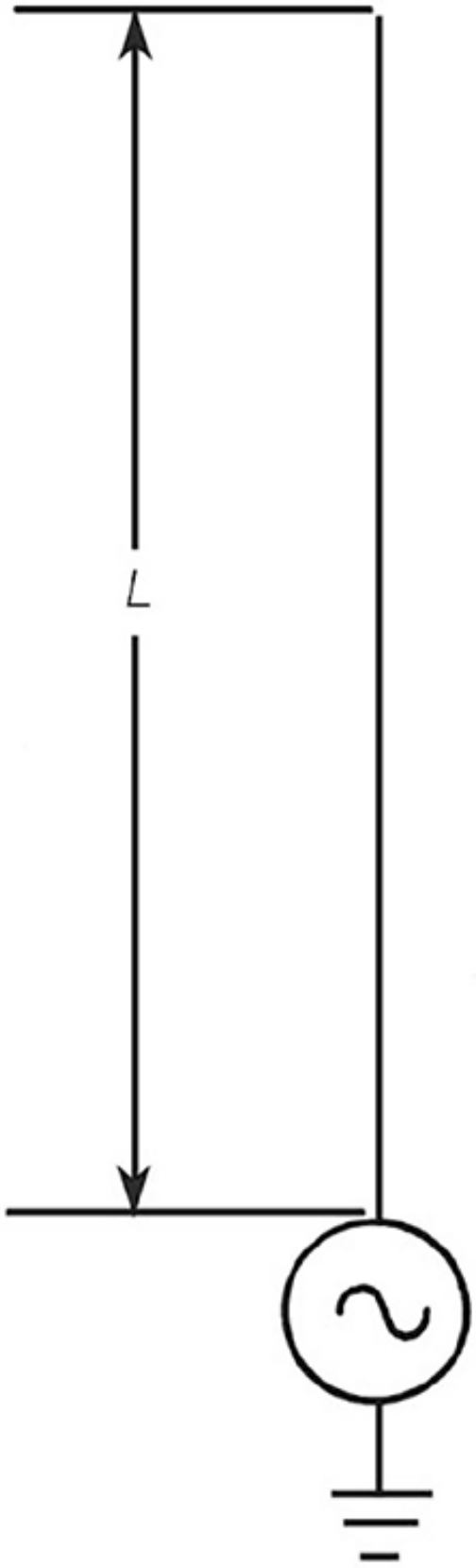
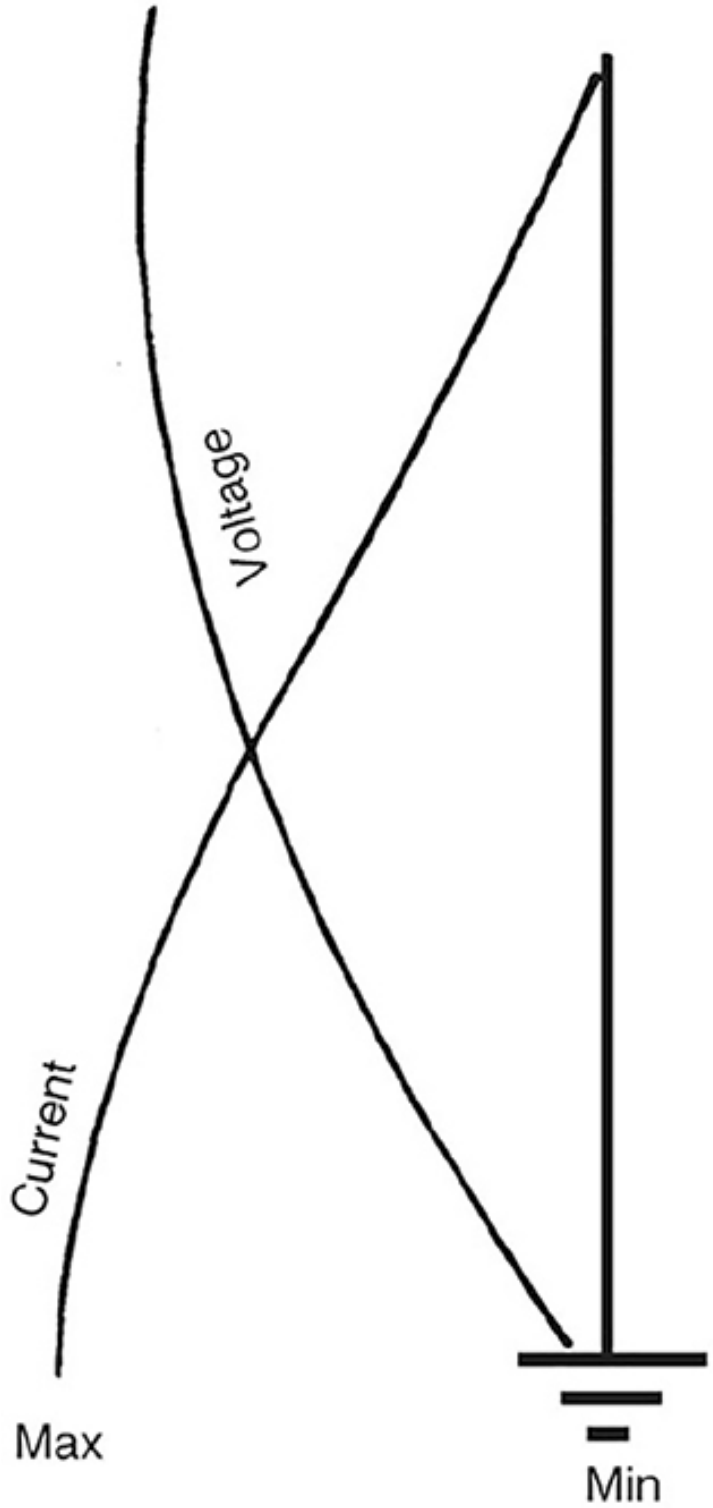

L

FIGURE 11.4 Basic vertical monopole.

Figure 11.5 shows the current and voltage distribution in a quarter-wavelength vertical. Like the half-wave dipole, the $\lambda/4$ vertical is fed at a current node, so the feedpoint impedance is at a minimum—typically less than 37 Ω, depending upon nearby objects, diameter of the radiating element, and other factors.



F IGURE 11.5 Current and voltage distribution along a $\lambda/4$ vertical.

Figures 11.6 and 11.7 show two configurations for a $\lambda/4$ HF vertical antenna. In Fig. 11.6, the radiating element is mounted at ground level but electrically insulated from the ground and fed with a 52-Ω coaxial cable. The inner conductor of the coaxial cable is connected to the radiator element, while the cable shield is connected to ground at the base of the vertical. For a $\lambda/4$ radiator, the feedpoint impedance will be lower than 52 Ω, but in most cases, the resulting voltage standing wave ratio (VSWR) is an acceptable tradeoff for simplicity without a matching network or ATU. If the antenna has a feed-point impedance of 37 Ω, the VSWR will be 52 Ω/37 Ω, or 1.4:1 at the design frequency.

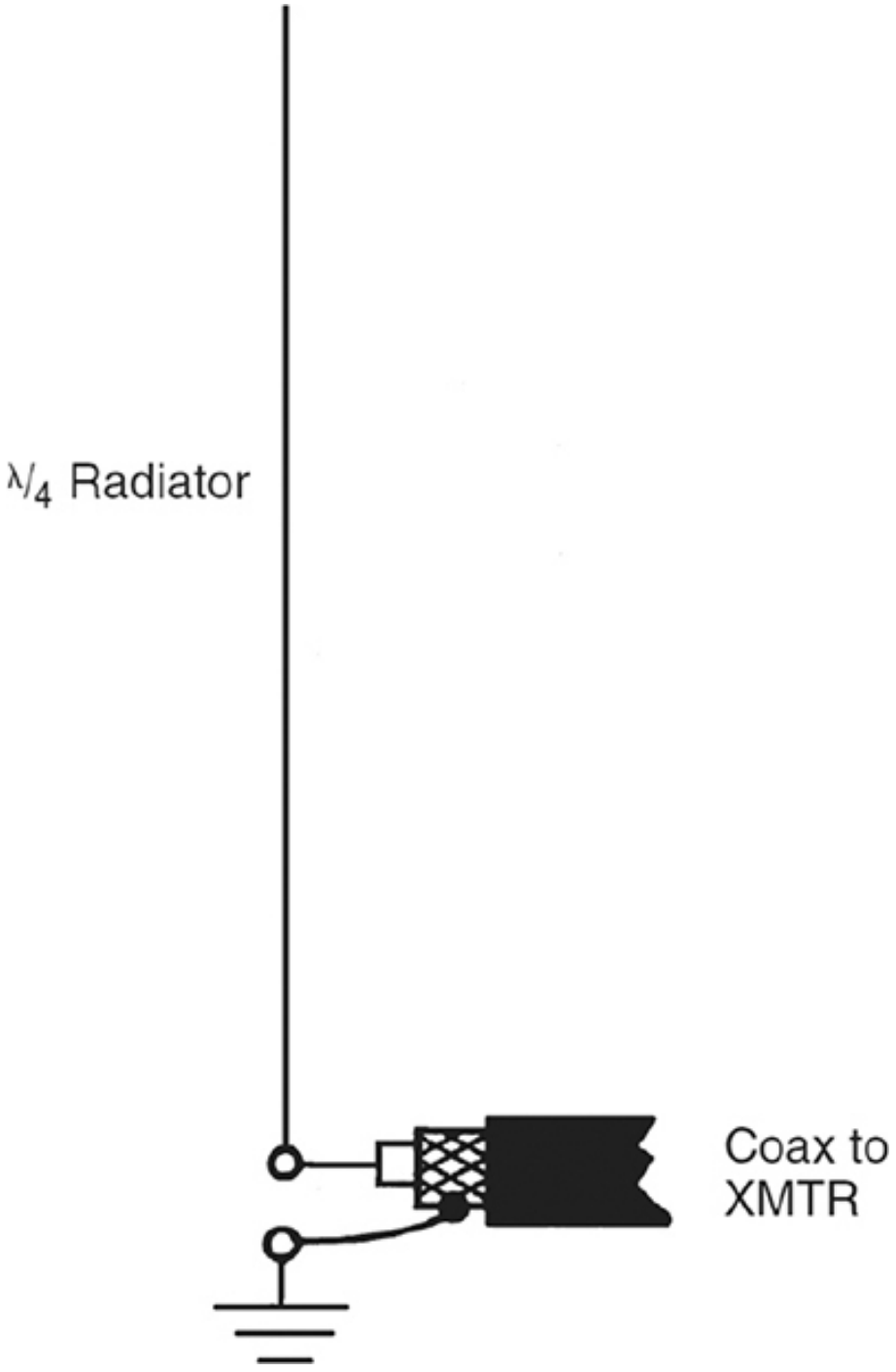


FIGURE 11.6 Simple coaxial-fed vertical antenna.

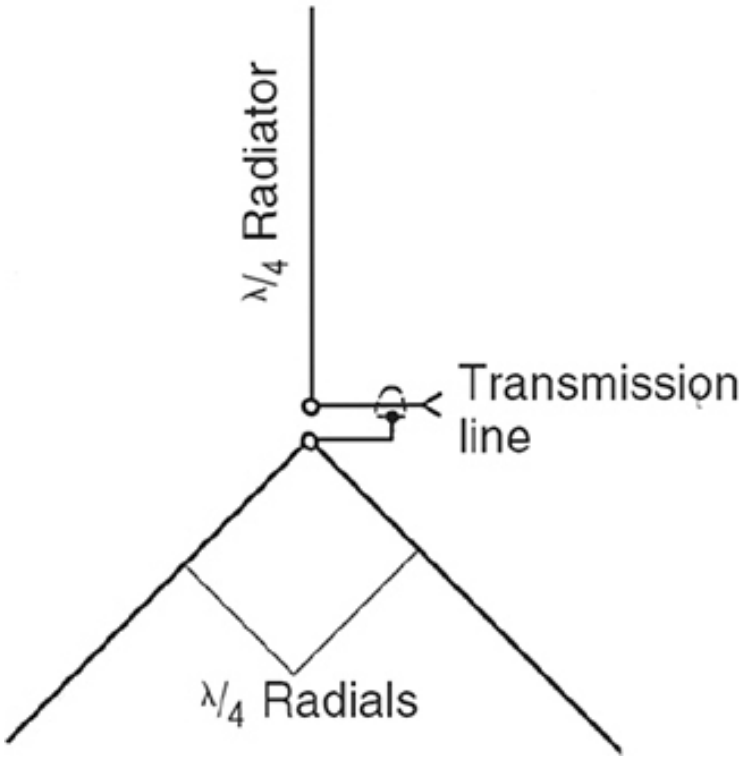


FIGURE 11.7 Mast-mounted vertical uses radials as a counterpoise ground.

An elevated ground-plane vertical is shown in Fig. 11.7. This antenna is as popular as the ground-mounted quarter-wave vertical, especially on 40 m and above. Amateurs and CB operators find it easy to construct this form of HF antenna because the lightweight vertical can be mounted at a reasonable height (15–60 ft) using a television antenna slip-up telescoping mast that is reasonably low in cost. As discussed earlier, this antenna replaces the lower half of a vertical dipole with an artificial ground comprised of quarter-wavelength radials.

The radials of a ground-plane vertical antenna are typically installed at some vertical angle between horizontal (0 degrees) and around 45 degrees below horizontal. Figure 11.7 is an example of the latter case, showing a ground-plane antenna made with "drooping" radials. Similarly, Fig. 11.8 is an example of a ground-plane vertical having perfectly horizontal radials. Sometimes the

feedpoint VSWR of the antenna can be improved by adjusting the degree of droop in the radials.

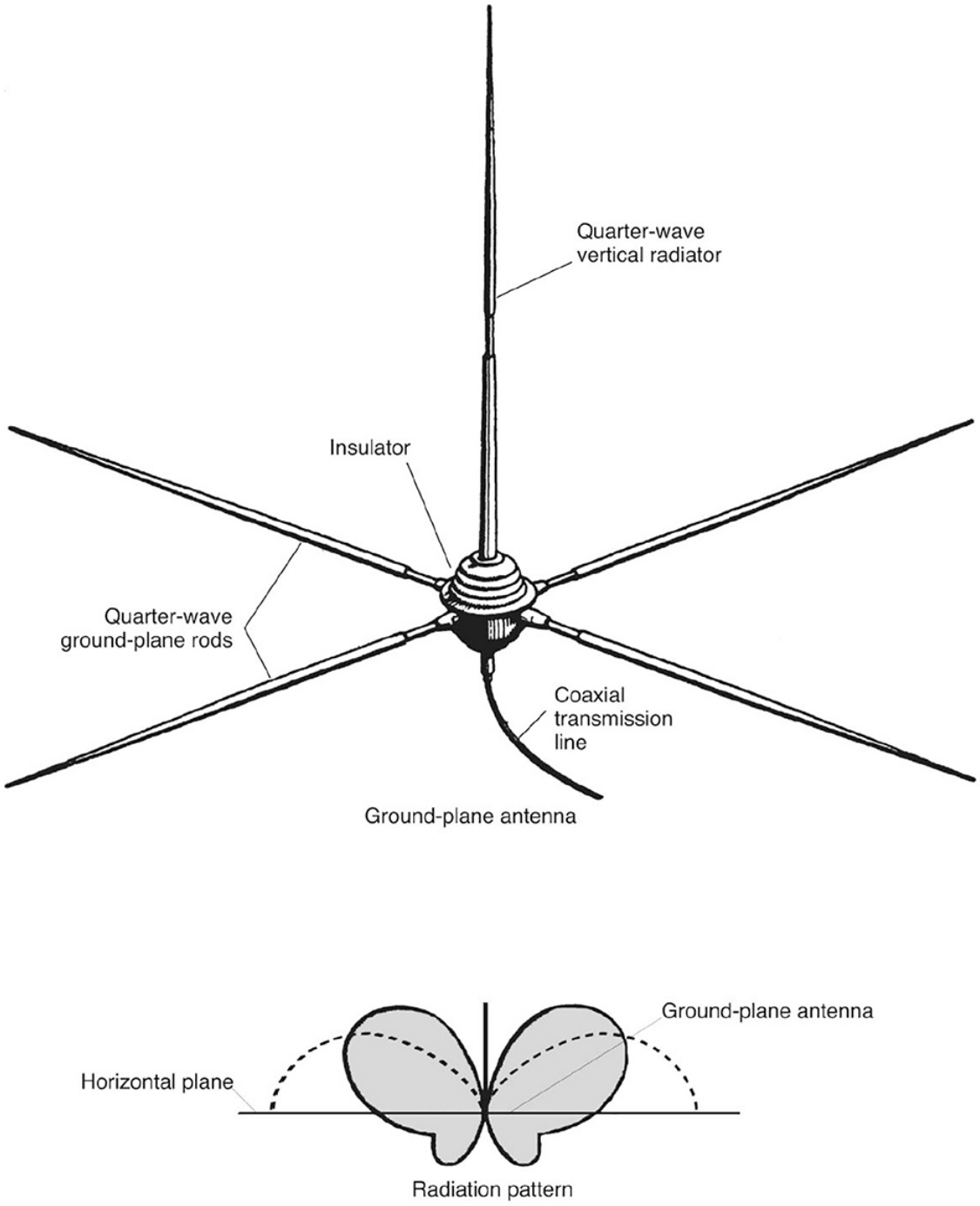


FIGURE 11.8 Ground-plane vertical antenna.

Unlike ground-mounted verticals, the imperfect ground plane formed by the radials of Fig. 11.8 does not preclude the existence of a far E-field below the horizontal plane of the radials, as suggested by the (solid line) pattern in the figure.

11.3 Non-Quarter-Wavelength Vertical Monopoles

If we gradually increase the length of our grounded monopole, we observe the same flattening of the donut shape as we saw with the dipoles of Chap. 8. Lengthening the monopole beyond $\lambda/4$ results in additional RF energy in the low-angle main lobe of the pattern, in exchange for reduced radiation at higher elevation angles. Figure 11.9 shows the approximate patterns for vertical monopoles of three different lengths: $\lambda/4$, $\lambda/2$, and 5⁄8λ. The quarter-wavelength antenna has reasonable gain over the widest range of elevation angles, but also the lowest maximum gain of the three cases. The 5⁄8-wavelength antenna, which is the grounded monopole equivalent of the extended double Zepp (EDZ), enjoys both the lowest angle of radiation and the highest maximum gain.

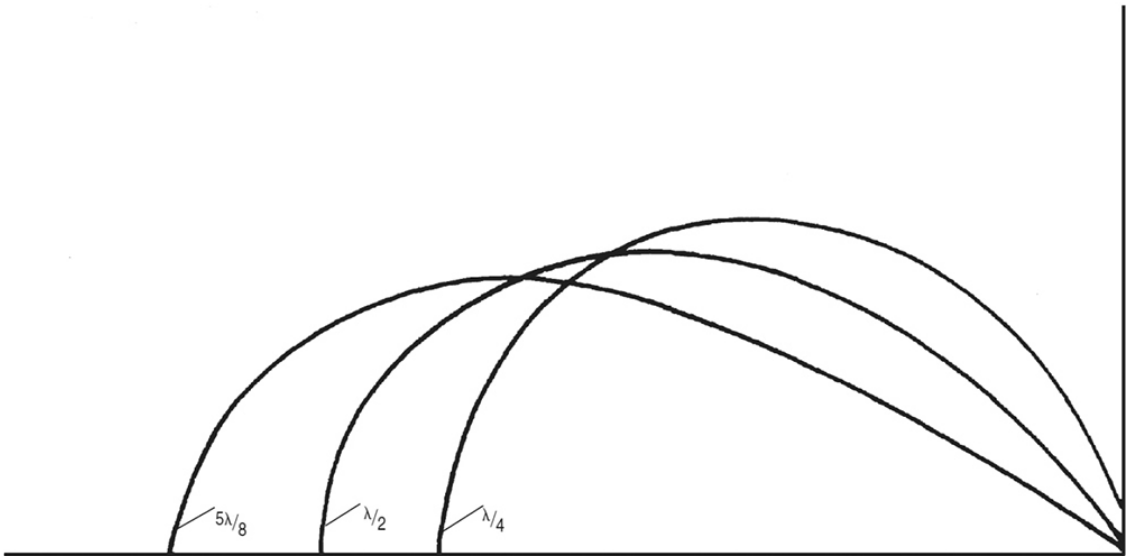


FIGURE 11.9 Vertical radiation pattern versus a function of element length over perfect ground.

The patterns shown in Fig. 11.9 assume a perfectly conducting ground surrounding the antenna. The effect of real-earth ground losses is to eliminate any hint of the radiated field at distant receiving sites for elevation angles near the horizon (Fig. 11.10).

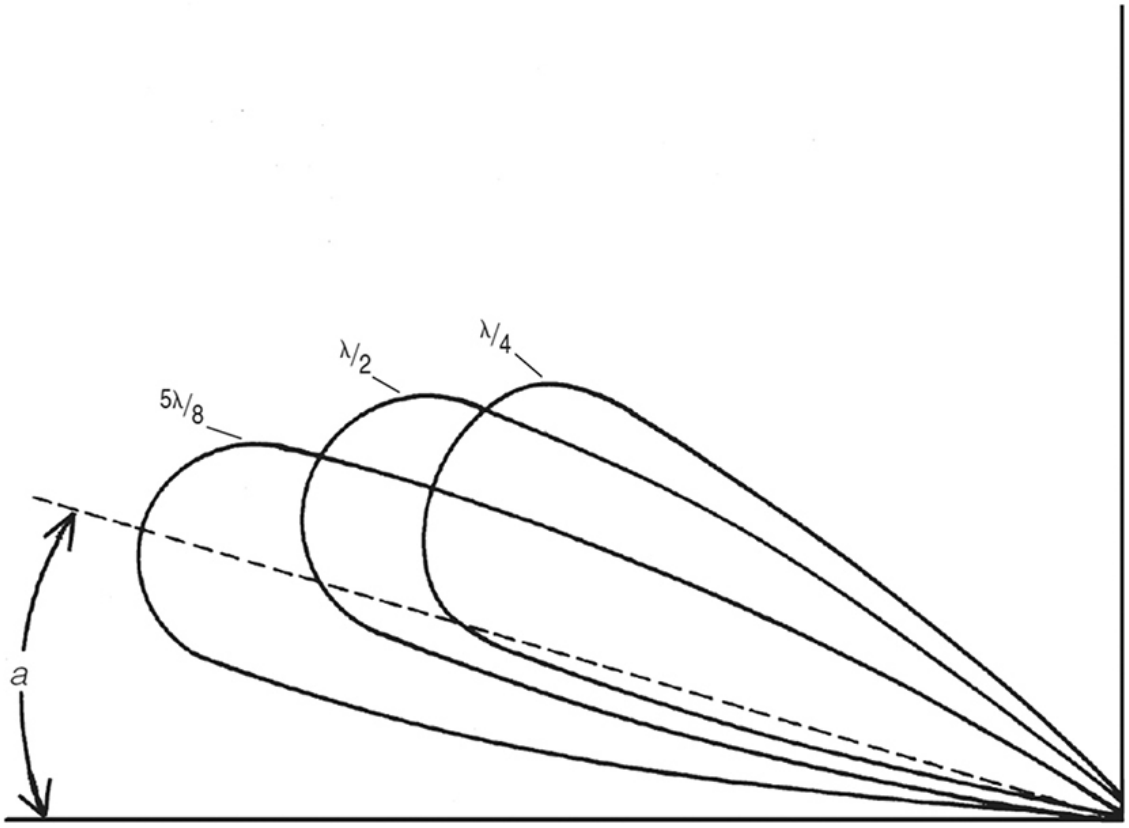
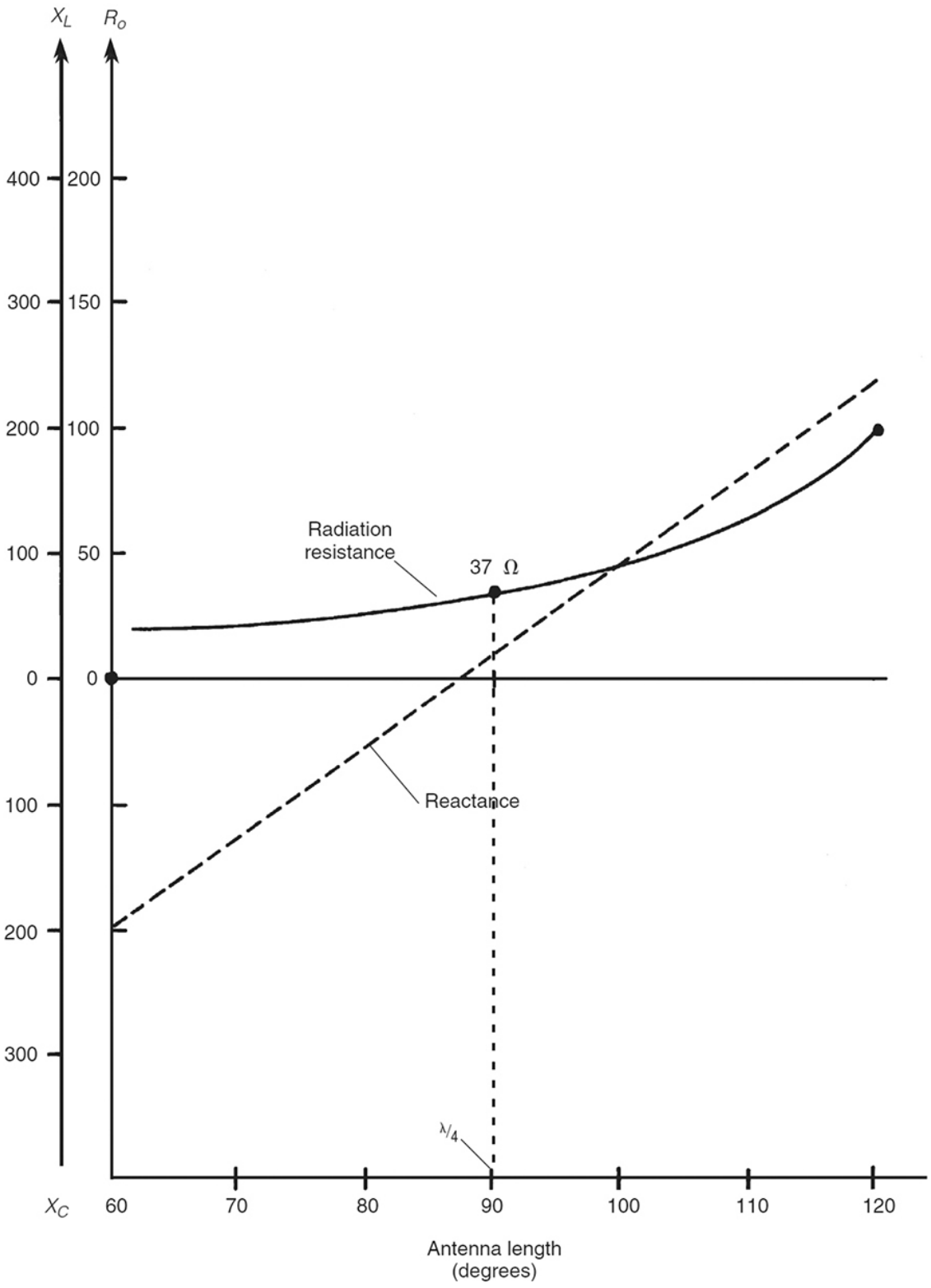


FIGURE 11.10 Accounting for ground losses over real earth.

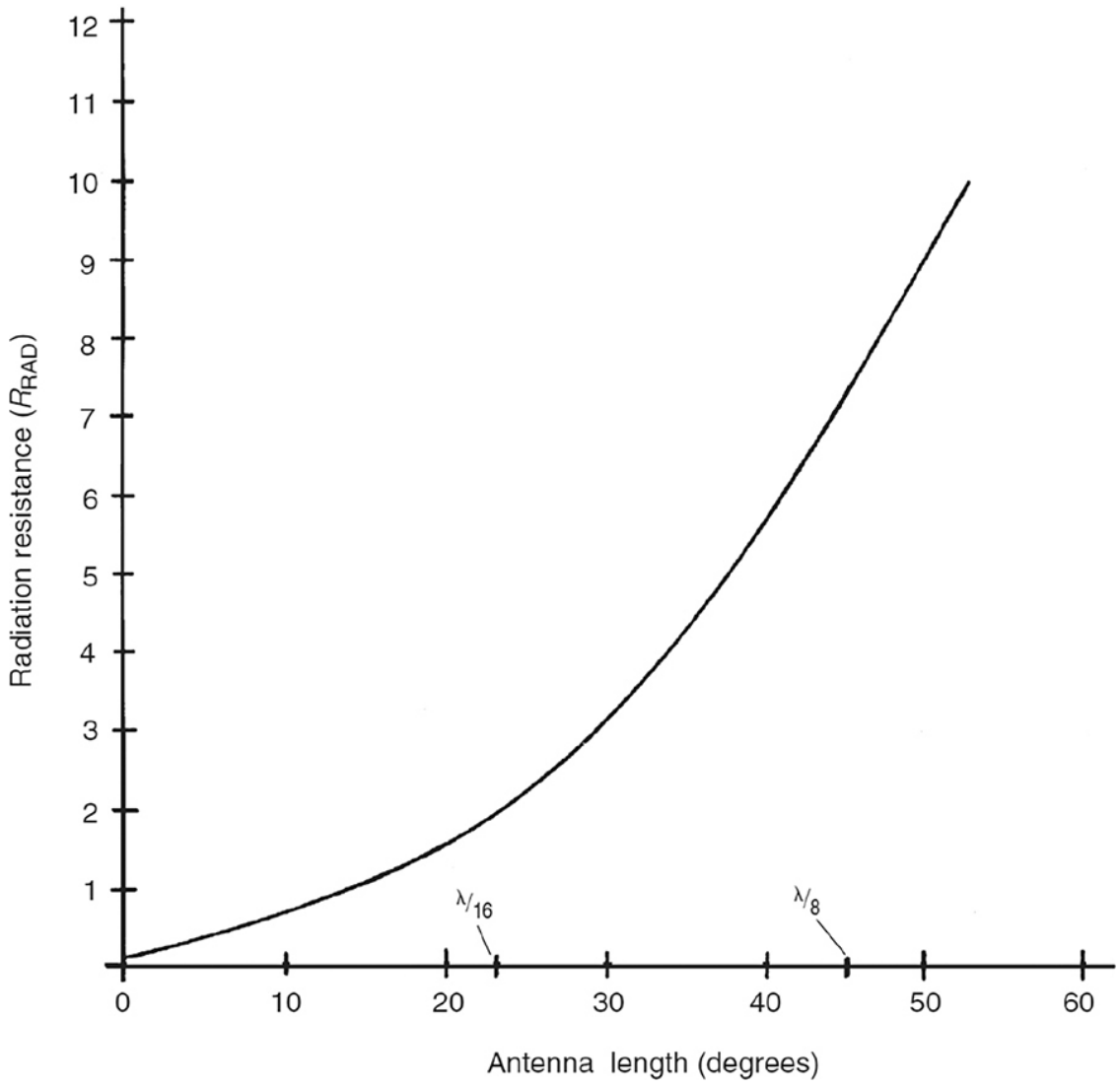
The feedpoint impedance of a grounded monopole is a function of the length of the radiator. For the standard quarter-wavelength antenna over perfect ground, the feed-point radiation resistance is roughly 37 Ω (i.e., half that of a

$\lambda/2$

dipole), with only a small reactance component. Figure 11.11 plots the reactive and resistive components of the vertical monopole's feedpoint impedance for radiator electrical lengths from 60 degrees to 120 degrees (360 degrees is one wavelength), while Fig. 11.12 shows the radiation resistance for antennas shorter than $\lambda/6$ (60 degrees).

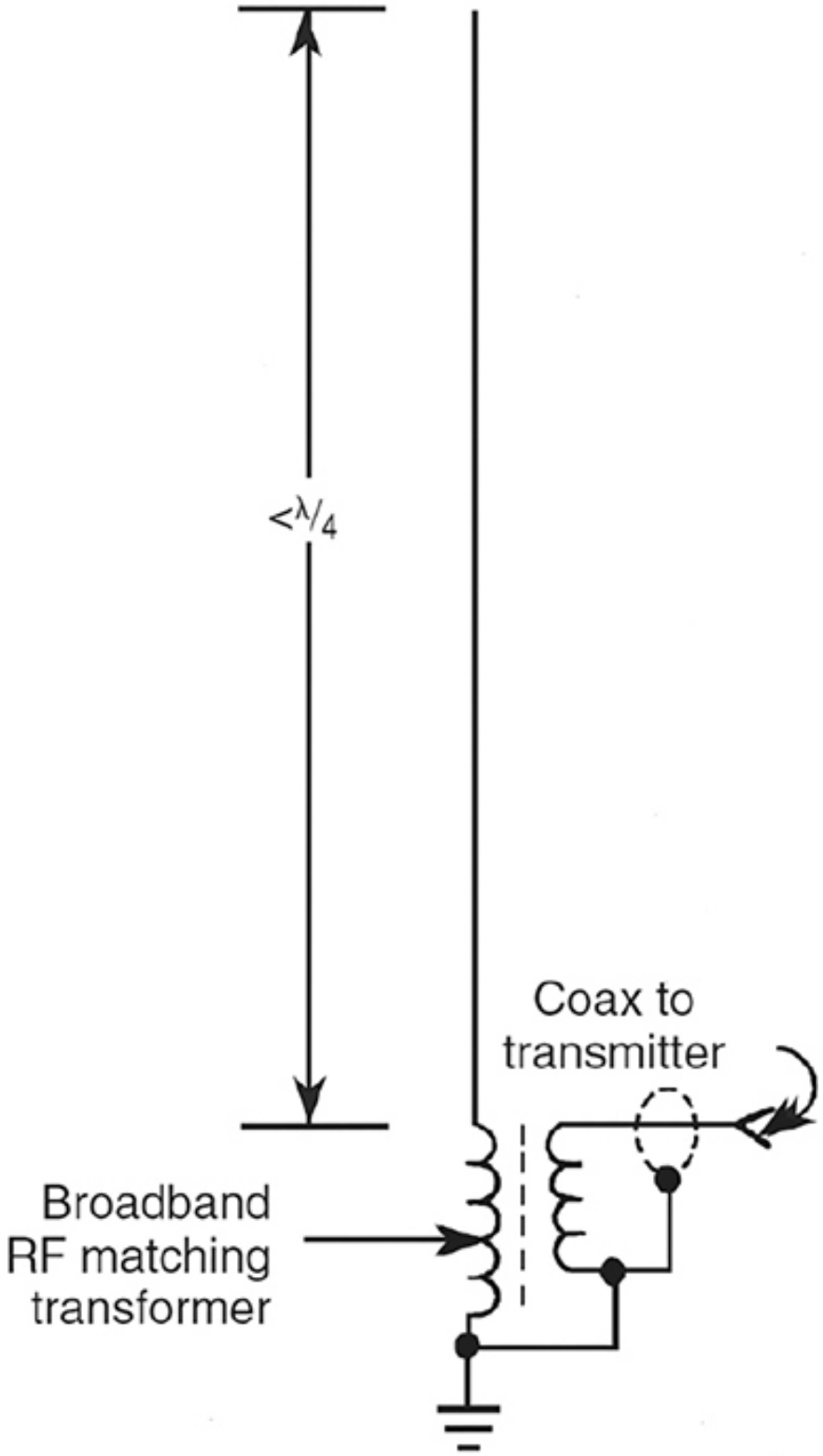


F IGURE 11.11 Antenna impedance as a function of antenna length.



F IGURE 11.12 Radiation resistance as a function of antenna length.

The radiation resistance for a short antenna is normally very small. A monopole that is 30 degrees long ($30/360 = 0.083\lambda$) has a resistance of only 3 Ω. General practice for such antennas is to use a broadband impedance-matching transformer to raise the impedance of such antennas to a higher value (Fig. 11.13). With such low radiation resistances, it is difficult to avoid dissipating most of the transmitter power in lossy grounds, matching network components, and the connections themselves.



F IGURE 11.13 Impedance matching with a broadband transformer.

11.4 Ground Systems for the Grounded Monopole

Earlier in this chapter, we discussed how the vertical antenna radiates equally well in all azimuthal directions. Over the years, some have jokingly said that the vertical radiates equally poorly in all directions. Which of these views is correct? Either can be true. The grounded vertical monopole antenna works well only when placed over a good ground system (this does not apply to a vertical dipole in free space, many wavelengths away from the nearest conductors). For a ground-mounted vertical the ground system provides an electrical return for antenna currents and replaces the lower half of a dipole. To the extent that the ground return is lossy or incomplete in its coverage of the radiation field of the vertical, a significant ground loss resistance is added to the total resistance seen by the RF energy delivered to the antenna terminals by the transmitter and feedline. Ground losses and antenna radiation resistance are in series and the transmitter output power is divided between the radiation resistance of the antenna and the loss resistance of the ground return system. With radiation resistances in the 2- to 37-Ω range, it does not take much loss resistance to steal most of your output power.

Ground return loss is one of the primary reasons mobile installations generally have poorer performance than home station antennas. Mobile whips (usually operated as grounded verticals) are so much shorter than a quarter-wavelength that their input impedance is very low—usually just a few ohms, at best.

The usual way to provide a good ground for a ground-mounted vertical is to use a system of radials such as those shown in Fig. 11.14. There are 16 radials equally spaced to cover the full circle around the antenna. At one time it was commonly thought that each radial was to be a quarter-wavelength. But radials in direct contact with the ground are sufficiently detuned from their free-space characteristics that an exact length is unimportant. Instead, what matters is that enough radials are installed that their far ends

are typically no more than 0.05 λ apart. It is even possible that, up to a point, the shorter your radials are the fewer you need, from the perspective of antenna input resistance. On the other hand, the limited extent of the ground system can lead to higher ground losses. As a rule of thumb, if you have a fixed amount of wire to apply to your system of radials, a few long radials are not as helpful as a greater number of shorter ones.

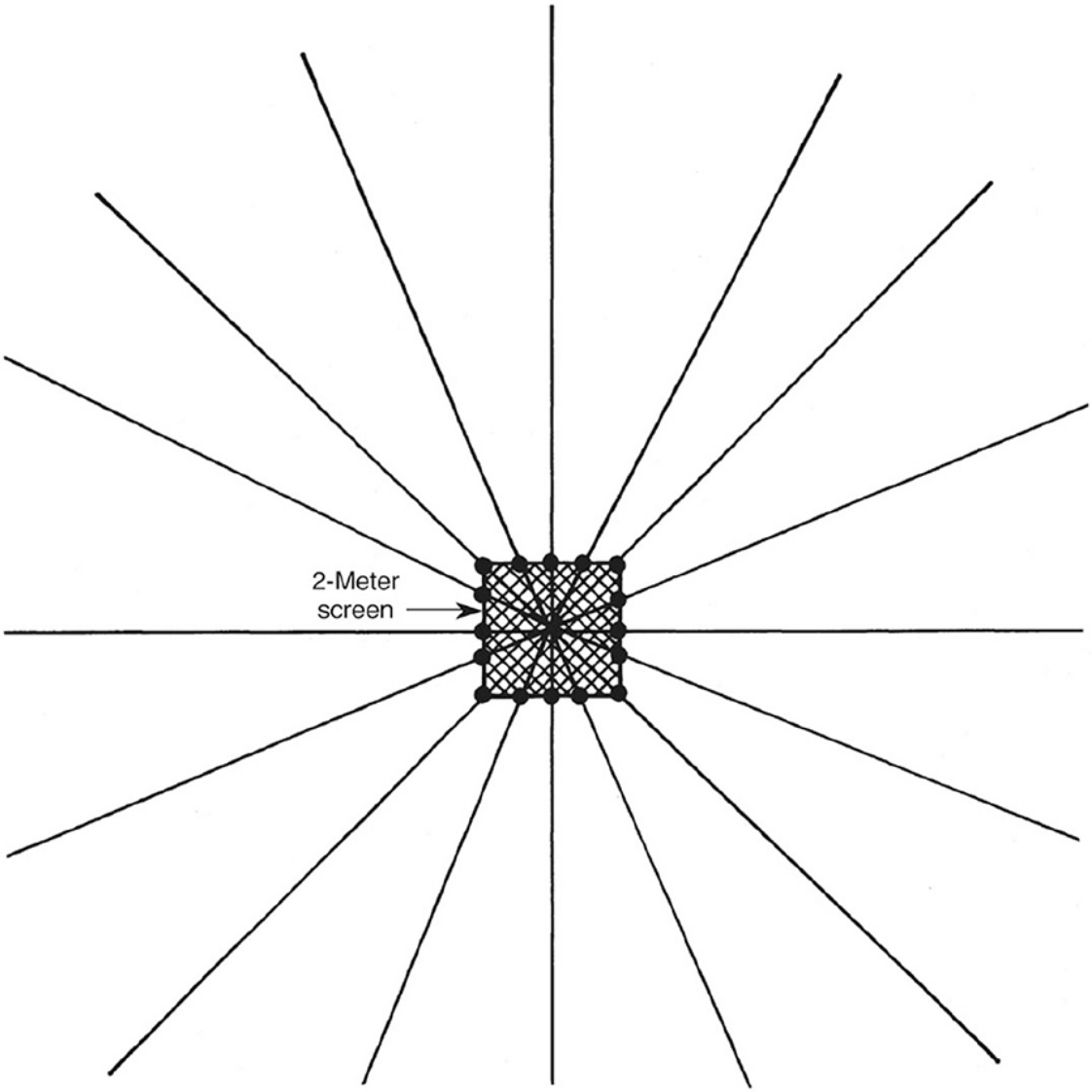


FIGURE 11.14 Comprehensive ground system for a vertical antenna.

Keeping the ends of the radials at 0.05 *λ* apart leads to a formula for the number of radials required:

$$N = L \times 128/\lambda$$

where *L* is the radial length. For quarter-wavelength radials, *N* is about 32. This is just an estimate. Increasing the spacing from 0.05 *λ* to 0.10 *λ*, for instance, means a slight increase in ground system losses and a corresponding decrease in signal strength at the distant receiving station in return for needing to install half as many radials. There is also nothing magic about making the radials equal in length. If you are boxed in by property boundaries, a city lot, or structures that get in the way, simply do the best you can. Radials are important to the efficient operation of verticals, but many different geometries can lead to good results.

The taller the vertical, the longer the radials should be for optimum antenna system efficiency. A taller vertical tends to induce return currents into the earth at a greater distance from its base than a shorter vertical. Again, as the radials are lengthened, the number of them should increase. Radials should be around the same length as the vertical's physical height.

Transmitter output power is divided between feedline losses, antenna and ground system losses, and antenna radiation resistance. Of those, only power dissipated in the antenna radiation resistance contributes to the received signal strength far away. The lower the radiation resistance of the antenna, the more attention that should be paid to reducing feedline losses and losses in the ground system. For optimal performance, a radial field used in conjunction with a short transmitting vertical (less than $\lambda/4$ tall) requires more wire than one for a full-size $\lambda/4$ or taller monopole if the radiation efficiencies of the two antennas are to be comparable.

Prior to World War II, scientists and engineers performed a series of tests on radial fields for AM broadcast stations. Figure

11.15 is typical of the kind of information that these experiments provided. While some folklore regarding ground systems may remain, no more than 50 to 60 radials is adequate for a quarter-wave ground-mounted vertical and the exact length of radials is immaterial.

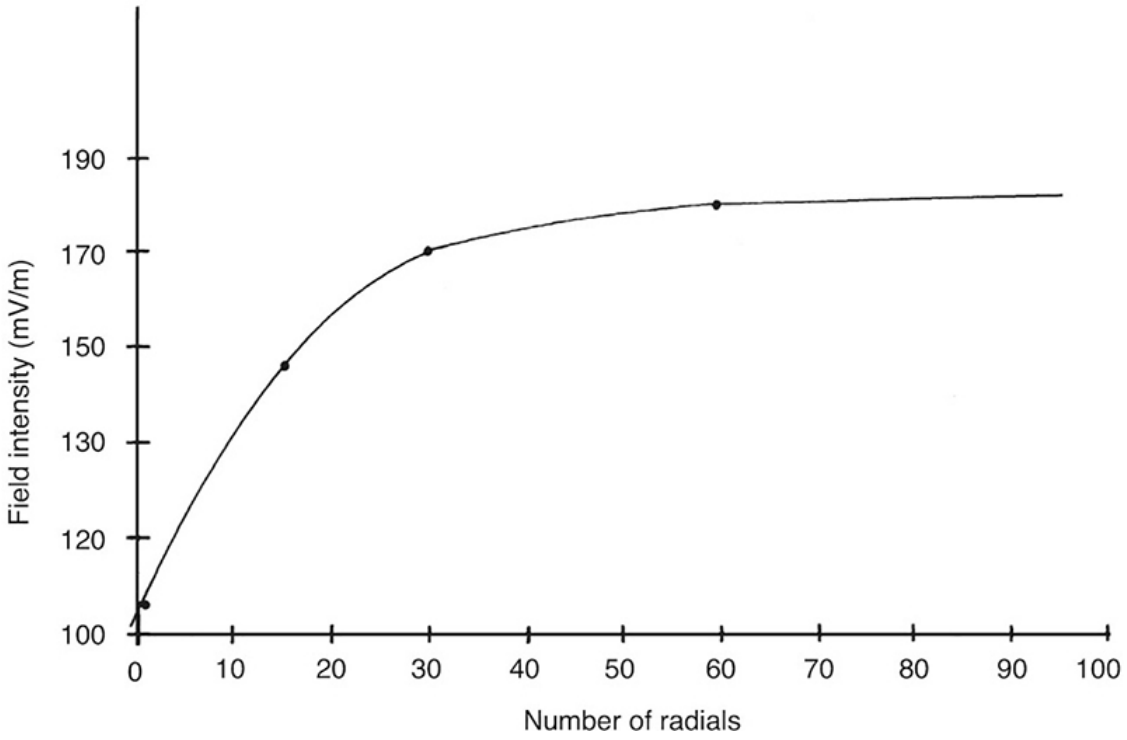


FIGURE 11.15 Typical test results for a vertical antenna as a function of the number of radials in the ground system.

One of the least important characteristics of radials is wire size. Antenna return currents are split among the many radials, and the conductivity of even tiny wire sizes is superior to that of soil. Bigger issues are survivability, mechanical strength, and visibility. Wires laid on the ground or on grass are subjected to a variety of stresses, whether from being stepped on or tripped over or after easing their way into the sod. Some types of terrain do not lend themselves to burial techniques and highly visible insulation may be important to prevent unsuspecting visitors from tripping over them.

Caution *If you decide to use an above-ground radial system, be sure to include measures to prevent people from tripping over it or (if it is an elevated system) running into it. There may even be liability implications when a visitor gets injured.*

The return currents in radials result from displacement currents flowing in the capacitor formed by the vertical element and the ground system. Displacement currents (see Chap. 3) are RF alternating currents that do not require a conducting medium for their existence. Consequently, radials can be either insulated or bare, although bare metals used outdoors will develop an oxide insulating layer on them sooner than insulated conductors. Radials can be copper or aluminum. Many amateurs have formed their radial fields out of scrap wiring from multiple sources, while others have managed to create additional wire by separating multiconductor cable into individual strands.

Because of the orientation of the electromagnetic fields around the vertical, only the ground resistance along radial lines from the base of the vertical is important. That is to say, there is little merit to connecting the tips of the radials together or adding other nonradial wire interconnections to the basic radial system. The same is true of ground rods, which are virtually useless in improving the performance of grounded verticals.

Typically, radials are connected together at the base of the antenna and the ground side of the transmission line is connected to this system. When installing radials, avoiding burying them too far below the surface of your soil; the farther beneath the surface they are, the less effective they become.

While a radial system can be designed to serve as part of your lightning protection and grounding system (Chap. 33), it does not have to be. Conventional solder connections may eventually be destroyed outdoors and will vaporize in the path of a lightning strike, but as long as it's understood they're not "forever," there's nothing wrong with soldering your copper radials to a copper bussbar or flexible plumber's tubing circling the base of your vertical (see Fig. 11.16). To avoid corrosion, copper should not be

in direct contact with galvanized tower components. Appropriate electrical clamps or a sheet of stainless steel between copper and tower legs should be used.



FIGURE 11.16 Example of radial connections at the base of a vertical antenna.

Some amateurs place a copper wire screen at the center of the radial system. For it to be useful, the required size is a function of the vertical's height and operating frequency. A possible concern is the creation of unwanted rectifying junctions at metal interfaces after extended exposure to the elements at each location where two wires in the screen cross.

11.5 Feeding the Ground-Mounted Vertical

The feedpoint resistance of a vertical monopole over perfect ground is one-half that of a dipole in free space. The input impedance of an ideal
$\lambda/4$
vertical radiator fed against ground is about 37 Ω. For many applications, this is close enough to 52 Ω that no further matching is required, and 52-Ω coaxial cable or hard line can be connected directly to the antenna. If minimizing the standing wave ratio (SWR) on the transmission line is important, the matching network can be an antenna coupler (protected from the weather by some form of housing) or a simple lumped element L-network. In either case, the purpose of the coupler or network is to raise the resistive part of the impedance seen by the transmission line to 50 or 75 Ω while canceling the reactive portion of the antenna impedance. The coupler settings or L-network component values will normally be suitable for only a narrow range of frequencies. The values used for 40 m, for example, will not be useful on other amateur bands. Direct feed implies that the base of the vertical is not in electrical contact with the ground system beneath it. This is often called series feed.

Alternatively, the base of the vertical radiator can be tied to the ground, and the resulting antenna fed through a tap point on the vertical section some distance above the base. Known as shunt feed, many forms of this have been devised and are discussed elsewhere in this book. In the shunt-feed arrangement, the connection partway up the vertical element functions like an autotransformer. Since the tip of a quarter-wave vertical is a high-impedance point (no current, high voltage) and the base is a low-impedance point (high current, low voltage), the impedance observed along the vertical element gradually increases as one moves the feedpoint from the base toward the tip. The shunt-feed network generally allows for close matching of the feedpoint impedance to an arbitrary transmission line impedance.

Three methods of shunt-feeding a grounded vertical antenna are delta, gamma, and omega. The shunt feed is useful when the resistive component of the antenna impedance on the centerline of a balanced element or the junction of the vertical element and ground is significantly lower than the characteristic impedance of the transmission line feeding it.

The delta feed system is shown in Fig. 11.17. A taut feed wire is connected between a point on the antenna, which represents a specific impedance on the antenna, and an antenna tuner. Although you might think that the sloping feed wire would change the antenna pattern, the distortion is minimal.

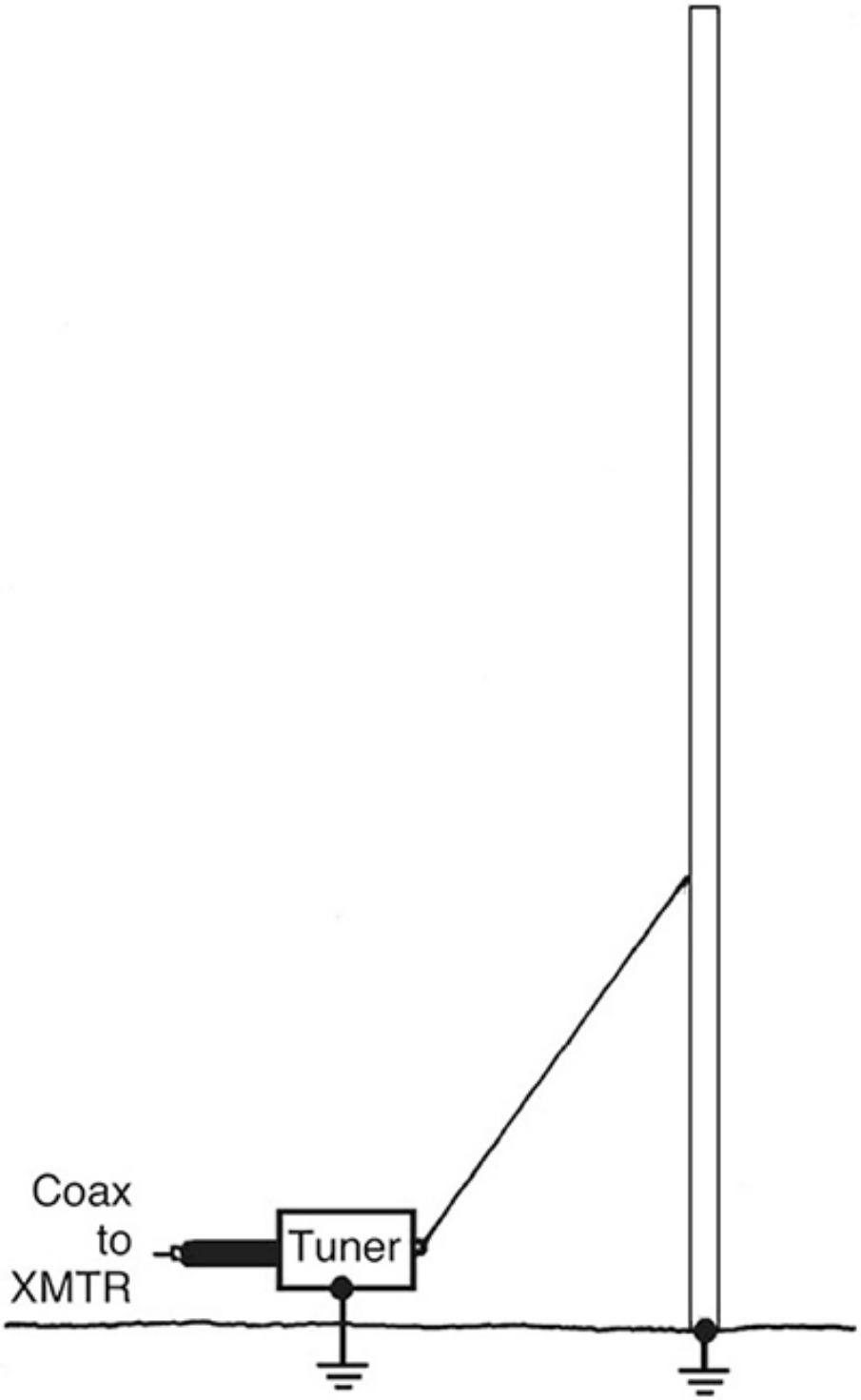


FIGURE 11.17 Delta-fed grounded vertical.

The gamma feed system is shown in Fig. 11.18. This method is used by amateurs to feed Yagi beam antennas (see Sec. 14.7.2).

The gamma match and its cousins use distributed components to accomplish the same impedance-matching functions a coupler or L-network does in the series-feed case. The critical dimensions for a gamma match are the distance from the base or centerline of the antenna to the tap point on the antenna (i.e., the active length of the gamma rod) and the spacing of the gamma rod away from the antenna element. The lower end of the gamma rod is not grounded; rather, it is connected to the hot side of the transmission line. It is important that the rod not be anywhere near a quarter-wavelength, or it would become a vertical antenna in its own right, and in fact would resemble the J-pole antenna.

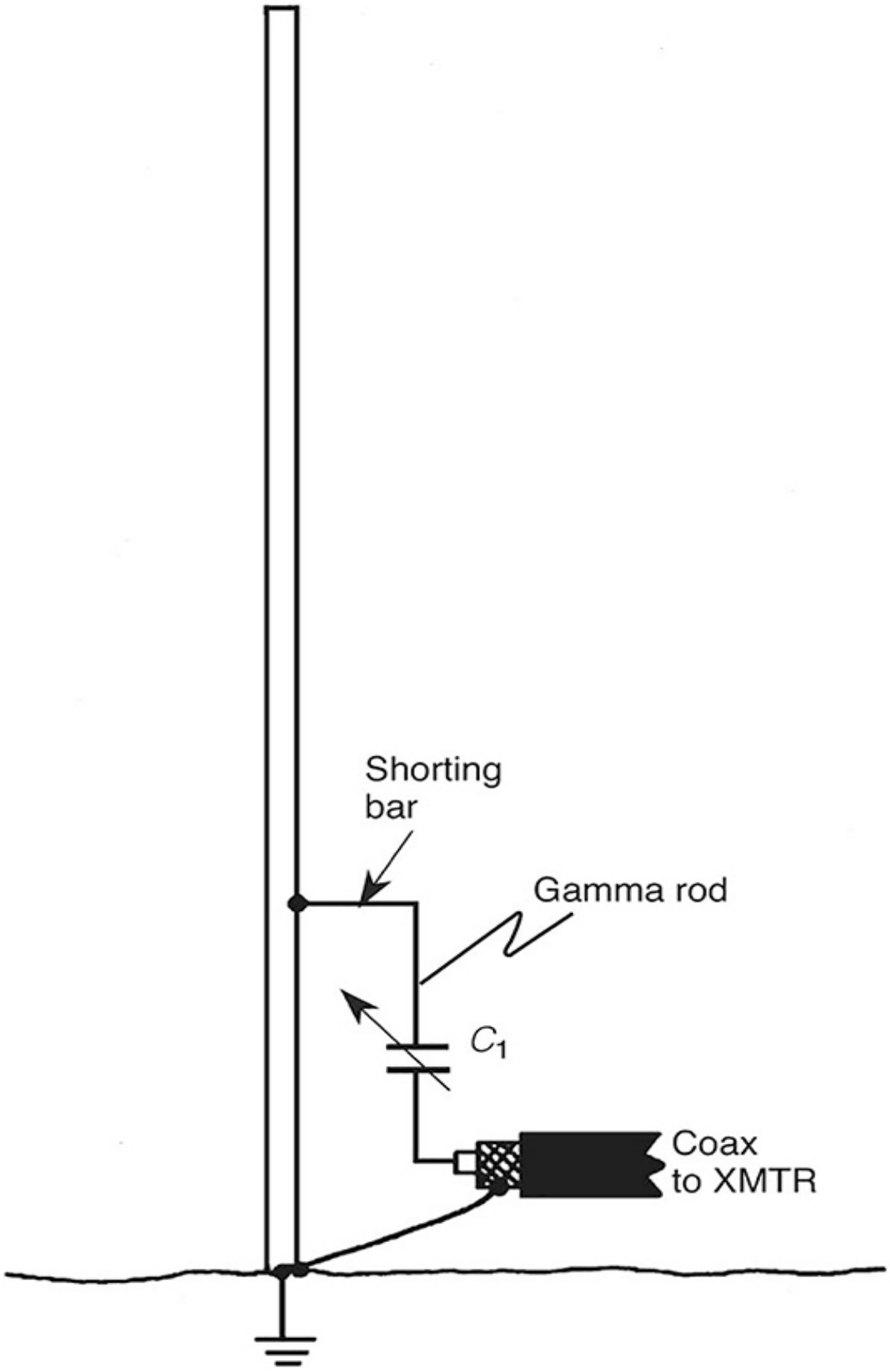

Shorting
bar
Gamma rod
C_1
Coax
to XMTR

FIGURE 11.18 Gamma-fed grounded vertical.

An advantage of the gamma match and other shunt-feed techniques is that the vertical antenna element retains a direct connection to ground, providing superior static discharge during thunderstorms. This is especially useful on the lower amateur frequencies, where verticals are often built using triangular guyed tower sections and it is less expensive to construct the tower with a direct electrical connection (through the tower legs) to ground.

A disadvantage of the gamma matched vertical is that the gamma rod tap point may be well beyond the user's reach on lower bands, making it difficult to adjust. Also, some multielement phased arrays use current-forcing techniques that are easier to control with the series feed.

Unless the user has the ability to adjust the gamma rod spacing and length and the overall length of the vertical element, the gamma match may require an additional matching network between the base of the antenna and the transmission line. In the easiest cases, a single series capacitor or inductor between the base of the gamma rod and the hot side of the feedline may provide a satisfactory match and SWR. More generally, an omega match (Fig. 11.19) can be inserted, although the user may need to experiment with which side of the series capacitor (C_S) the hot end of the gamma capacitor (C_G) should attach to. Chapter 20 includes details of a gamma match for a top-loaded tower on 160 m.

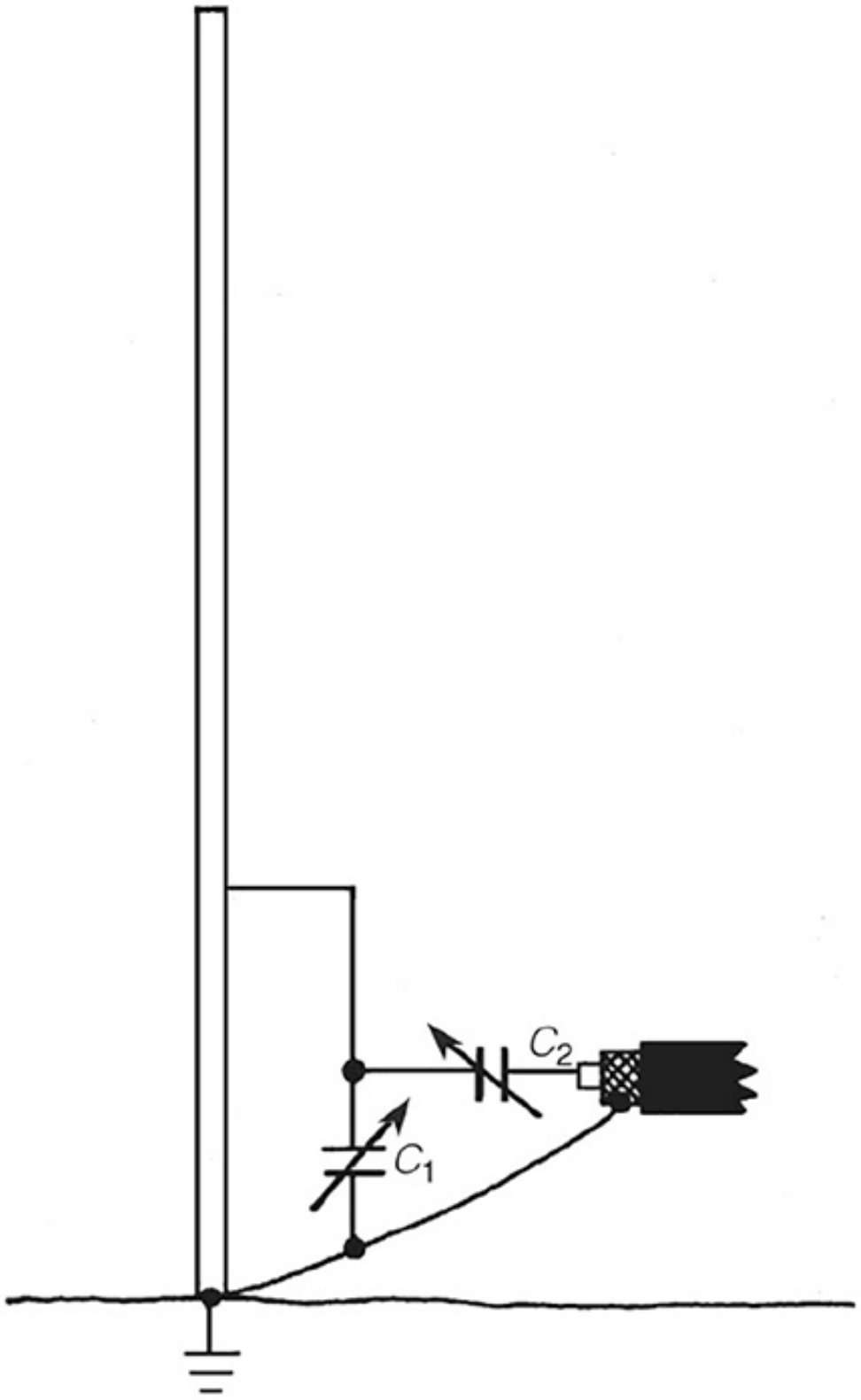


FIGURE 11.19 Omega-fed grounded vertical.

As the radiating element of a vertical monopole increases beyond $\lambda/4$, both the resistance and the reactance seen at the series feedpoint increase. For any significant increase in the length of the vertical radiator, direct connection to commonly available transmission lines will result in increased SWR on the transmission line, so some form of matching network or ATU should be employed. An alternative is to directly ground the base of the radiator and use some form of shunt-feed network to accomplish the proper match.

For a 5⁄8λ antenna, the input resistance is about 1600 Ω—certainly not an acceptable match for anything but a custom-designed transmission line. So some form of impedance matching is needed. For a single-band 5⁄8λ antenna, one option is to use a form of stub matching employing unbalanced coaxial cable, such as shown in Fig. 11.20. The lengths are given by

$$L_1(\text{ft}) = \frac{122}{f(\text{MHz})} \text{ or } \frac{38}{f(\text{MHz})} \quad (11.2)$$

and

$$L_2(\text{ft}) = \frac{30}{f(\text{MHz})} \text{ or } \frac{9}{f(\text{MHz})} \quad (11.3)$$

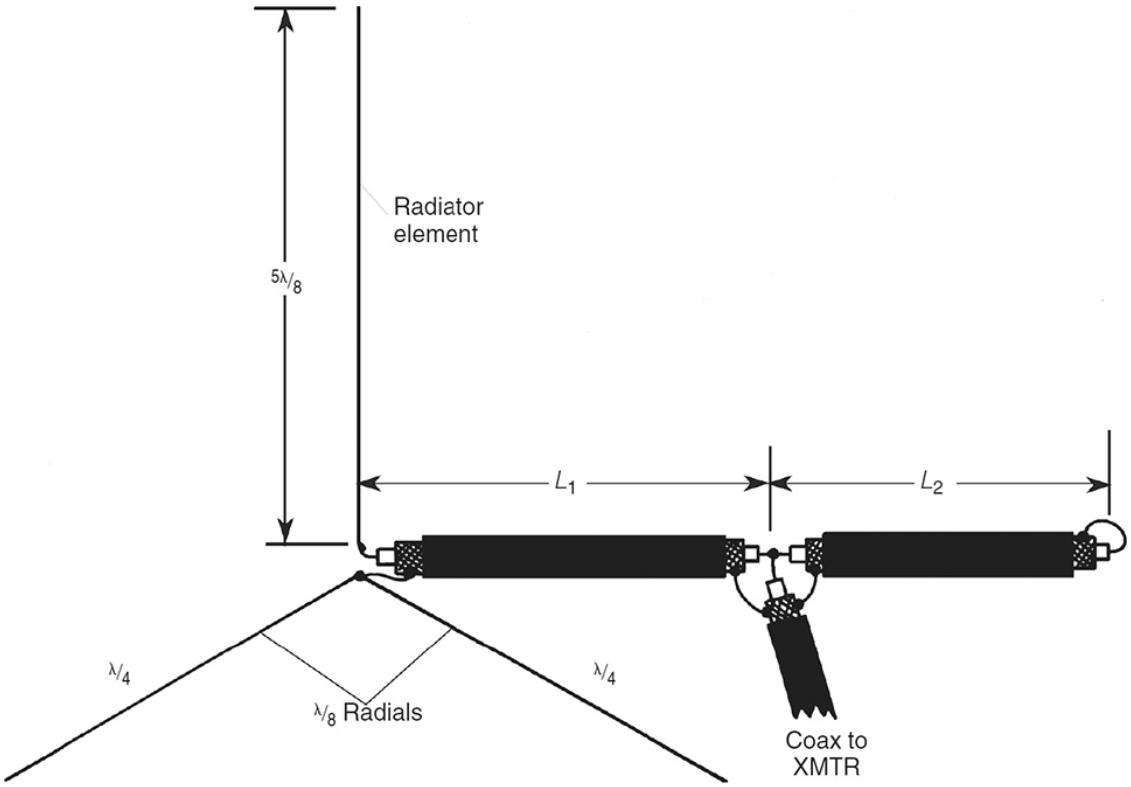


FIGURE 11.20 *Q*-section for match feedpoint impedance to line.

11.6 Elevated Ground-Plane Antenna

There are three main differences between a ground-mounted vertical and an elevated ground-plane vertical:

- Radials located some distance above ground must be treated as resonant elements.
- Fewer radials are needed to ensure that most of the transmitter power is delivered to the radiating element instead of being dissipated in lossy earth.
- An elevated GP vertical may have fewer nearby obstacles (e.g., trees and buildings) and thus enjoy a cleaner shot at low elevation angles.

Once the base of an elevated ground-plane vertical is raised above a certain height (many consider $\lambda/8$ a reasonable minimum),

radials act less like long capacitors and more like the other half of a vertical dipole. As such, it is important for maximum radiation and ease of matching that the lengths are approximately

$\lambda/4$

at the operating frequency. The exact length will vary somewhat with radial droop—the vertical angle at which the radials come away from the base of the monopole.

The farther the lossy earth from these elevated radials, the less effect it has and the easier it is to capture the bulk of the return currents of the vertical radiating element in a small number of radials. Between two and eight radials per band is adequate for elevated antennas. Unless at least three equispaced radials are used, there will be some variation in field strength with azimuth. The improvement in radiated field characteristics for more than three or four radials is small if the GP antenna is at least $\lambda/8$ off the ground. For a

$\lambda/4$

vertical, there may be little difference between horizontal and drooping radials other than to raise the feedpoint impedance from around 25 to 40 Ω.

11.7 Vertical versus Horizontal Polarization

A question frequently asked is: "Would I be better off putting up a horizontal dipole or a vertical?" Here are some factors often found to be helpful in arriving at an answer:

- Different polarizations are used for different applications and in different communities. The 11-m citizens band uses vertical polarization because of the high percentage of vehicular mobile users. Almost all house- or tower-mounted antennas—even multielement Yagis and quads—for 27 MHz are vertically polarized. The same is true for most FM mobile and repeater operation on 2 m and above. On the other hand, weak signal SSB or CW DXing on those bands is accomplished with long-boom, high-

gain Yagis employing horizontal polarization. FM broadcasting is done with the transmitted signal split between horizontal and vertical polarization, since antennas of both polarizations are in use.

- Propagation is polarization-dependent. On the HF bands there is generally a 6-dB ground reflection advantage at each end of the path at higher elevation angles for horizontal antennas. Below about 10 MHz, however, verticals often outperform all but the highest dipoles and beams at low radiation angles. Otherwise, the source polarization is immaterial for skip propagation because the ionosphere tends to randomly rotate the polarization of waves passing through it. If you are an inveterate DXer, you will most likely enjoy more success on the MF and lower HF bands with a vertical than with a dipole, but if your principal operating activity is a regional or local network of friends, a low horizontal dipole or loop may be your best choice.
- Local noise sources may favor one polarization over the other. An improvement in received signal-to-noise ratio (SNR) through the proper choice of polarization may be possible.
- Space or support limitations may dictate one antenna over another. A residential backyard may be better suited for an 80-m ground-mounted $\lambda/4$ vertical than for a $\lambda/2$ dipole, while apartment residents may choose a simple dipole or endfed longwire stretched from rooftop to rooftop across the courtyard below.

Another solution is to use both types of antennas. Switching from one to another as conditions change may lead to SNRs that change by as much as 20 dB.

11.8 Supporting the Vertical Antenna

Verticals for frequencies below about 5 MHz are substantial structures. The support requirements are akin to those of 60-ft or greater towers. Information on supporting verticals for the MF and lower HF bands is found in Chap. 32. Above 5 MHz, supporting a $\lambda/4$ vertical is a simpler task. For these antennas, the techniques of Chap. 31 are typical.

PART III

Directional Antenna Arrays

CHAPTER 12
Wire Arrays
CHAPTER 13
Vertical Arrays
CHAPTER 14
The Yagi-Uda Beam Antenna
CHAPTER 15
Cubical Quads and Delta Loops

CHAPTER 12
Wire Arrays

Instead of a simpler antenna, to get higher forward gain, better rejection off the back or sides of the antenna, or a combination of both, wire arrays can be used. Few antenna projects bring as much reward at such a low cost as wires, and that's especially true of wire arrays.

12.1 Longwire Antennas

The end-fed longwire (or, more simply, longwire) is a name given to any of several types of resonant and nonresonant antennas. The longwire is a single antenna, not an array, but we can use the concepts of array antennas to understand its operation, so we consider it here in this chapter on wire arrays. It is capable of providing gain over a dipole and (if it is long enough) a low angle of radiation. These advantages become substantial when the antenna is several wavelengths long. Longwire antennas have been popular for decades—often because the layout of a particular property lends itself to locating the end of an HF wire close to the radio room.

Depending upon the operating frequencies used, a longwire antenna may be either resonant or nonresonant. Figure 12.1 shows the classic random length, nonresonant long-wire antenna. It consists of a wire radiator that is at least a quarter-wavelength or more commonly much longer. The specific length is not critical, but by convention it is a wavelength or more at the lowest frequency of operation. A 70-ft wire fed at one end against a good RF ground will acquit itself well on all HF bands above 3.5 MHz.

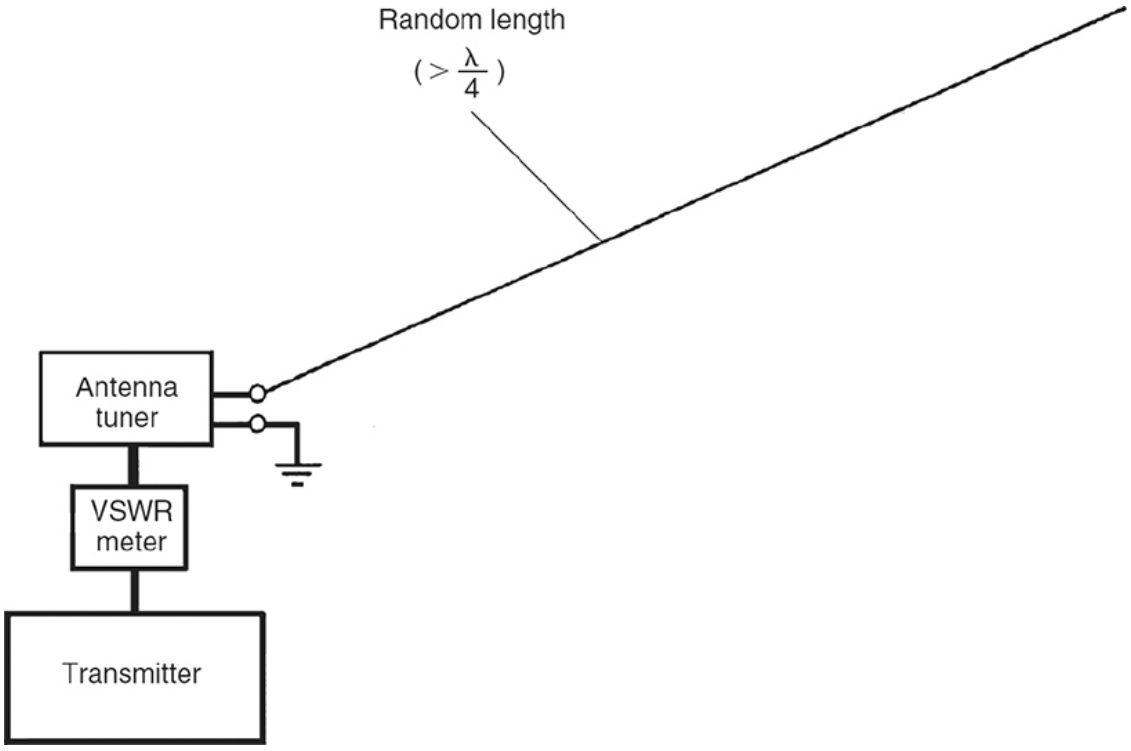


FIGURE 12.1 Random-length antenna.

The longwire is end-fed. When it is an even number of quarter-wavelengths it will exhibit a high feedpoint impedance of a few thousand ohms. For lengths that are odd multiples of $\lambda/4$ it will present an impedance of perhaps 25 to 100 Ω. An antenna tuner (ATU) should be used. Even transceivers with built-in antenna tuners typically cannot match the very high feedpoint impedances of end-fed wires that are approximately a multiple of $\lambda/2$ in length.

The most important factor affecting the performance of an end-fed longwire can be the ground system. While we know that a $\lambda/4$ monopole radiates well only with the help of the ground system beneath it, any antenna with a single-wire feedline depends on the nearby ground to pass return currents back to the transmitter circuits. If the transmitter is trying to push electrons into a longwire

on each half-cycle of RF, it must necessarily be receiving or pulling an equal number of electrons from a return circuit of some kind. Where no attention is paid to the ground return paths, the return path is mix of anything and everything connected to the chassis and the ground of the power amplifier: AC power line, mic cable shield, audio cable shields, etc. Lack of an appropriate ground return path is a major reason why equipment in a radio transmitting environment becomes "hot" whenever the transmitter is operated. Since the fed end of the antenna is often in the radio room, the issue is especially relevant to the installation of a longwire antenna.

Failure to include an adequate ground return for a longwire raises the apparent feed-point impedance and causes a large portion of the transmitter output power to be dissipated in the resistive losses of other ground return paths. Ideally, many radials of lengths greater than or equal to

$\lambda/4$

at the lowest frequency of operation, as shown in Fig. 12.2, should be included as part of any longwire antenna project.

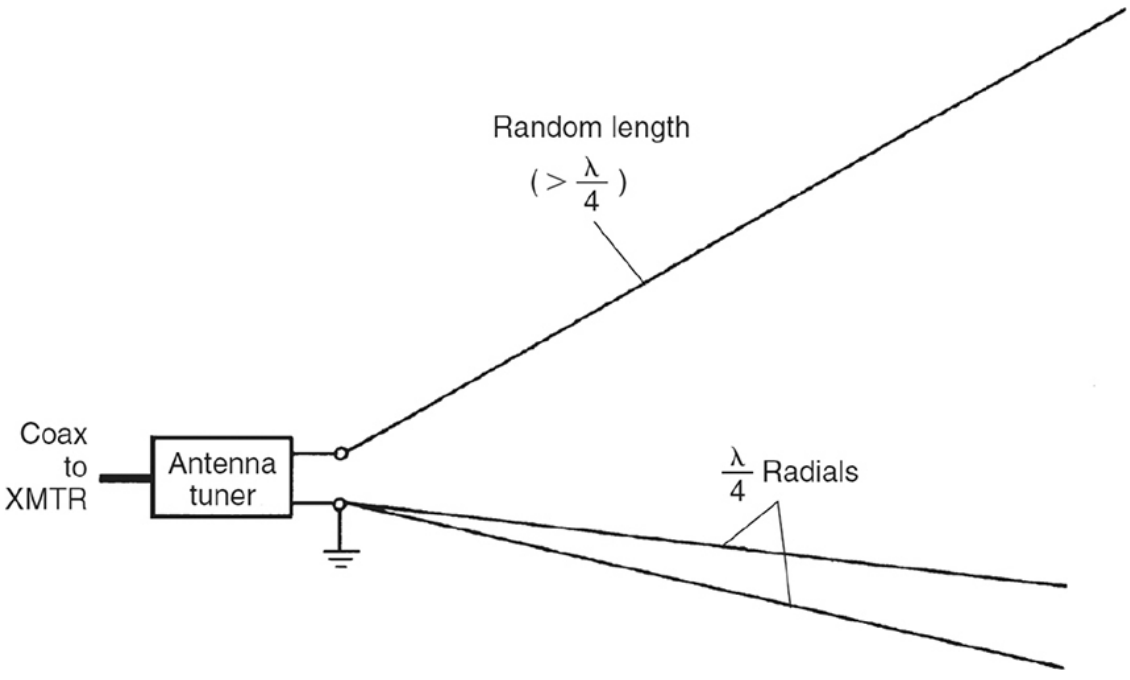


FIGURE 12.2 Radials improve the ground for a random-length antenna.

An RF ground is different from the ground in an electrical supply system. RF grounds for antennas are discussed in detail in Chap. 33. At the very least, a $\lambda/4$ radial can be connected to the chassis of the ATU (or the transmitter if no ATU is used) and run outside. This effectively converts the longwire into an off-center-fed wire of unpredictable performance, but it's better than having no defined ground at all.

For multiple bands, a single radial may be inadequate. Just as antenna tuners can be used to match to an antenna with a challenging impedance, tuners are available for the ground system. The MFJ model MFJ-931 artificial RF ground is installed between the radial wire and the ATU or transmitter chassis ground connection, as shown in Fig. 12.3. With band changes, the tuning controls can be adjusted for maximum ground current. An artificial ground is often inferior to a well-designed ground system and is normally used only when a better option is unavailable.

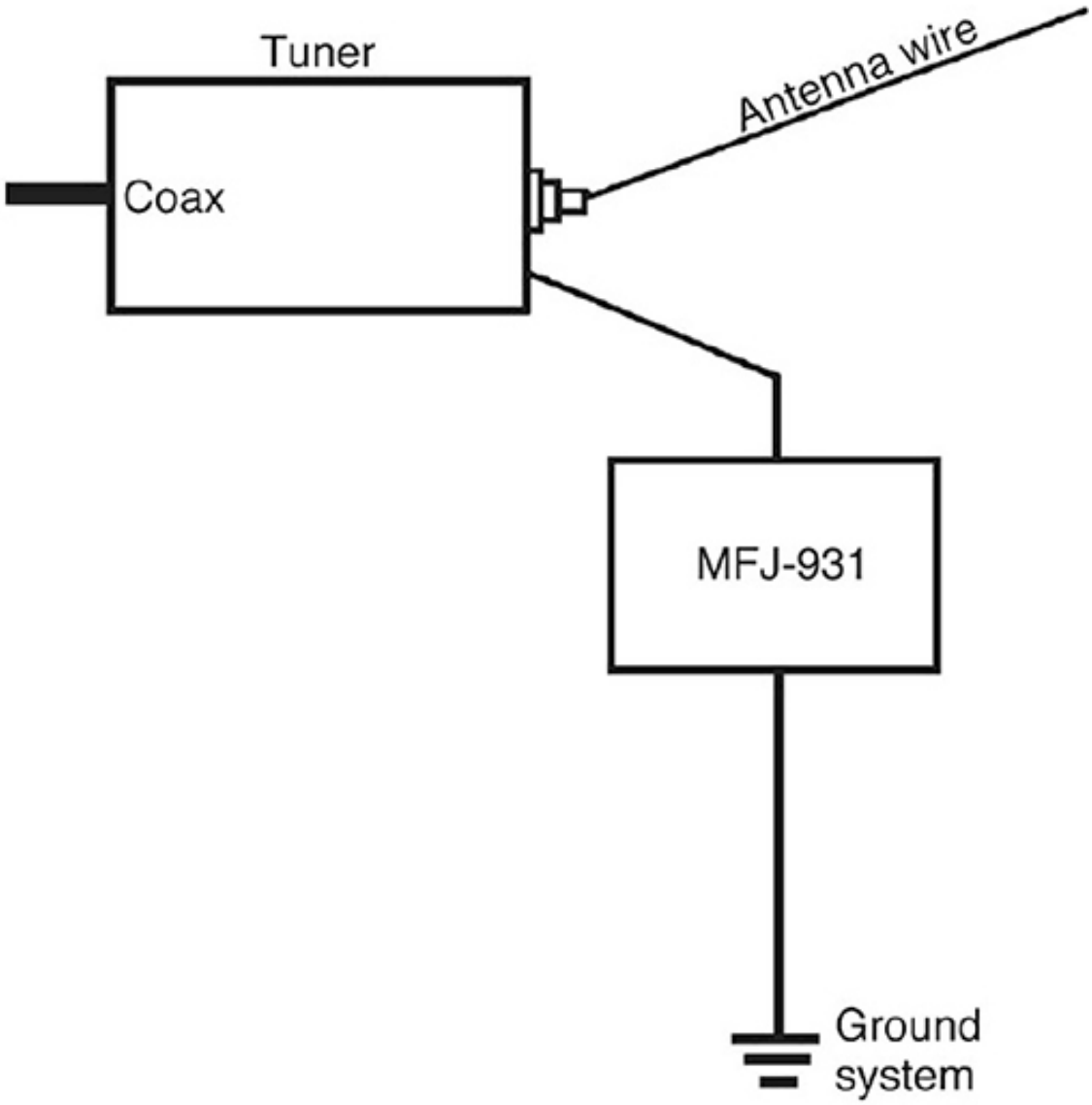


FIGURE 12.3 MFJ ground line tuner installation.

12.1.1 Longwire Pattern and Gain

As the electrical length of an end-fed wire antenna increases—either by using a fixed-length wire on a higher frequency or by lengthening the wire while keeping the operating frequency fixed—the pattern and the direction of maximum gain change. At $\lambda/2$, the pattern of an end-fed wire is the broadside "donut" of a $\lambda/2$

dipole, perhaps very slightly skewed as a result of being fed at one end instead of in the center. At a length of λ, the figure eight pattern has broken into a four-lobed cloverleaf with peaks at 45 degrees from the antenna axis and the maximum gain of the antenna (in the center of each lobe) has increased by almost 1 dB.

As the frequency is further increased, the pattern breaks into more lobes, such as that depicted in Fig. 12.4 for a 2λ longwire. The lobes shift toward the antenna axis and the amplitude of these main lobes grows, reaching an additional 5 dB compared to the $\lambda/2$ wire at a length of 4λ, which corresponds to using an 80-m 135-ft $\lambda/2$ dipole on 10 m. Figure 12.5 plots the increase in gain of the main lobe relative to a $\lambda/2$ dipole as a function of longwire length.

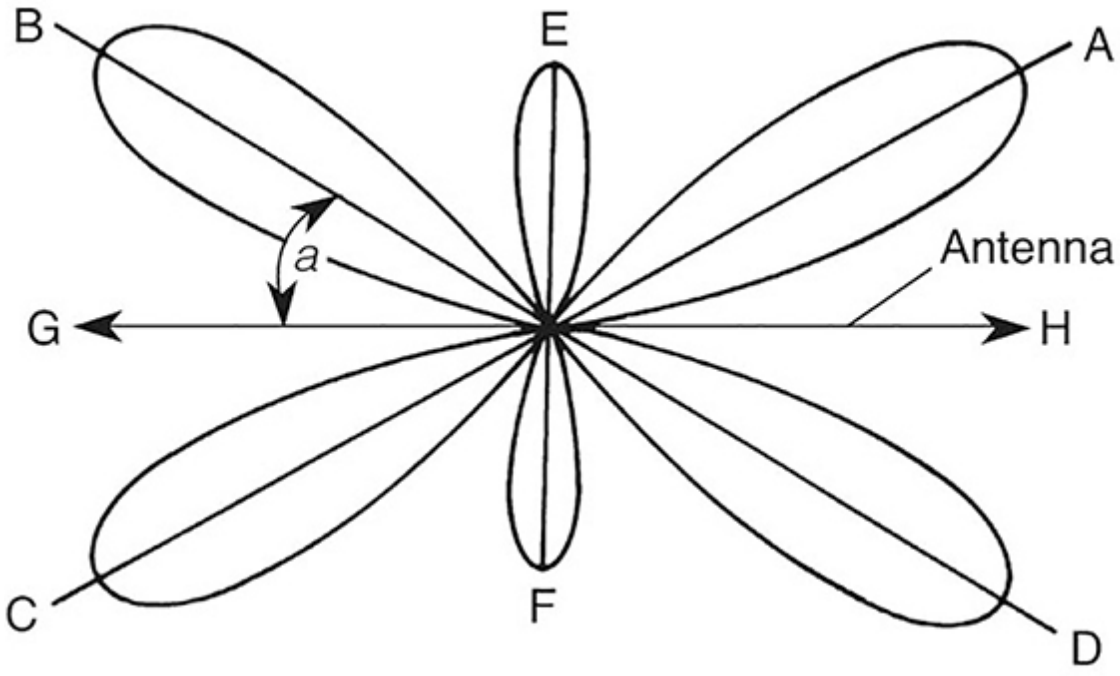


FIGURE 12.4 Radiation pattern of longwire antenna.

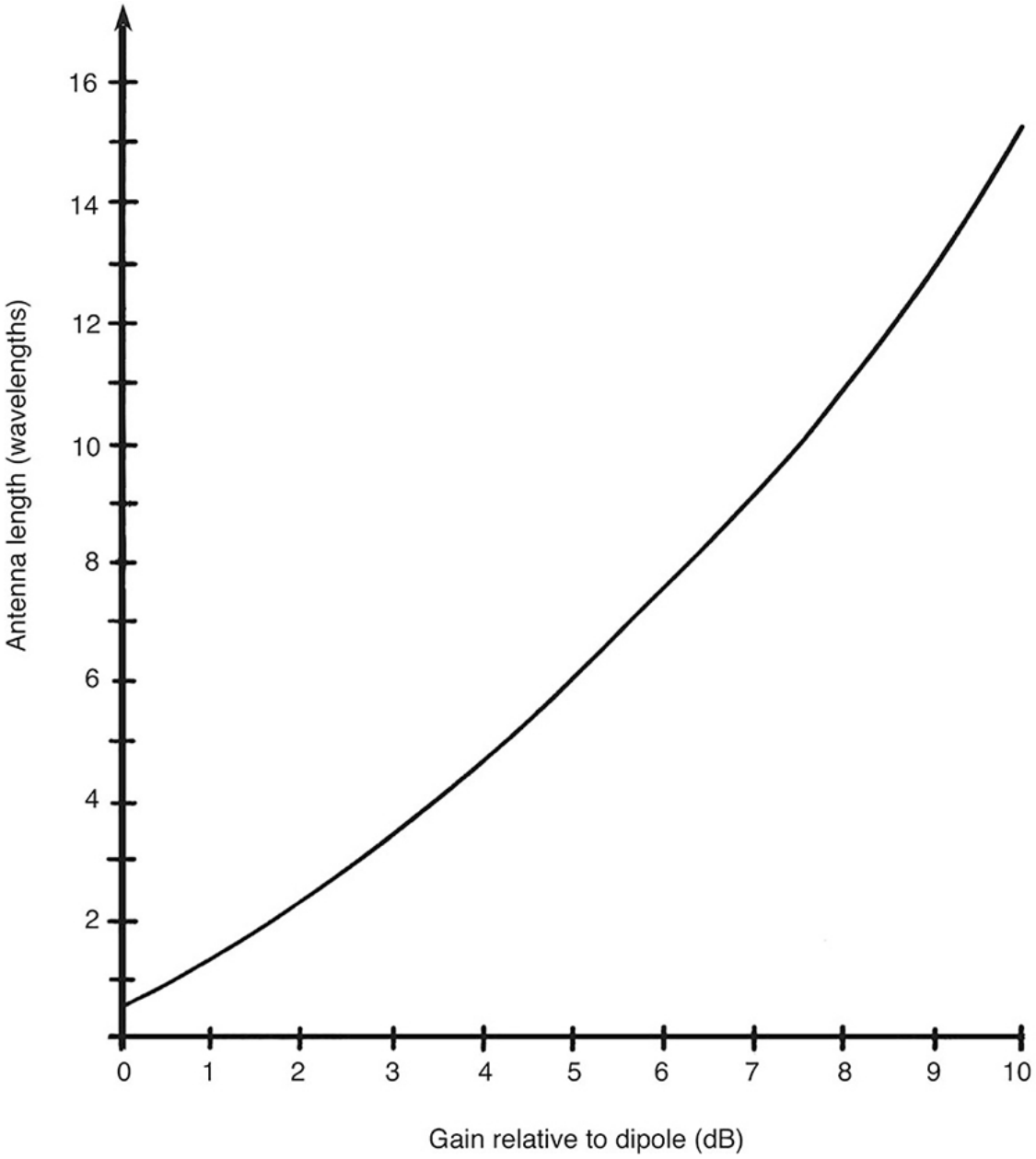


FIGURE 12.5 Antenna length versus gain over dipole.

Most longwires are fixed in one location. Having the direction of peak signal gradually precess from broadside to nearly 90 compass degrees away while changing frequency may be less than ideal. Often an amateur or short-wave listener (SWL) with acreage available installs multiple longwires for a preferred band. An ideal configuration might have the fed ends at a common point where an ATU is located and can be switched from one wire to another.

For long wires, electrostatic fields can be a safety hazard. Thunderstorms many miles away can produce enough charge build-up on long antennas to cause serious damage to a receiver or injure the operator. A simple solution is to connect a transmitter-rated RF choke between the antenna terminal and the ground connection of the ATU. In a pinch, a dozen 10-MΩ, 2W resistors in parallel can be substituted.

12.1.2 Nonresonant Longwire Antennas

The resonant longwire antenna is a standing wave antenna because it is unterminated at the far end. A signal propagating from the feedpoint toward the open end will be reflected back toward the transmitter from the open end; the resulting superposition of the forward and reflected waves sets up stationary (i.e., "standing") currents and voltages along the wire, similar to the process described in Chap. 5.

In contrast, a nonresonant longwire is terminated at the far end in a resistance equal to its characteristic impedance (see Fig. 12.6). Thus, the outward-bound wave is absorbed by the resistor, rather than being reflected. Such an antenna is called a traveling wave antenna. Terminating resistor

R_1

is selected to be equal to the characteristic impedance

Z_0

of the antenna (i.e.,

$R = Z_0$

). When the wire is 20 to 30 ft above average ground,

Z_0

is about 500 to 600 Ω.

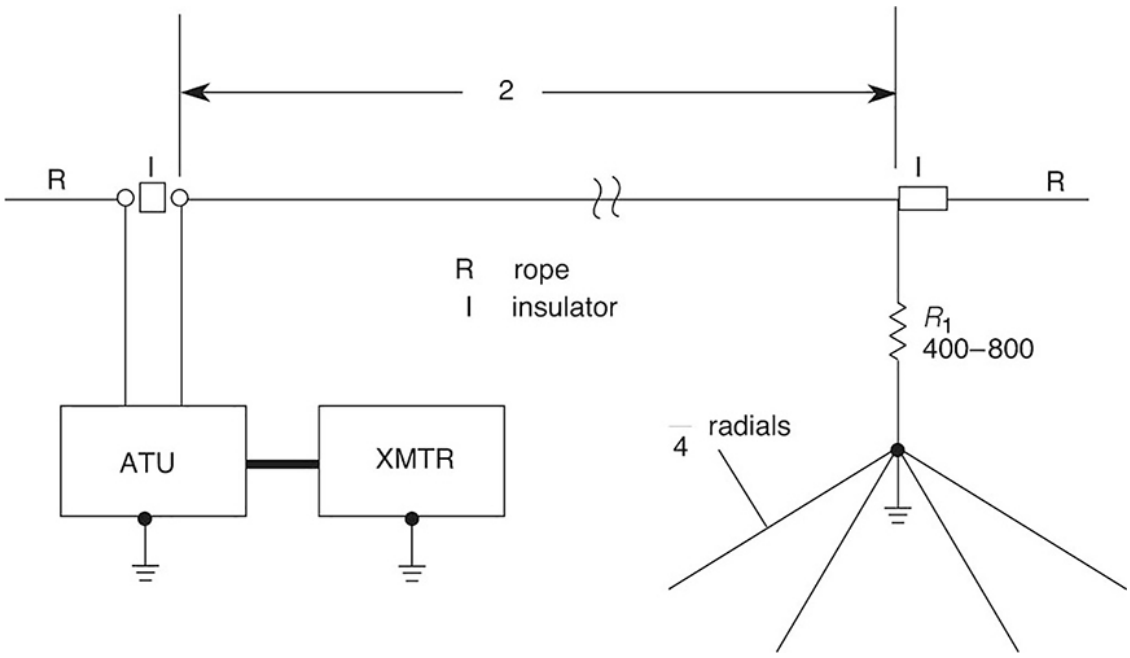


FIGURE 12.6 Longwire (greater than 2λ) antenna is terminated in a resistance (typically 400 to 800 Ω). Radials under the resistor improve the grounding of the antenna.

The radiation pattern for the terminated longwire is a unidirectional version of the multilobed pattern found on the unterminated longwires. The orientation of the lobes varies with frequency, even though the pattern remains unidirectional. The directivity of the antenna is partially specified by the angles of the main lobes. It is interesting to note that gain rises almost linearly with $n\lambda$, while the directivity function changes rapidly at shorter lengths (above three or four wavelengths the rate of change diminishes considerably). Thus, when an antenna is cut for a certain low frequency, it will work at higher frequencies, but the directivity characteristic will be different at each end of the spectrum of interest.

Of course, if a terminated longwire is used for transmitting, approximately half the transmitter power will be dissipated in the termination resistor, so one should subtract 3 dB from the overall gain relative to unterminated antennas.

12.2 Arrays of Longwires

Longwire antennas can be combined to increase gain and directivity. Two of the most popular of these are the vee beam and the rhombic. Both forms can be made in either resonant (unterminated) or nonresonant (terminated) versions.

12.2.1 Vee Beams

The vee beam (Fig. 12.7) consists of two equal-length longwire elements (wire 1 and wire 2), fed 180 degrees out of phase and oriented to produce an acute angle between them. The 180-degree phase difference is provided by connecting the two wires of the vee to opposing sides of a single parallel conductor feedline.

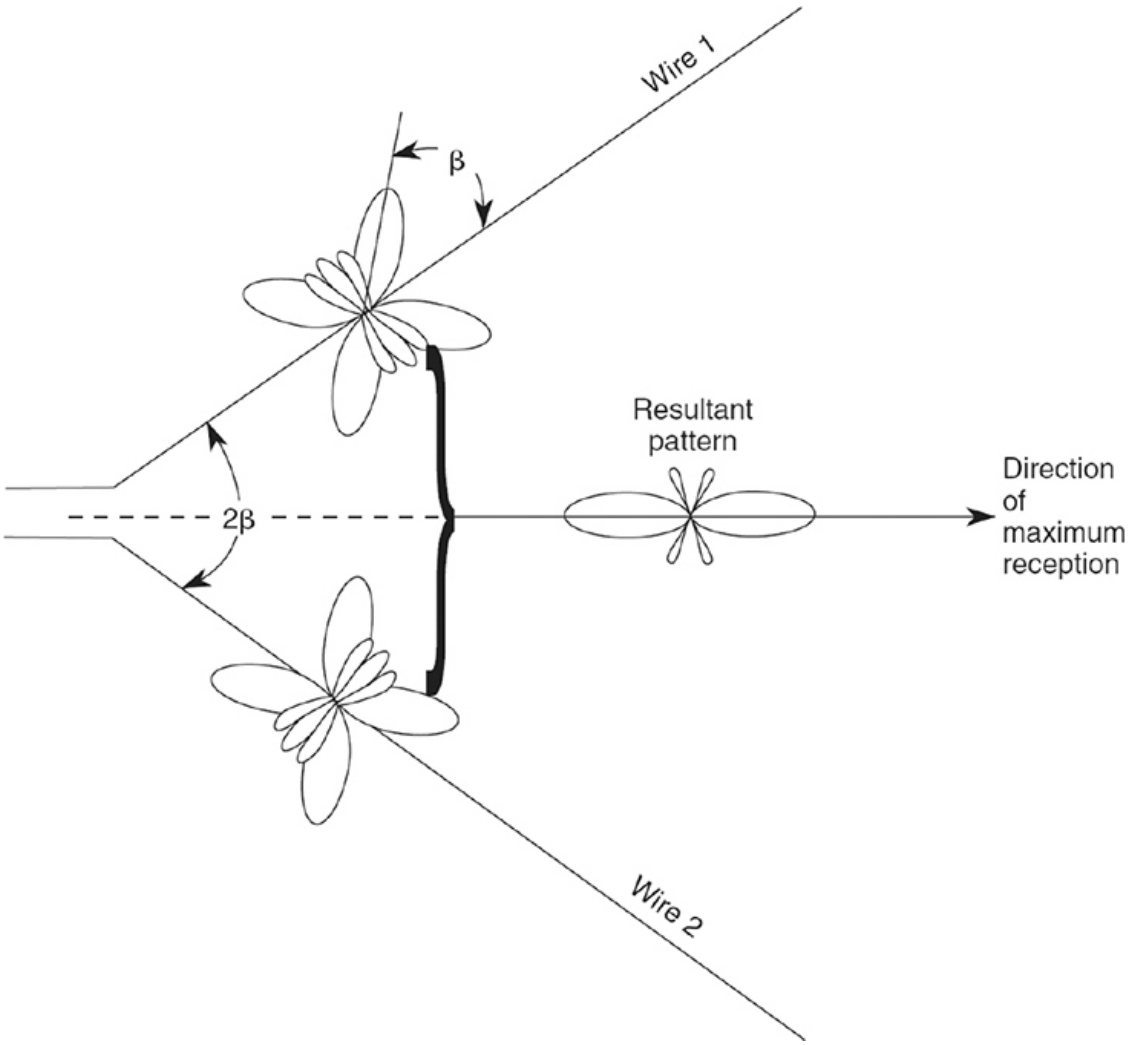


FIGURE 12.7 The radiation pattern of vee beam antenna consists of the sum of the two longwire patterns that make up the antenna.

The unterminated vee beam of Fig. 12.7 has a bidirectional pattern that is created by summing together the patterns of the two individual wires. Proper alignment of the main lobes of the two wires requires an included angle between the wires of twice the radiation angle of each wire. If the radiation angle of the wire is β, then the appropriate included angle is 2β. To raise the pattern a few degrees, the 2β angle should be slightly less than these values. It is common practice to design a vee beam for a low frequency (e.g., 75-/80- or 40-m bands), and then to also use it on higher

frequencies that are harmonics of the minimum design frequency. A typical vee beam works well over a very wide frequency range if the included angle is adjusted to a reasonable compromise. A common included angle is between 35 and 90 degrees, depending on how many harmonic bands are required.

Vee beam patterns are usually based on an antenna height that is greater than a half-wavelength from the ground. At low frequencies, such heights may not be possible, and ground reflection causes pattern distortion.

The gain of a vee beam is about 3 dB higher than the gain of the single-wire longwire antenna of the same wire length and is considerably higher than the gain of a dipole (see Fig. 12.8). At three wavelengths, the gain is 7 dB over a dipole. The gain may be still higher mutual impedance effects (around 1 dB at 5λ and 2 dB at 8λ).

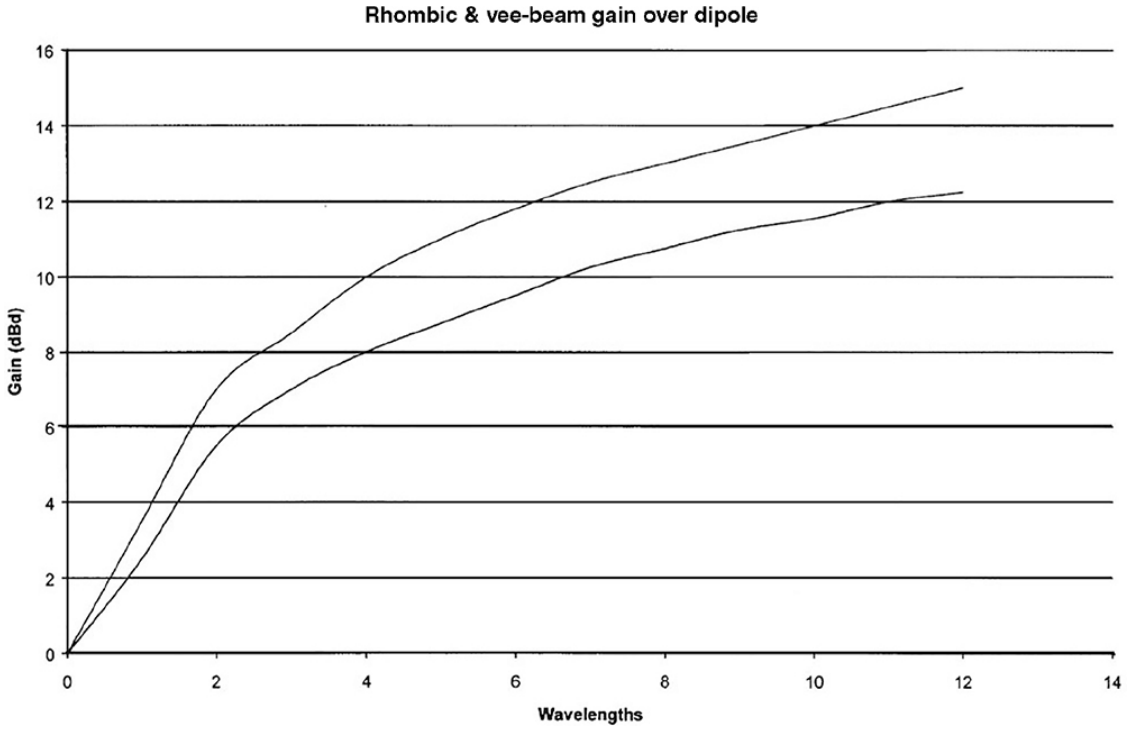


FIGURE 12.8 Gain versus length of vee beam and rhombic beam antennas.

12.2.2 Nonresonant Vee Beams

Like single-wire longwire antennas, the vee beam can be made nonresonant by terminating each wire in a resistance that is equal to the antenna's characteristic impedance (Fig. 12.9). The regular vee is a standing wave antenna, whereas the terminated version is a traveling wave antenna. Because the terminating resistor absorbs the incident wave after it has propagated to the end of the wire, standing wave antennas are unidirectional. In a standing wave antenna, any energy reaching the far end is reflected back toward the source, so it can radiate in the opposite direction relative to the incident wave.

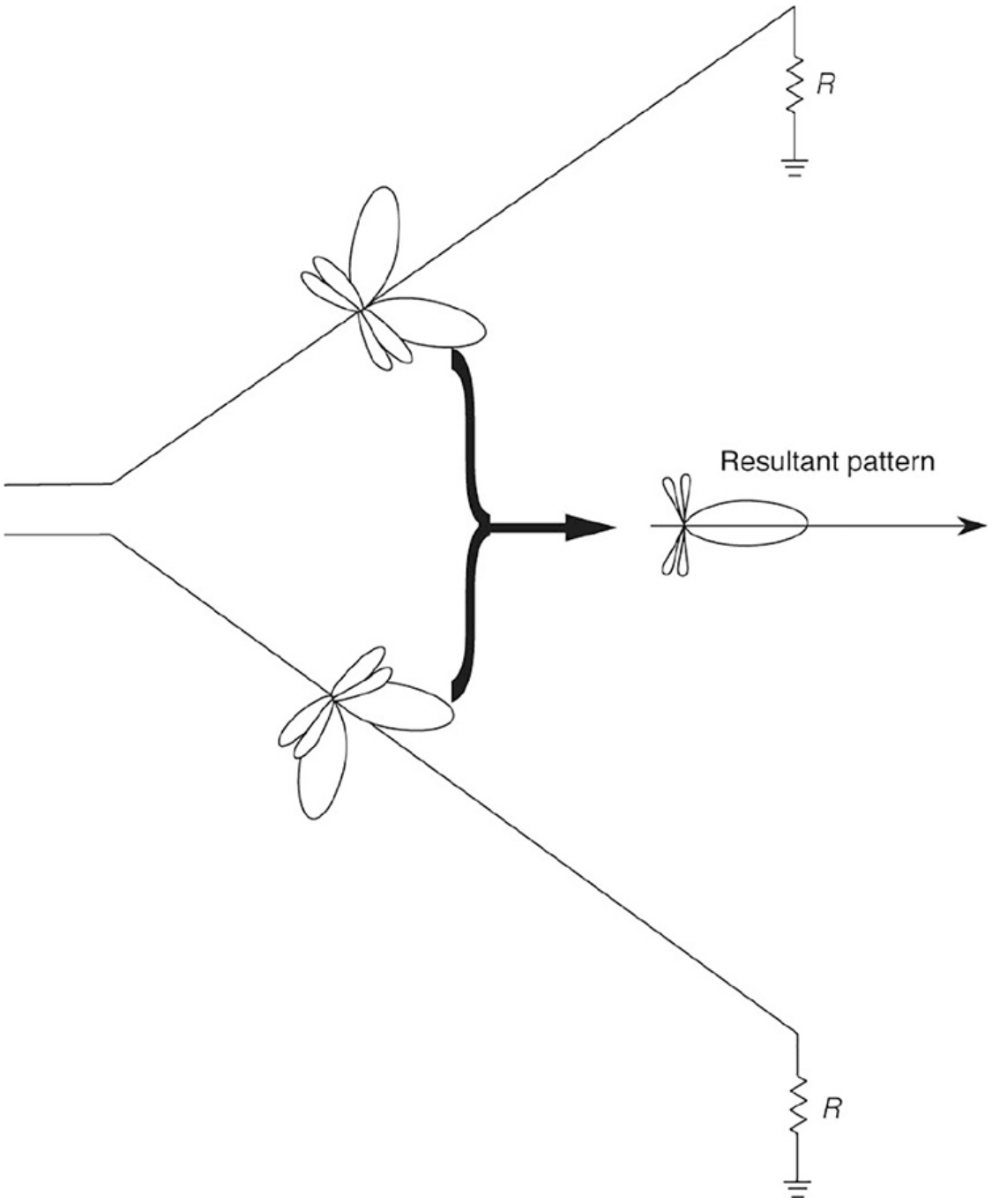


FIGURE 12.9 Nonresonant vee beam antenna.

12.2.3 Rhombics

The rhombic antenna, also called the double vee, consists of two vee beams positioned to face each other with the tips connected. The unidirectional, nonresonant (terminated) rhombic shown in Fig.

12.10 develops approximately the same gain and directivity as a vee beam of the same size. The nonresonant rhombic has a gain of about 3 dB over a vee beam of the same size (see Fig. 12.8 again).

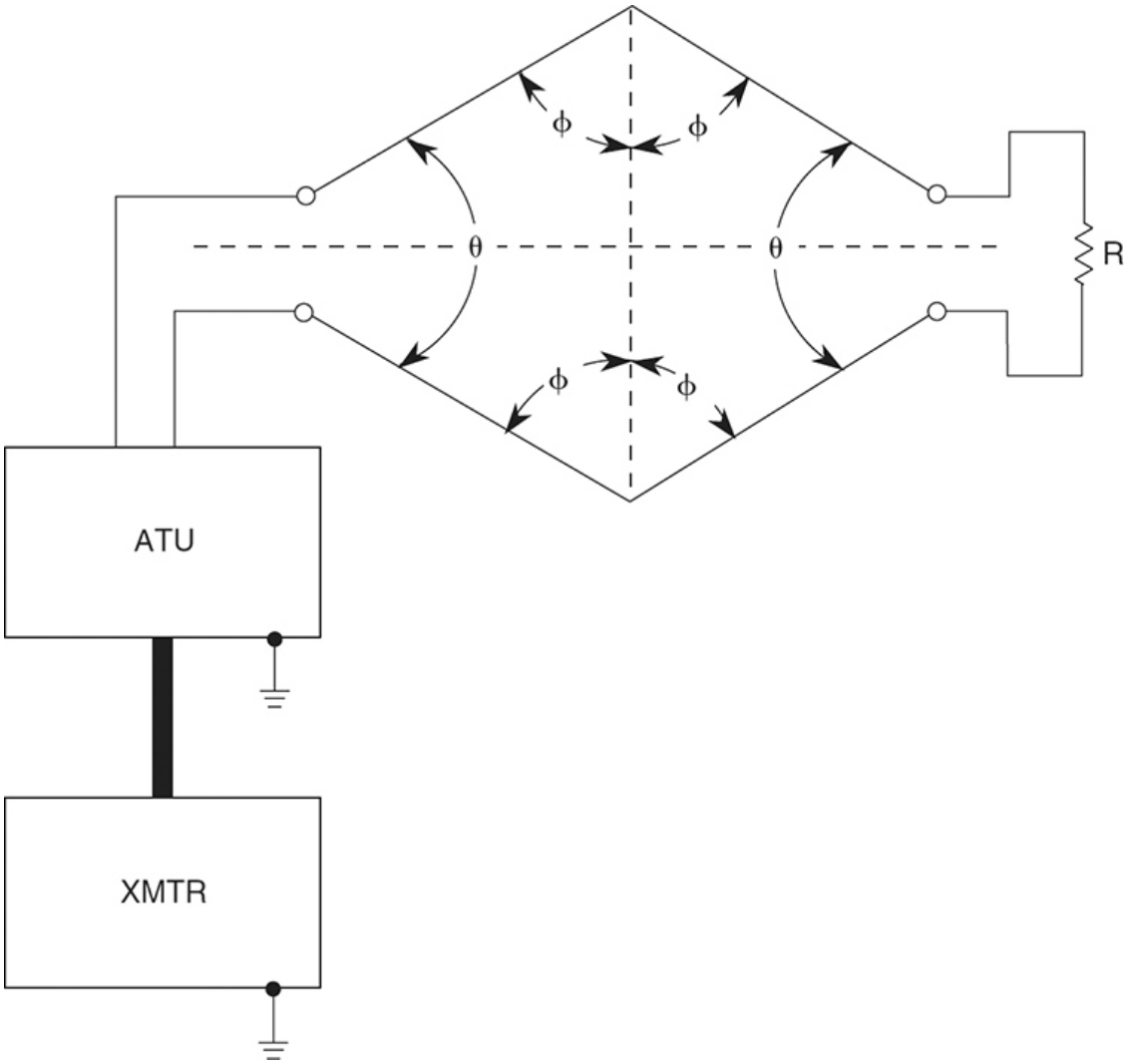


FIGURE 12.10 Terminated rhombic antenna.

The layout of the rhombic antenna is characterized by two angles. One-half of the included angle of the two legs of one wire is the tilt angle ϕ, while the angle between the two wires is the apex angle θ. The angles are related; from Fig. 12.10 we can see that

$$\theta/2 = 90 - \phi$$

. A common rhombic design uses a tilt angle of 70 degrees, a length of 6λ for each leg (two legs per side), and a height above

the ground of 1.1λ. For this antenna, θ is 40 degrees.

The termination resistance for the nonresonant rhombic is 600 to 800 Ω, and it must be noninductive across the entire range of operating frequencies. For transmitting rhombics, the terminating resistor should be capable of dissipating at least one-third of the average power of the transmitter. For receive-only rhombics, the termination resistor can be a 2W carbon composition or metal-film type. Such an antenna works nicely over an octave (2:1) frequency range.

12.3 Bobtail Curtain

The bobtail curtain of Fig. 12.11 is an array of three quarter-wavelength vertical elements in a line. Adjacent elements are spaced a half-wavelength apart and attached electrically to a horizontal wire across their tops that serves both as a phasing wire and as the other half of an L-shaped dipole. The array is fed at the bottom end of the center element. This is a high impedance point, so a parallel-tuned tank circuit or an ATU capable of matching high impedances is required.

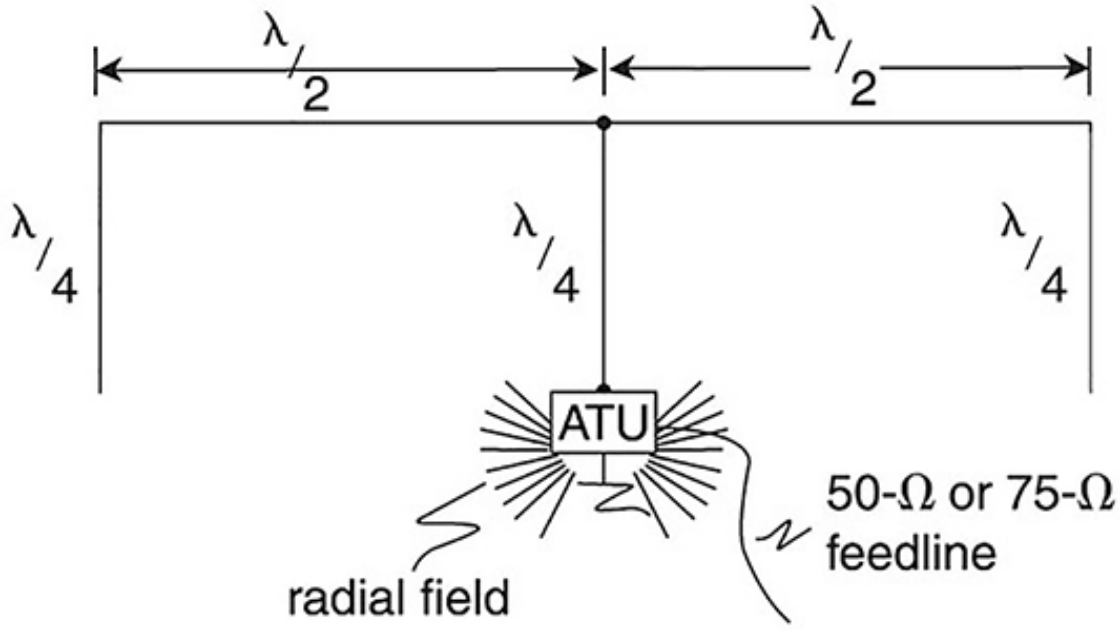


FIGURE 12.11 Bobtail curtain.

The array must be kept high enough in the air that the bottoms of the $\lambda/4$ vertical elements are some distance above ground, so that humans and animals cannot touch the high-voltage tips. The feed wire to the center element may require special protection. A bobtail curtain for 80 m should be suspended about 75 or 80 ft up. With catenary effects causing the phasing line to sag, end supports of at least 100 ft are needed.

RF energy introduced at the bottom of the center element reaches the phasing line at the top and splits equally to the two end elements. Since each end element is $\lambda/2$ from the center element, the currents of all three elements are in phase. The current magnitude in the center element is twice that of either of the others. This is an example of the binomial array discussed in Chap. 7. Since the currents feeding the end elements are traveling in opposite directions along the top the horizontal radiation is partially canceled out. Because all three elements are in phase, the principal lobe is broadside to the plane of the three radiating elements. For maximum signal north and south, for instance, the array should be suspended between supports on an east-west line.

Unless there is an RF ground at the base of the middle element for the matching network to connect to, there is no low-loss return path for RF currents flowing to and from the ground end of the matching network and transmitter power will be wasted in unintended lossy ground return paths. The matching network at the base of the middle element may require a good system of radials. Since radials lying on the surface of the earth are not resonant, their exact length is not particularly critical. In some cases, a radial fields may be needed under the end elements as well. The bobtail curtain has a low angle of radiation and may work well for long haul contacts while being useless for anything nearby.

12.4 Half-Square

If one end element of the bobtail curtain of Fig. 12.11 and its associated phasing line are removed, the resulting antenna is known as a half-square. This antenna has two vertical radiators joined by a
$\lambda/2$
horizontal phasing line. It is fed at the bottom of one element—a high impedance point.

The half-square is a two-element broadside array. The fed element and half the horizontal wire constitute an L-shaped
$\lambda/2$
dipole feeding a second
$\lambda/2$
dipole. Since the current in consecutive
$\lambda/2$
sections reverses, the currents in the vertical elements are in phase. If the antenna is hung from east and west supports, maximum radiation is north and south. Off either end, contributions by the two vertical radiators are 180 degrees out of phase, so the array has a null to the east and west. Because the currents in the two halves of the horizontal wire are out of phase, they mostly cancel and the array exhibits little high-angle radiation.

Over average ground, the half-square has about 4 dB additional gain broadside to the axis of the antenna compared to a single
$\lambda/4$
vertical and the peak of the main lobe occurs at a slightly lower elevation angle (20 degrees versus 23 degrees). Overall system efficiency and the specifics of the matching network may dictate the use of radials directly connected to the ground and chassis of the ATU that feeds the vertical wire. Adding radials under both vertical elements adds about 0.4 dB to the low-angle broadside gain over average earth.

12.5 20-m ZL-Special Beam

The antenna shown in Fig. 12.12 is related to the Yagi beam. It consists of a pair of folded dipoles mounted approximately 0.12 wavelengths apart. Each element is 30.5 ft in length, and the spacing between the two is 7.1 ft. The elements can be made of two parallel wires separated with spacers or from 300-Ω television-type twin lead. Alternatively, if the antenna is to be rotatable, spaced aluminum tubing or wires inside small PVC tubes supported at their ends by a truss assembly can be used.

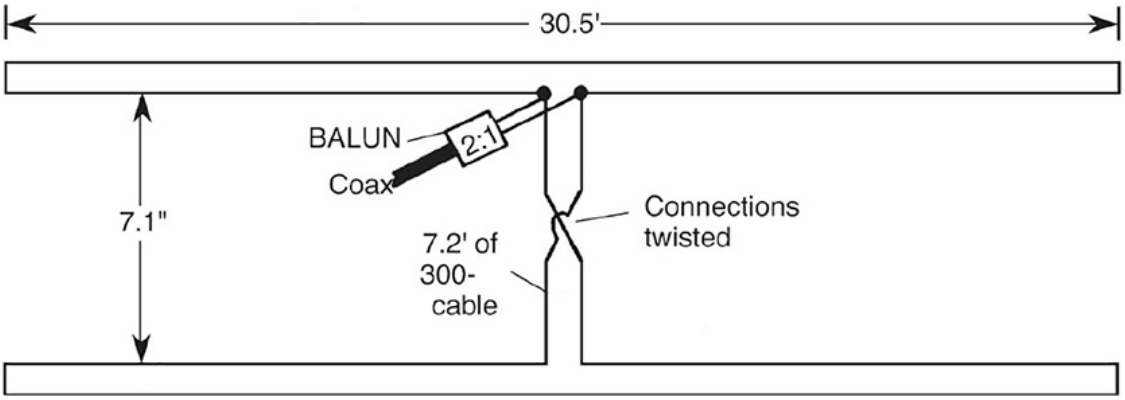


FIGURE 12.12 ZL-special beam antenna.

The two $\lambda/2$ folded dipole elements of the ZL-special are fed 135 degrees out of phase with respect to each other. The feedline is connected to one of the dipoles directly. The other is fed through a length of 300-Ω twin lead that has an electrical length of about 45 degrees ($\lambda/8$) and has been given a half-twist to add another 180 degrees of phase shift. Its physical length will depend on the velocity factor of the particular line used. The feedpoint impedance is on the order of 100 to 150 Ω, so the antenna is a good match for either 52-Ω or 75-Ω coaxial cable if a 2:1 impedance matching transformer is used.

CHAPTER 13
Vertical Arrays

We will frame our discussion of arrays of vertical antennas by considering the merits of the single vertical antenna of Chap. 11. The vertical antenna may be praised or cursed by its users, depending upon their experiences with it. As amateurs and others have come to appreciate the importance of a good RF ground system to the vertical's performance, the vertical has come into its own on the 1.8- and 3.5-MHz bands because it provides superior low-angle radiation compared to dipoles.

A vertical antenna is omnidirectional in its coverage. It transmits and receives equally well in all horizontal directions. For some applications, this is desirable. There are only a few things that can be done to increase the signal level at a specific faraway receiving location. The transmitter power can be increased, the antenna efficiency can be improved by reducing resistive losses and ensuring that the system of radials at the base of the antenna is adequate, or the electrical length of the vertical radiator can be increased (up to a maximum of $5\lambda/8$). This causes the low-angle field strength for a given transmitter power to increase because high-angle radiation is being reduced as the vertical is lengthened.

For receiving applications, how the antenna responds to noise sources is also important. For bands where propagation is heavily influenced by the ionosphere, atmospheric band noise arrives from compass headings where the band is "open." After sunset at the site of the vertical antenna, atmospheric noise on the MF and lower HF bands is likely coming only from points east of the antenna, where darkness and ionospheric propagation are prevalent. To the west, daylight absorption in the lower layers of the ionosphere substantially reduces the noise levels at the vertical. This is true of the noise ("QRN" in the radio shorthand known as Q-signals) from

distant thunderstorms. Nearby thunderstorms can be heard equally well regardless of their location relative to the vertical, since the propagation mode is likely direct line of sight from clouds in the region to the antenna. Human-caused noise ("QRM") is usually nearby and unaffected by ionospheric propagation conditions. QRM often originates in digital circuits and electronics. We have little or no control over the direction of these sources from our vertical, so their compass headings are likely to be randomly distributed.

We can improve our ability to hear and to be heard on any frequency—especially on the MF and lower HF bands—by operating two or more vertical antennas in an array. An array of verticals involves multiple antennas in close proximity, each fed at a specified relative phase angle and amplitude to produce the desired radiation pattern.

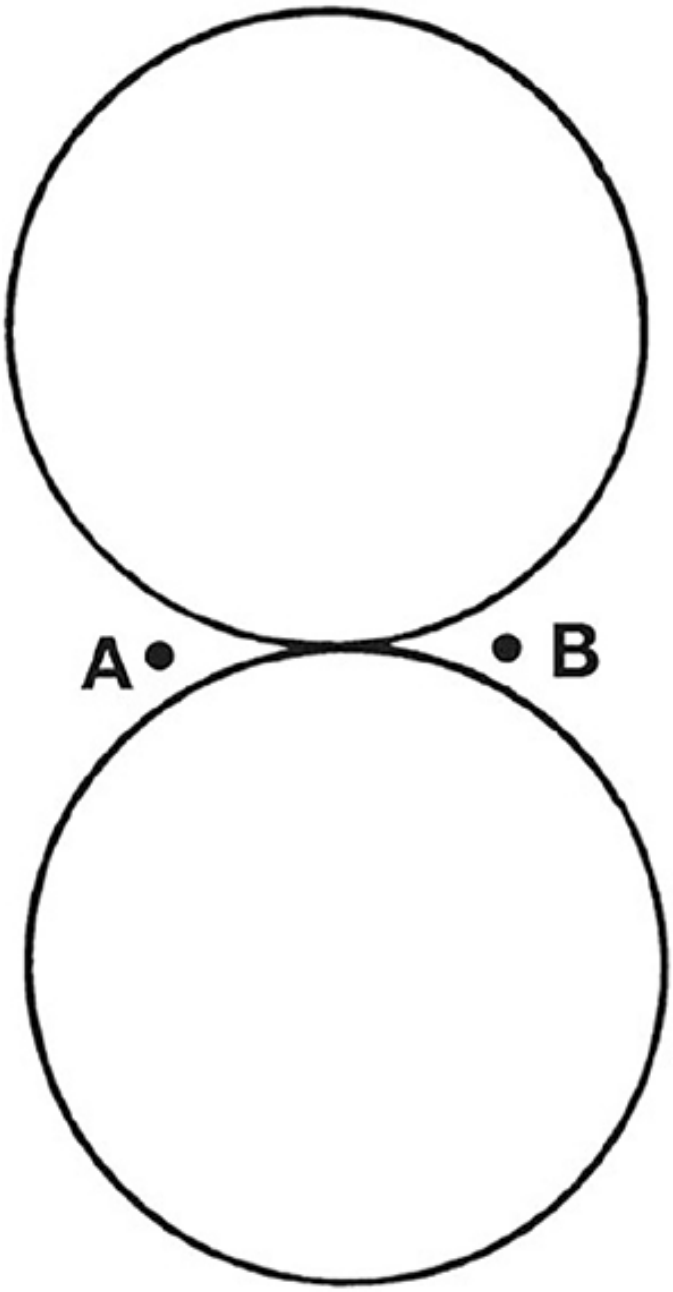
As we saw in Chap. 7, arrays can provide increased field strength in some directions at the expense of strength in other directions. In other words, the wire array increases the antenna gain in a specific direction. On receive, directivity can increase the received signal power from a given direction and decrease the noise received from undesired arrival angles.

AM broadcast stations have used arrays of vertical antennas for decades. Engineers have designed many combinations of verticals. Standard patterns for two-, three-, four-, and five-vertical arrays dating from before World War II exist in the literature. The more complicated ones are mainly relevant to the AM broadcast band, but the simpler combinations have many virtues useful to amateurs and others. Amateurs and shortwave or broadcast band listeners often have a requirement that AM broadcast stations do not: most arrays are designed so they can be switched or steered to more than one compass heading.

To do the subject of vertical arrays full justice would require a separate book at least as big as this one. This chapter will serve as an overview of some simple but effective arrays that can be formed with proper placement and feed systems.

13.1 Two-Element Array

Figures 13.1 and 13.2 show the radiation patterns for a pair of vertical antennas spaced a half-wavelength (180 degrees) apart and fed in phase or (180 degrees) out of phase. Although other phasing relationships between the drive currents to the two antennas are possible, these are most common because of the simplicity of the required feed networks.



F IGURE 13.1 Pattern of two radiators fed in phase, spaced a half-wavelength apart.

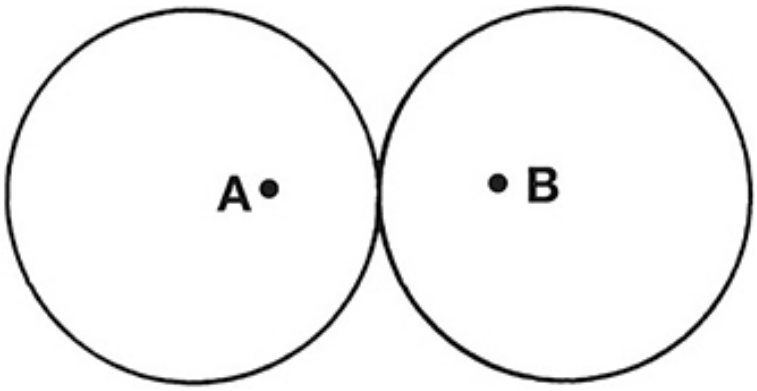


FIGURE 13.2 Pattern of two radiators fed out of phase, spaced a half-wavelength apart.

When the two verticals are fed in phase with equal currents, the radiation pattern of Fig. 13.1 results. It is a bidirectional figure eight that is maximum in the direction that is broadside to the axis of the array. If there are no nearby obstructions and the ground around the two verticals is reasonably flat and featureless, and if the drive currents to the two antennas are well matched in both amplitude and phase, a deep null is formed along the axis of the array.

If the phase of the drive current to one of the elements is shifted by 180 degrees, the pattern rotates 90 degrees and exhibits directivity along the line between the centers (A-B), as depicted in Fig. 13.2. This is often called an end-fire pattern or an end-fire array.

The simplest way to make sure the amplitudes of the feed currents to the two array elements are identical is to bring a common feedline from the transmitter to a point midway between the two elements and then feed each element from there. Figure 13.3 shows the feedline configuration for driving the array elements in phase. The lengths

L_1

and

L_2

must be the same, and ideally they should be cut from the same roll of coaxial cable or hardline so that their impedances and velocities of propagation are matched.

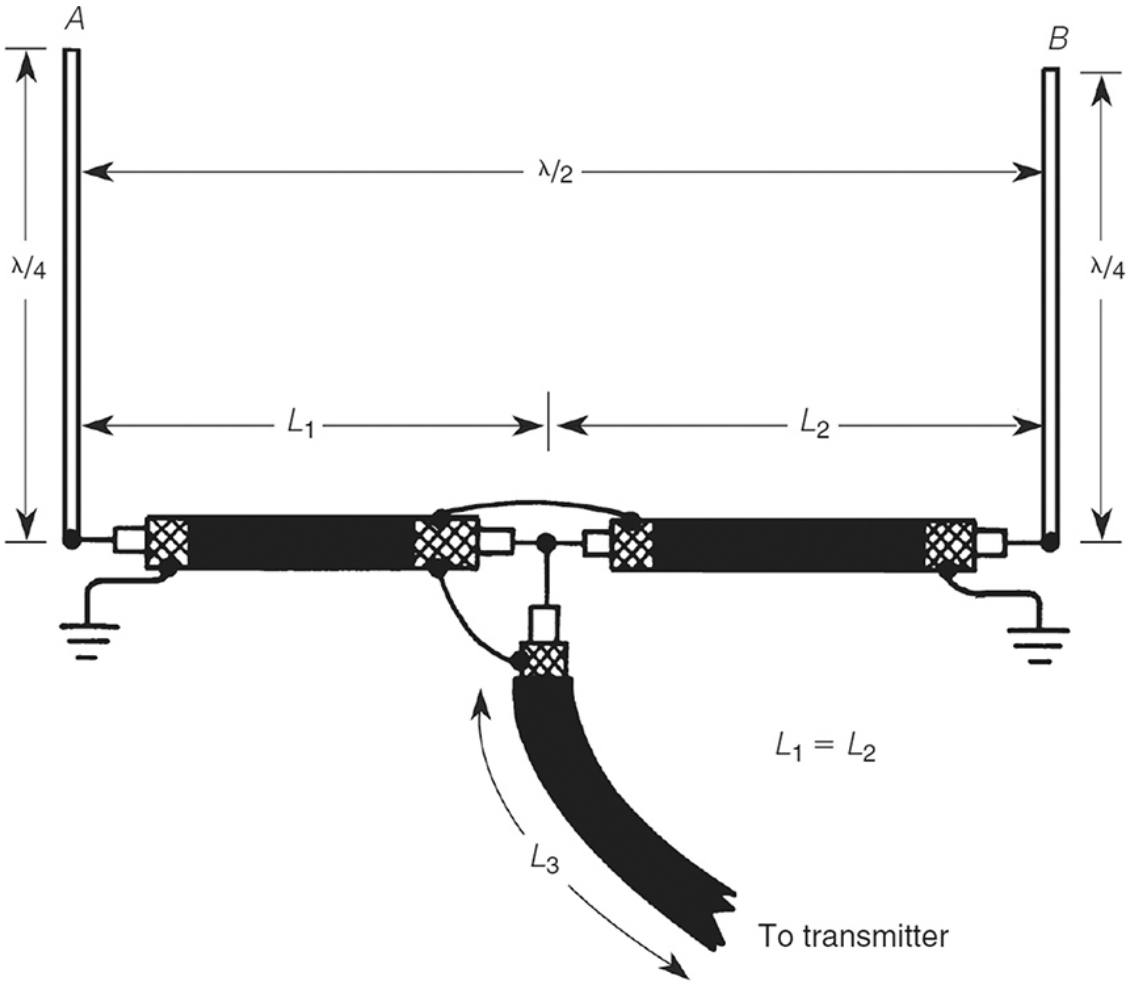


FIGURE 13.3 Feeding a phased array antenna in phase.

To switch the two-element array to its end-fire mode, a 180-degree phase shift must be added to the drive signal delivered to one of the two elements. This phase shift can be provided with a discrete component LC network or a

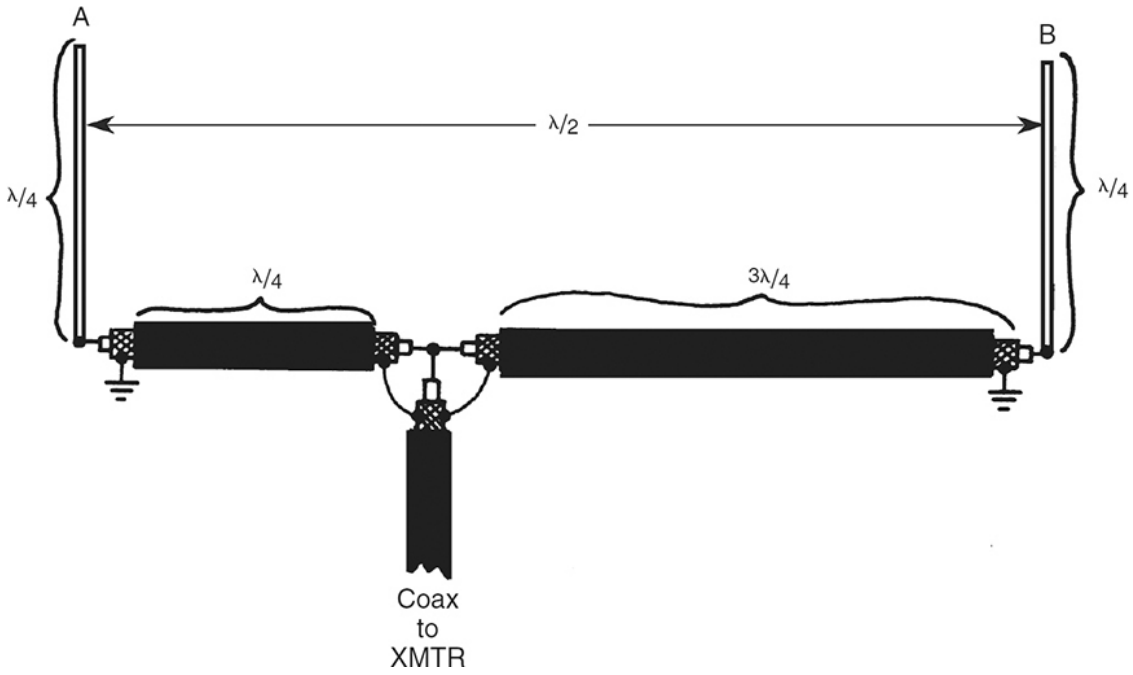
$\lambda/2$

section of the same cable in the line to one element. The section should be half of the wavelength in the coaxial line, which will be smaller than the free space wavelength by the cable velocity factor.

Example 13.1 With standard RG-8U coaxial cable having a velocity factor of 0.67, a half-wavelength section of cable for 3.6 MHz is

$$L\ (\text{ft}) = \frac{492 v_F}{f(\text{MHz})} = 91.57\ \text{ft} \quad (13.1)$$

If we are interested only in the end-fire pattern of Fig. 13.2, we can feed the array as shown in Fig. 13.4. Here again, the lengths are relative to the wavelength at the operating frequency in the coaxial line. To switch back and forth easily between broadside and end-fed modes, we can use the circuit of Fig. 13.5. Two convenient, but equal, lengths of coaxial cable (L_1 and L_2) are used to carry RF power to the antennas. One segment (L_1) is fed directly from the transmitter's coaxial cable (L_3), while the other is fed from a phasing switch. The phasing switch is used to either bypass or insert a phase-shifting length of coaxial cable (L_4).



F IGURE 13.4 Minimum length feed for a phased array antenna.

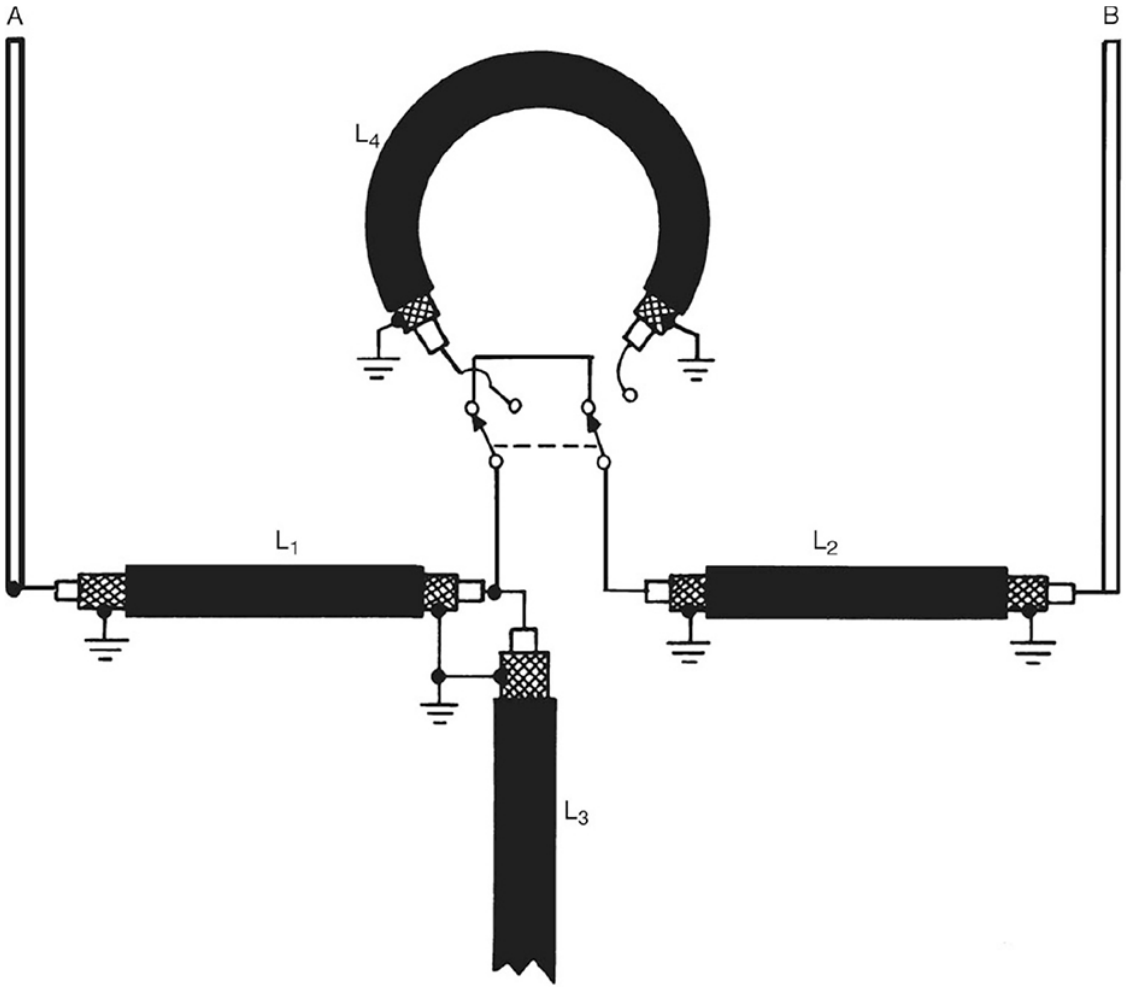


FIGURE 13.5 Phase-shifting antenna circuit.

In principle, this technique allows the selection of varying amounts of phase shift from 45 degrees to 270 degrees. At angles other than 0 degrees and 180 degrees, the currents no longer split equally between the two elements because the element feedpoint impedances are not the same. For angles other than 0 degrees or 180 degrees, more complicated feed systems, including current-forcing techniques, may be needed.

With unequal lengths of transmission line feeding the elements of an array, unequal amplitudes are a likely by-product. The most noticeable effect for small differences, such as would occur when adding a

$\lambda/2$

section of cable to implement the end-fire mode, will be to cause the nulls to fill in and lose some of their depth.

With the addition of a phasing transformer, a two-element vertical array can be remotely switched between broadside and end-fire modes while preserving balanced drive currents at the two elements. Figure 13.6 shows how a two-element array is fed through such a two-port phasing transformer. The phase-reversing switch

S_1

may consist of coaxial relays controlled from the transmitter end of the transmission line. The transformer itself is made from a 1:1 toroidal balun kit such as those available from Amidon Associates and others.

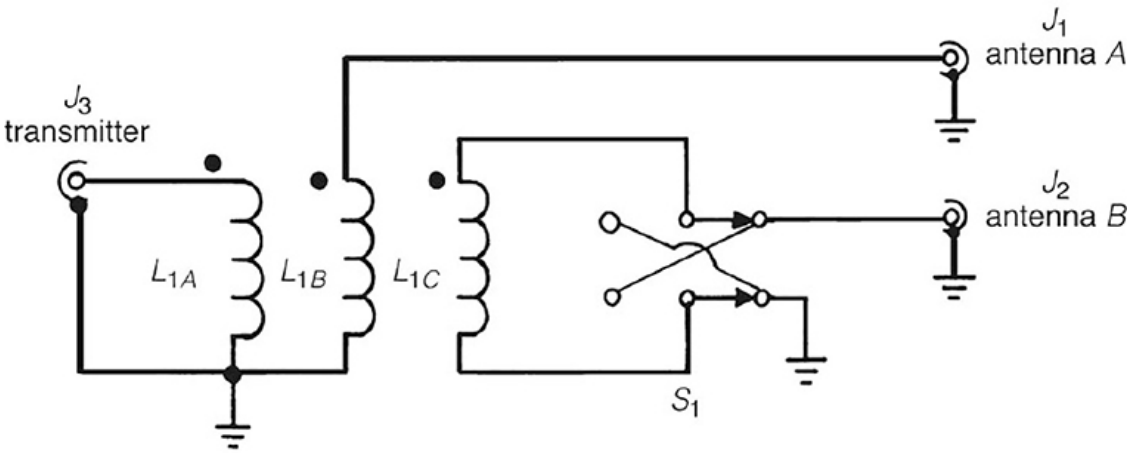


FIGURE 13.6 Phasing transformer circuit.

The three coils are in trifilar style. The dots in Fig. 13.6 show the sense of the coils, and they are important for correct phasing. If the dot end of the first coil is connected to

J_3

(where the feedline from the transmitter or receiver attaches), connect the dot end of the second coil to the 0-degree output (J_1 , which goes to array element A). The third coil is connected to J_2

through a DPDT low-loss RF relay or switch rated for the power levels involved. Although the transformer adds a small amount of loss, it is applied equally to both elements, so the amplitudes of their drive currents remain equal regardless of which mode the array is operated in.

As shown in Fig. 13.7, the coaxial cables to the individual elements have identical lengths from the point where they connect to

J_1

and

J_2

on the transformer housing to the base of either element.

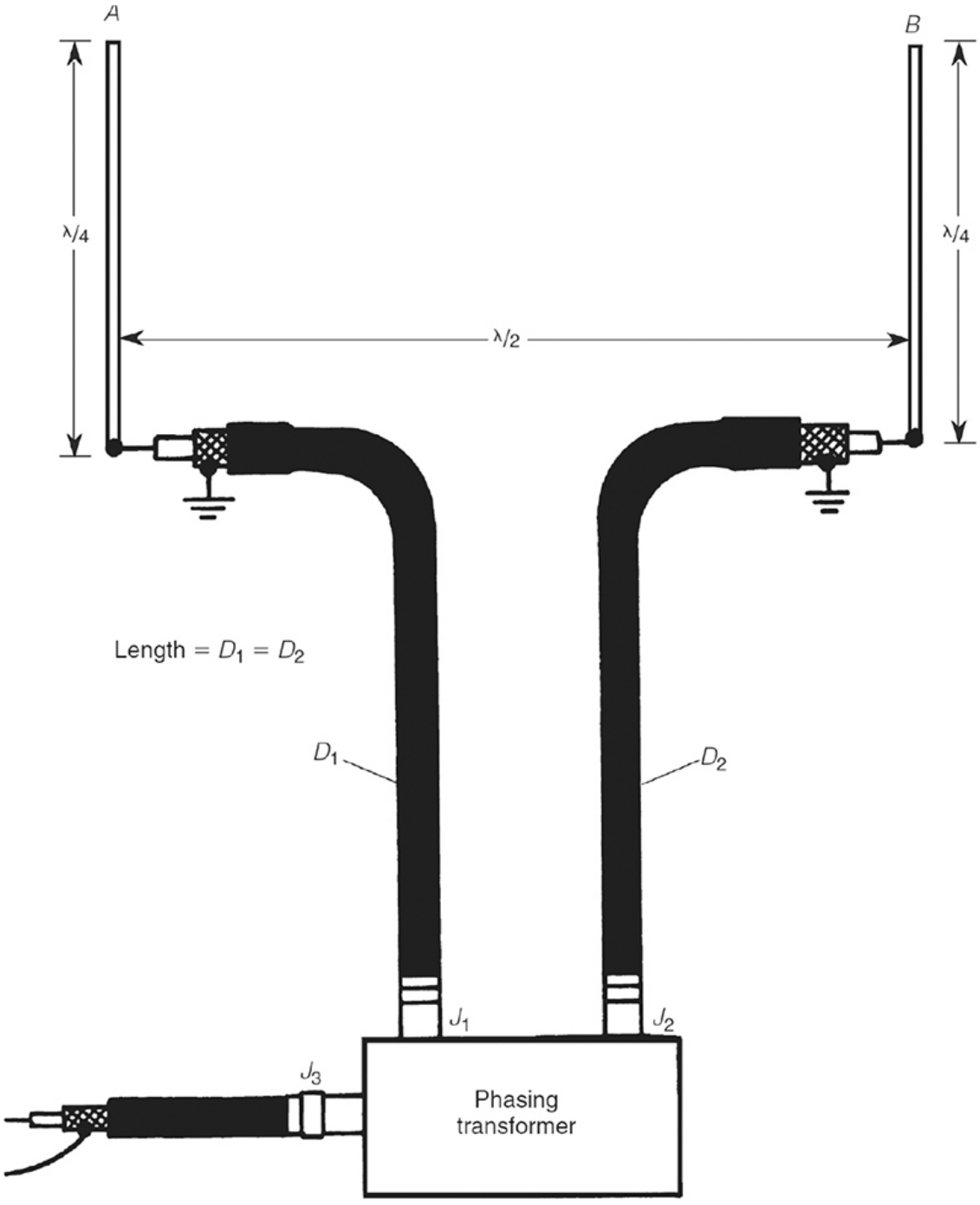
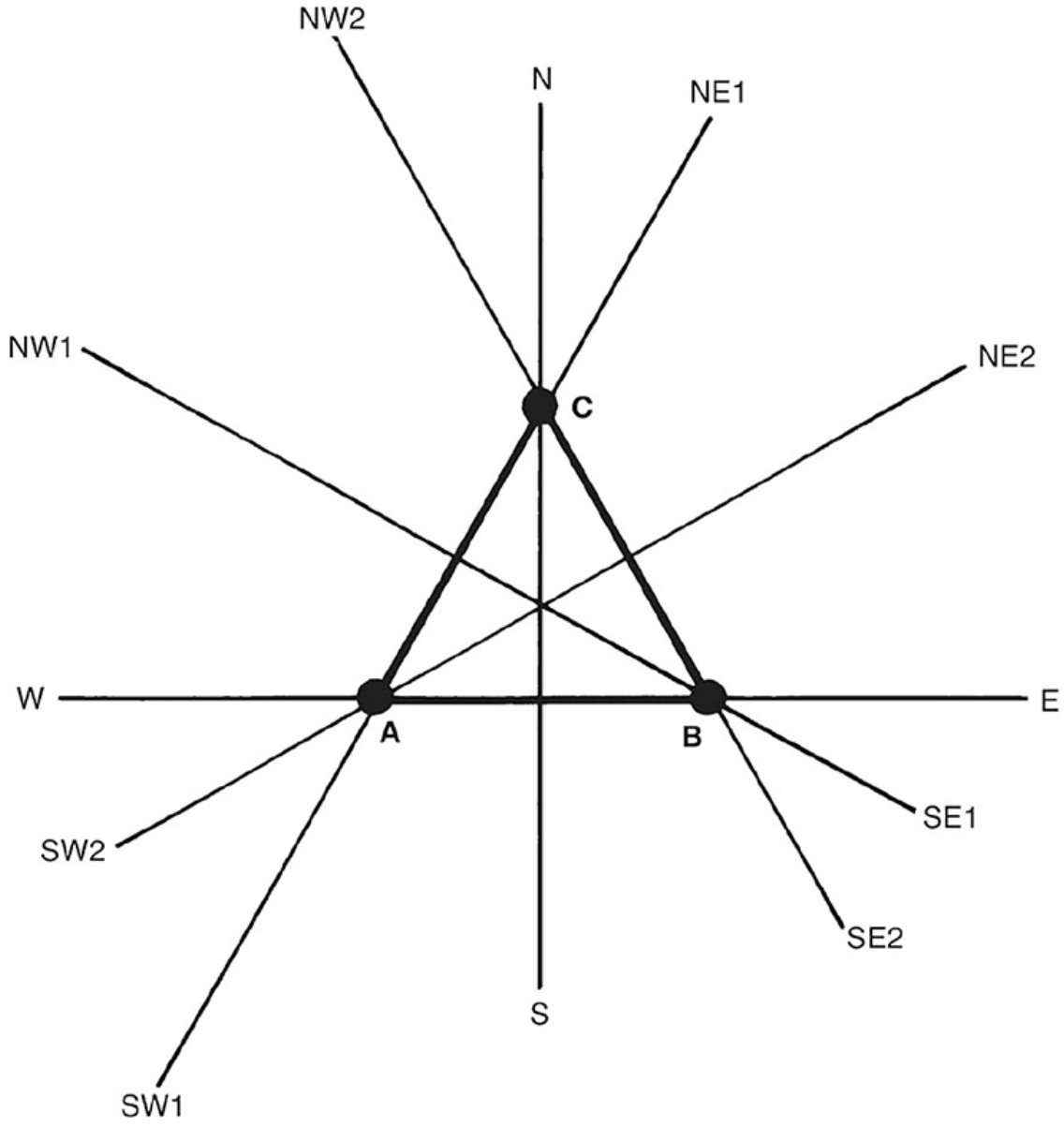


FIGURE 13.7 Connection to antennas.

13.2 360-Degree Directional Array

The vertical antenna array concept can be used to provide round-the-compass steering of the antenna pattern. Figure 13.8 shows

how three
$\lambda/4$
verticals (arranged in a triangle that is a half-wavelength on each side) can be used to provide unidirectional and bidirectional patterns through combinations of in-phase, out-of-phase, and grounded elements. For a specific pattern, any given element (A, B, or C) is grounded ("passive"), fed at 0-degree phase, or fed with 180-degree phase shift. By adjusting the phasing to the vertical array or switching elements between active and passive, the main beam of the array can be steered in 30 degree increments around the compass. Depending on the driving phase and which elements are active, the gain can vary by a few dB. The elevation angle of the main beam also shifts somewhat.



F IGURE 13.8 Three-element phased array.

If space is at a premium, or if multiband operation of an array of trap verticals is desired, the same equilateral triangle can be used with shorter spacings between the verticals. With spacings other than

$\lambda/2$

, 180-degree phase shifts in drive current are no longer matched by the phase shift undergone by the radiated wave as it heads from one element to another. With $\lambda/8$ spacing at 80 m, for example,

phase shifts of 45 degrees and 90 degrees occur on 80 m and 40 m, respectively, and multiples of 360 degrees (i.e., in phase) on bands above 20 m. Nonetheless, directive radiation patterns are possible with such a system.

13.3 Four-Square Array

Another way to obtain directivity that can be switched to multiple compass headings is the four-square (or 4-square) array. In its original implementation, four
$\lambda/4$
vertical monopoles are located at the corners of a square that is
$\lambda/4$
on a side. Drive currents of equal amplitudes are applied to all four elements with three different phases. To develop maximum gain along a diagonal of the square, if the rear vertical on the diagonal is defined to have zero phase shift, the front vertical (i.e., the one in the direction of maximum gain) is fed at -180 degrees and the two side verticals are both fed at -90 degrees. By switching the phase of the drive current going to each specific vertical, the array pattern can be shifted by any multiple of 90 degrees around the compass.

An amateur DXer in the northeastern United States would likely orient his or her 4-square so as to put the center of the four switchable pattern lobes at 60-, 150-, 240-, and 330-degree compass headings to favor Europe and North Africa, South America, Australia and New Zealand, and the Far East.

Variations of the 4-square utilize different combinations of phase shifts or smaller spacings between elements. In most cases, 4-squares are constructed with ground-mounted monopoles and extensive radial fields, but many employ elevated radial systems. If the user has the means to supply drive currents with arbitrary phase and amplitude relationships to the four elements, many customized azimuth patterns can be obtained.

13.4 Receive-Only Vertical Arrays

Much work has been done on phased arrays for receiving at MF and the lower HF bands. In contrast to the challenges of weak signal reception at VHF and above, high atmospheric noise below 10 MHz gives the system designer some latitude in selecting techniques for teasing weak signals out of the background noise. At these low frequencies, receive-only arrays differ from their transmitting antenna counterparts in two ways:

- At bands where sky and environmental noise is dominant, antenna efficiency is of only secondary importance. Array elements much shorter than $\lambda/4$ with low feedpoint impedances and inexpensive tubing and support requirements can be employed. If needed, preamplifiers can be added at the array.
- Because no transmitted RF energy is directly applied to the elements, there is no need for high-voltage or high-current components that are normally required for transmit applications.

Over the past decade or two, 160-m and 80-m DXers fortunate enough to have the necessary acreage have designed and built six-circle and eight-circle arrays for these bands. Typical of these arrays, which provide outstanding receive directivity and signal-to-noise ratio (SNR) on these notoriously weak-signal frequencies, is the eight-circle for 160 m designed by Tom Rauch, W8JI. The output signals from four of eight short vertical elements equispaced around the circumference of a circle are combined with appropriate relays and phasing lines to provide more than 7 dB gain in the desired direction, relative to a single such element. Because the array is eight-way symmetrical around the compass, the main lobe can be switched in 45-degree increments.

At any given switch position, only four of the eight elements are active: two adjacent elements on one side of the circle form an end-fire two-element array, as do two adjacent elements directly opposite, on the far side of the circle. These two end-fire arrays are combined in a two-element broadside array, thus enjoying high

forward gain and deep nulls to the rear, thanks to the principle of pattern multiplication. The eight-circle and the four-element combined end-fire/broadside array building block are described in detail by John Devoldere in the excellent book *ON4UN's Low-Band DXing*, published by ARRL.

Although circle arrays for a given frequency range require less acreage than four bidirectional Beverage antennas, a 1.8-MHz eight-circle requires at least three acres. John Kaufman, W1FV, has designed and published a compact dual-band version of a nine-circle array requiring a very modest 140-ft-diameter footprint (corresponding to slightly more than a half-acre). The W1FV array consists of eight elements around the periphery of the circle, surrounding a single element at its center. It functions as a three-element in-line array that can be switched in 45-degree heading increments. The 160-m design acquits itself quite nicely on 80 m, eliminating the need for a separate 80-m receiving array.

High-performance arrays of short receiving elements require careful design and construction. Close matching of signal amplitudes and differential phase shifts from element to element are needed. Poor feedline installations or sloppy grounding techniques can lead to common-mode signal pickup and other ills that may nullify the directivity and SNR improvements possible with these antennas.

13.5 Grounding

Like ground-mounted or ground-plane vertical antennas, vertical arrays need an adequate ground system. Each element of an array of vertical monopoles should have its own extensive ground system. The ground may be saltwater, sheets of copper flashing, or two dozen or more radials each roughly as long as the vertical is tall. If radials or other conductors are employed, they should not be allowed to overlap unless they are insulated or unless all points of possible contact have been soldered or mechanically joined. Intermittent or oxidized joints can raise the background noise level

when receiving and result in the radiation of spurious out-of-band signals when transmitting. More information can be found in Chaps. 11 and 33.

CHAPTER 14
The Yagi-Uda Beam Antenna

Many antenna arrays have only driven elements connected to signal sources, but some types of arrays also use elements not connected to signal sources, or parasitic elements. Parasitic elements scatter the fields radiated by driven elements. By adjusting the length and other characteristics of the parasitic elements, the phase and amplitude of the scattered field can be controlled so that the overall radiation from the array has high gain in a particular direction.

One of the most popular parasitic arrays is the Yagi-Uda beam antenna or, more commonly, the Yagi. The Yagi was first described publicly by its inventors in 1926. The Yagi is a balanced antenna and is often used for horizontal polarization, but vertically polarized Yagis for the 11-m citizens band are also used. Some arrays of verticals used by broadcast stations and amateurs can be viewed as being similar in operating principle to a Yagi as well.

The advantages of the Yagi are gain, directivity, simplicity of construction, single-element feedpoint, and a relatively compact size. A three-element Yagi for 20 m can, in the footprint of a 50-ft-diameter circle, provide a more constant forward gain rotatable through all compass directions than can a half-dozen fixed rhombic wire antennas, each 400 ft long on its longer axis.

The Yagi is sometimes referred to as a beam, perhaps referencing both its horizontal extent in the most common form and its high gain. Usage examples include, "this antenna is a three-element beam" or a "three-element triband beam."

A directional beam antenna provides several advantages over a more omnidirectional antenna. A dipole has a gain of approximately 2 dB in its best directions. With one additional element, the gain can increase by 3 dB, and with two, the increase can be up to 6 dB.

That’s equivalent to swapping a 25W transmitter for a 100W unit. On receive, a high-gain antenna receives more signal power from a desired direction and less noise and interference from the sides and rear of the desired signal direction, improving the received SNR. When considering an upgrade, it can be better to improve the antenna rather than the power amplifier, as a higher gain antenna helps both receiving and transmitting. The beam antenna pattern has sharp nulls in the azimuthal and the elevation response patterns, which can be used for direction finding. By slightly adjusting the heading of the beam antenna to put interfering signals in the nulls and suppress them by as much as 30 to 40 dB.

14.1 Theory of Operation

Figure 14.1 shows in schematic form the basic three-element Yagi-Uda antenna. Elements of a Yagi normally lie in the same plane, are parallel to each other, and have their centers in a straight line. In a rotatable Yagi, these constraints are best met by attaching the centers of all elements to a common boom. The boom may serve as an electrical connection.

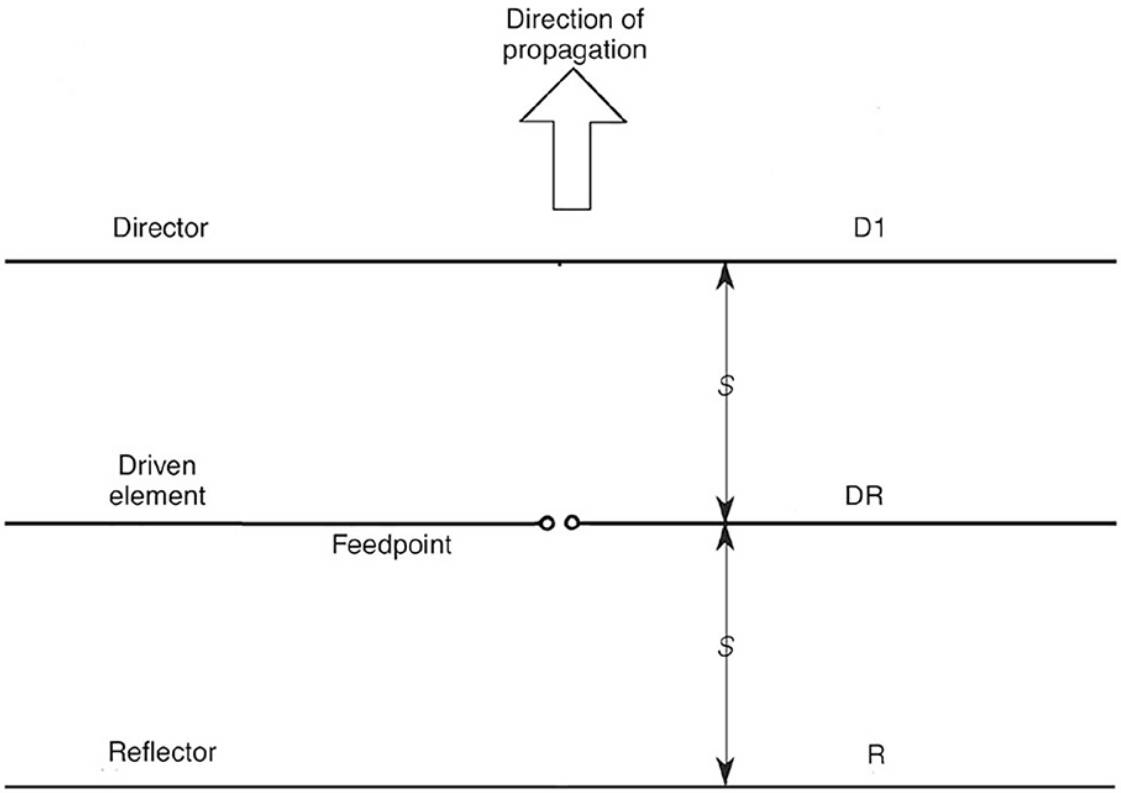


FIGURE 14.1 Basic Yagi-Uda antenna.

The azimuthal pattern of a Yagi antenna is shown for low takeoff angles in Fig. 14.2. The half-power beamwidth is indicated on the plot. The radiation pattern also exhibits sidelobes and back lobes. Sidelobes represent wasted power transmitted and unwanted opportunities for interference while receiving. The goal of the antenna designer for high-gain antennas is to increase the amplitude of the main lobe while decreasing that of the sidelobes and back lobes. With careful design, it is possible to reduce sidelobes and back lobes with only a slight decrease in the main lobe gain or other desirable parameters.

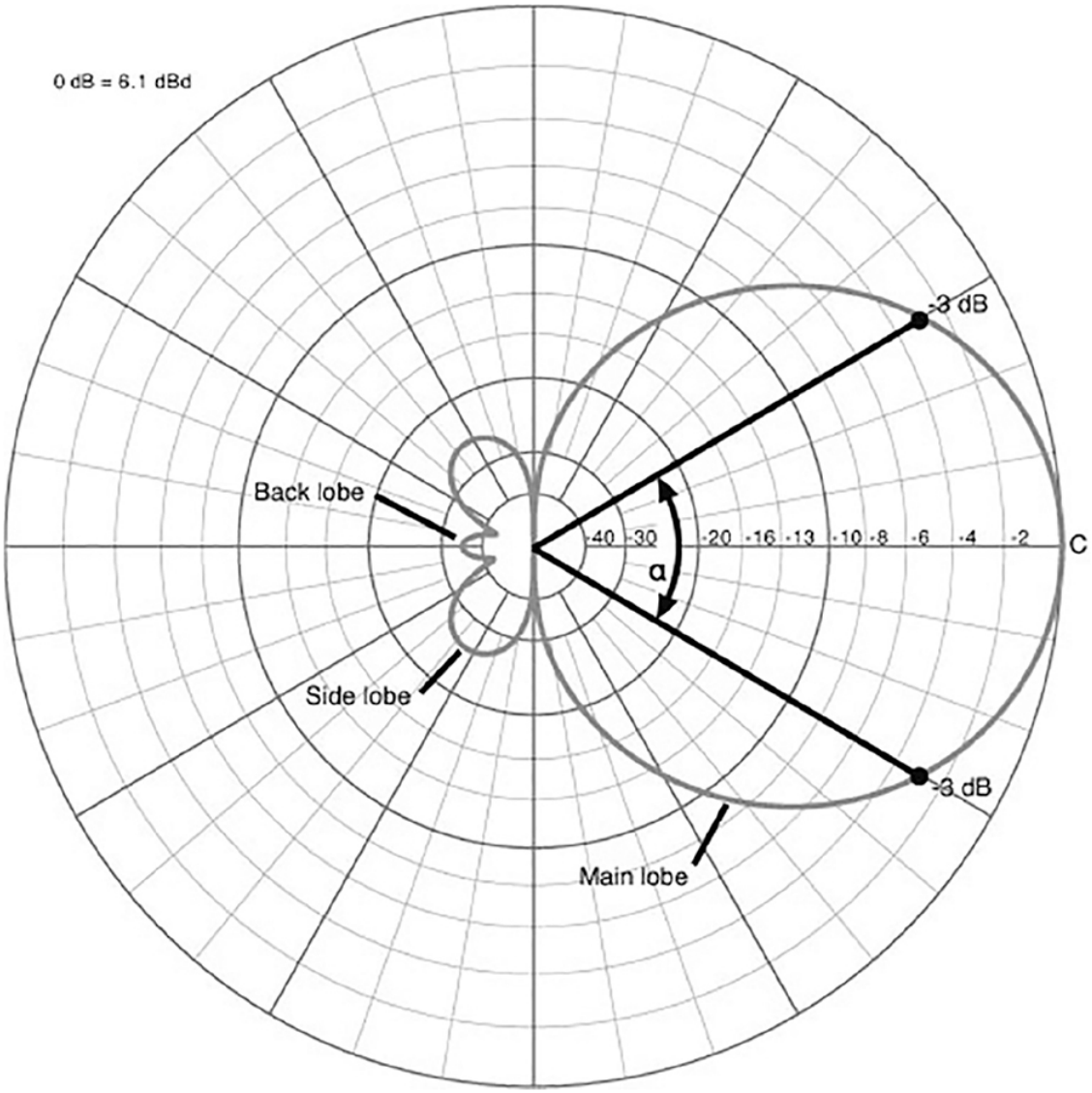


FIGURE 14.2 Pattern of a typical beam antenna.

In addition to having a driven element—which is similar to a standard half-wavelength dipole fed at or near the center by any of a number of different means that we will discuss later—Yagi antennas employ two additional types of elements: reflectors and directors. These are called parasitic elements because they are not directly connected to the feedline but instead receive energy radiated from the driven element and then reradiate it. Reradiated energy from the parasitic elements combined with radiation from the driven element causes peaks and nulls in the radiation pattern as a function of angle.

A detailed analysis of the Yagi antenna is beyond the scope of this book, but its operation can be understood qualitatively. When fields radiated by the driven element, typically a dipole, reach a parasitic element, they induce currents in that element. Each parasitic element is shorter (when used as a director) or longer (when used as a reflector) than a half wavelength, so the induced currents change in phase relative to the fields from the driven element. The fields scattered by the parasitic elements combine in the far field of the antenna. Because the radiation from each element has its own unique phase and amplitude, the resulting radiation pattern exhibits peaks and nulls depending on the direction around the Yagi.

While any object around a driven element will scatter fields, for the regular, linear arrangement of the parasitics in this type of antenna, Yagi and Uda derived equations for the radiation patterns that can be used to design the lengths and locations of the parasitic elements to achieve a high-gain radiation pattern. From these equations, we can calculate optimized dimensions for the diameter, length, and spacing of the elements in a Yagi antenna.

If a reflector is used, it is placed behind the driven element relative to the direction of desired maximum signal radiation and it is typically a few percent longer than a half wavelength on the conducting element at the operating frequency. Directors are placed in front of the driven element. The director nearest the driven element is typically about 3 percent shorter than a half wavelength in the conductor.

Spacings between adjacent elements of a Yagi typically are between 0.1 λ and 0.25 λ, where λ is the wavelength in air. If weight, turning radius, and wind load are more important than bandwidth, a three-element Yagi can fit on a boom that is only 0.2 λ in overall length, but other attributes, such as standing wave ratio (SWR) bandwidth, front-to-back ratio, and forward gain, have been compromised.

The basic Yagi is a planar antenna, with all elements lying in the same plane and sharing the same orientation. There is no fixed rule regarding the number of either reflectors or directors, but it is

common to use a combination of a single reflector and one or more directors in addition to the driven element. A typical three-element beam consists of a driven element flanked by a reflector at one end and a director at the other end of a boom. Some four-element Yagis have two directors, but two driven elements can also be employed to provide maximum gain and directivity over a wider bandwidth.

14.2 Three-Element Yagi Antennas

For a three-element HF Yagi with full-size elements, there are six important dimensions:

- Length of the driven element (DE)
- Length of the reflector (R)
- Length of the director (D)
- Diameter of the elements (d)
- Centerline spacing between the driven element and the reflector (R-DE)
- Centerline spacing between the driven element and the director (DE-D)

At VHF and UHF and for some HF designs the dimensions of the boom-to-element attachment method need to be incorporated in the design calculations. The sum of the interelement spacings plus the width of one boom-to-element bracket determines the minimum required boom length for a specific design.

The dimensions work together to determine the antenna performance. If a Yagi design meets specifications for forward gain, front-to-back ratio, and SWR, a change in the spacing between the reflector and the driven element may worsen some or all of the parameters. The dimensions for a multielement Yagi constructed of #12 copper wire are usually different from those for the same beam with elements of 1-in aluminum tubing. Because of the close coupling between elements, multielement Yagis are more sensitive to changes in dimensions than simple dipoles.

Example 14.1 A
$\lambda/2$
-dipole for 20 m constructed from #12 copper wire is trimmed for a purely resistive feedpoint impedance at 14.1 MHz. The resulting length of the dipole is 33 ft 3 1/2 in. When the wire is replaced with 1-in aluminum tubing, the dipole resonates at 33 ft 2 1/2 in. The difference, 1 in, is usually inconsequential to the overall operation of the dipole.

Example 14.2 A three-element Yagi with elements of 1-in aluminum tubing meets the following specifications at 14.1 MHz:

- $\text{Forward gain} = 8.25 \text{ dBi}$
- $\text{Feedpoint impedance} = 13.8 + j0\ \Omega$
- $\text{Front-to-back ratio} = 31 \text{ dB}$
- $\text{Front-to-side ratio} = 25 \text{ dB}$

With the use of a simple 4:1 balun, the input impedance becomes 55 Ω and the 2:1 VSWR bandwidth using 50-Ω coaxial cable is 220 kHz.

If we replace the aluminum elements with #12 copper wire of exactly the same length, the performance of the resulting beam at 14.1 MHz becomes:

- $\text{Forward gain} = 7.5 \text{ dBi}$
- $\text{Feedpoint impedance} = 24.2 - j17.0\ \Omega$
- $\text{Front-to-back ratio} = 16.3 \text{ dB}$
- $\text{Front-to-side ratio} = 16.3 \text{ dB}$

Each performance parameter of the three-element beam has been compromised changing the element diameter. To obtain a reasonable SWR at the design frequency, we must switch to a 2:1 balun to obtain a transformed input impedance of 48.4 – *j*34 Ω. This corresponds to an SWR of 1.9:1.

The dimensions can be adjusted to improve performance with the wire elements, but for comparable front-to-back (F/B) and front-to-side (F/S) ratios and a feedpoint impedance that is a good match for the 4:1 balun, we can obtain about 8 dBi forward gain. The resulting 2:1 SWR bandwidth is only 125 kHz—one consequence of making the

element diameter much smaller. In the process, we have had to lengthen the director by 3 in and slightly shorten the reflector for this particular design.

Element diameter also comes into play when translating a design to a different operating frequency. If all dimensions are scaled, the resulting element diameter(s) may be unacceptable: they may be mechanically too fragile, too heavy, or difficult to obtain. If all of the listed dimensions are not scaled by the same factor, performance on the scaled band may decrease. One approach is to use scaling to get into the ballpark with dimensions for the new band based on the original band, then use antenna modeling software (Chap. 29) or experimentation to fine-tune all the dimensions so as to optimize performance around standard aluminum tubing diameters or copper wire gauges.

Many different Yagi designs exist. For a specific Yagi, one or more characteristics of the antenna usually must be compromised, so one design may be developed for a three-element Yagi with wide SWR bandwidth, while another may favor a high F/B ratio over a narrow bandwidth centered on the design frequency.

Some designs are based on mechanical tradeoffs rather than electrical performance specifications. One common constraint is total boom length. A manufacturer might choose a shorter boom length to keep material costs and antenna prices down. An amateur backyard experimenter might choose a shorter boom length because of the limitations of his rotator or the proximity of a nearby tree. For obvious reasons, boom lengths of 20 ft or slightly less are very popular. In particular, 3-in irrigation pipe makes a very strong boom. Some commercial Yagis butt two suitably thick-walled 10-ft aluminum tubes together to form a 20-ft boom, usually with interior and/or exterior reinforcing sections for some distance on either side of the junction.

Since the publication of the original Yagi-Uda paper, engineers and amateur experimenters have improved the available designs. Early work was aimed at VHF. Shortwave broadcasters and the US

military joined hands in bringing the Yagi to the HF bands, and Yagis are now available through microwave frequencies and beyond.

When working with Yagi antennas, these guidelines can be helpful:

- Exact interelement spacings are not critical on a Yagi; they are typically 0.1 to 0.3 wavelengths for popular proven designs. Once established for a design, changes to spacings interact with element lengths and do affect Yagi performance.
- Wider element spacings for a given number of elements provide greater SWR bandwidth and more constant forward gain and F/B ratio across the entire band. Of course, wider element spacings translate to longer booms.
- Maximum forward gain and maximum F/B ratio seldom occur at the same frequency. Longer boom lengths make it easier to optimize multiple parameters over a wider common range of frequencies.
- Yagi gain is roughly proportional to boom length. Adding elements onto a fixed boom usually worsens performance.
- A reflector 3 percent longer than a half-wavelength in the conductor is a good starting point. Said another way, start with a reflector length that is no longer than $\lambda/2$ in free space or air. Similarly, a director is roughly 3 percent shorter than a half-wavelength in the conductor.
- Most HF Yagis are constructed from telescoped lengths of aluminum tubing to minimize mechanical stresses on the inner portions of the elements and the overall load on the boom. Elements and booms are often tapered or stepped in diameter. Stepping or tapering alters the electrical lengths of the elements for a given physical length and must be taken into account in the design and modeling process.

- At VHF, UHF, and higher frequencies, elements of a single diameter are usually used. The details of how elements are connected through, over, or around the boom can have a significant effect on Yagi performance at those frequencies.
- Wire elements are usually longer than elements made from larger-diameter tubing, but stepped-diameter elements are longer than single-diameter ones.
- The length of the driven element plays a minor role in optimizing a Yagi's unique attributes (forward gain, F/B and F/S ratios, and the resistive part of the feedpoint impedance). The driven element length is usually adjusted following all other design choices to make the reactive part of the feedpoint impedance small.
- The input impedance at the center of a typical three-element Yagi is lower than that of a simple half-wave dipole; 10 to 30 Ω is common. A 4:1 balun for the lower end of the feedpoint impedance range, a stub, or a series-section transformer for the upper end of the range is often needed to provide a better match to the transmission line (see Chap. 6).
- Even though it is a balanced antenna, a Yagi can be fed with an unbalanced feedline such as coaxial cable with little impact on forward gain. There may be minor loss of symmetry in the radiation pattern and it may be necessary to add feedline chokes at one or more locations along the cable to minimize common-mode pickup and radiation by the feedline.
- Other methods of feeding a Yagi include the gamma match (also unbalanced), the T-match (a balanced gamma match), and the hairpin match (a balanced stub-matching technique; a variant of the hairpin is used on many Hy-Gain beams under the trade name "Beta match"). Often, a lumped-component balun of the

appropriate turns ratio and located right at the driven element feed terminals is the simplest approach.

- Yagi antennas are available with shortened elements to reduce wind load, rotational torque, and weight. The most common methods for shortening elements include traps, loading coils, and linear loading wires.
- At the top of a support, Yagi antennas are a target for lightning discharges and wind-driven precipitation static. Yagis on a tower may be part of a capacitive top-loading system for 160 or 80 m when the tower is used as a transmitting vertical on either of those bands. The Yagi designs that best address these concerns have both sides of all elements grounded to the boom, either directly or through the matching network.
- Notwithstanding the care with which the element lengths and mechanical details are typically designed, Yagi antennas are usually robust. Long-boom, multidirector Yagis have continued to perform even though ice and/or wind have ripped one or more half-elements from the antenna.

14.3 Yagi Antennas with Four or More Elements

As additional elements are added to a Yagi, they typically are directors. Until the advent of the dual-driven element variant previously described, a four-element Yagi meant a reflector at one end of the boom, followed by the driven element and the two directors. A five-element Yagi usually employed three directors. If the first reflector is properly designed and positioned on the boom relative to the driven element, virtually all of the energy radiated off the back of the beam (i.e., beyond the reflector) will be canceled by the radiation from all the elements. There will be little RF power passing beyond the reflector and there is little need for additional reflectors. Nonetheless, some long, high-performance VHF and UHF

Yagis use three reflectors mounted in a plane orthogonal to the plane of the other elements.

14.4 Yagi Construction

In this section, we will briefly discuss some simple construction methods and inexpensive materials for Yagis. The projects in this chapter are within the reach of most readers with a drill, screwdriver, driver, and hacksaw or tubing cutter, although for more complicated designs, parts can be purchased or harvested from used Yagis.

As the design frequency of a rotatable Yagi drops below 30 MHz or so, mechanical issues come to the fore. Reliable element self-support dictates the use of multiple tubing diameters for the elements, sometimes in conjunction with truss systems. As many as five or six different tubing diameters in each element of a 20-m beam may be used. As element lengths increase, the strength of the boom-to-element clamps or brackets becomes a critical design matter.

14.4.1 Rotatable Dipole

We will start with the driven element of a Yagi, which is simply a half-wave dipole antenna. As we have seen in earlier chapters, the half-wave dipole has modest gain, and by mounting the dipole horizontally and rotating the antenna the main beam can be steered in azimuth.

Basic considerations for the half-wave dipole such as the required antenna length can be found in Sec. 5.4. Because of its large size, rotating a horizontal

$\lambda/2$

dipole for the MF and lower HF bands is a mechanical challenge. Above 10 MHz or so, a rotatable dipole becomes practical, not only because the

$\lambda/2$

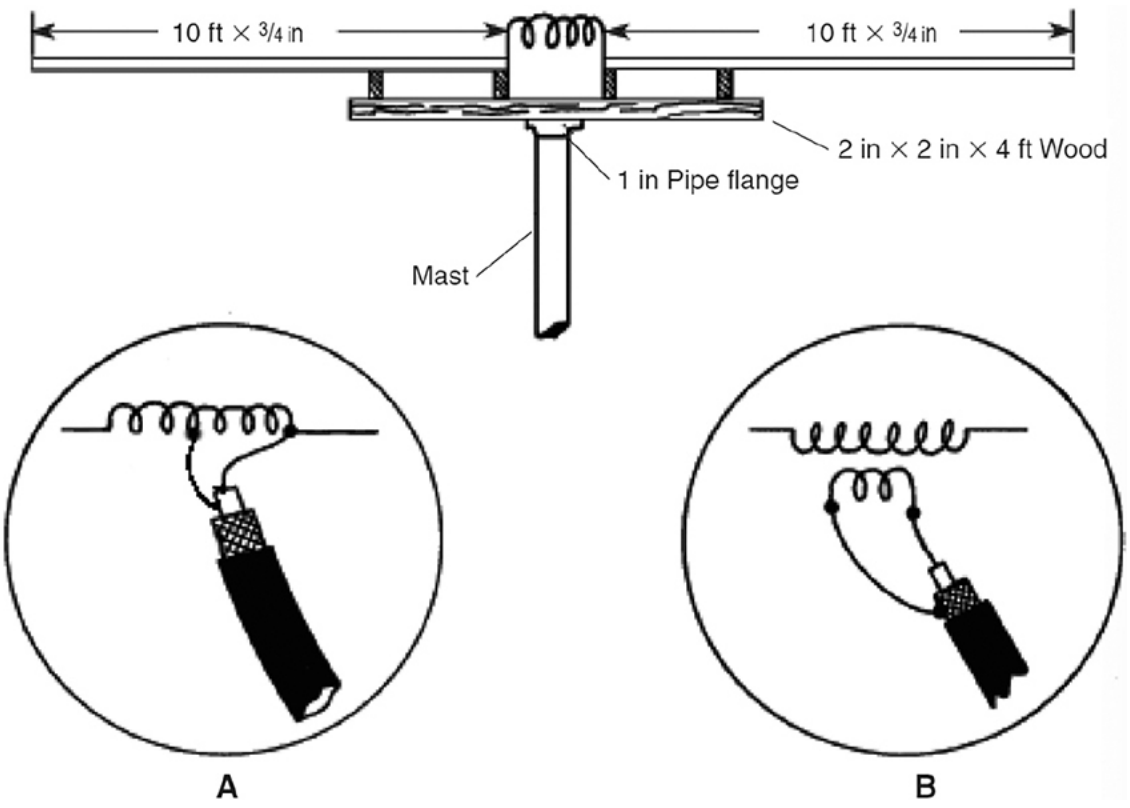
element has a manageable length but because the user doesn't need to lift it as high above ground to make it a reasonable performer. The length of a half wave dipole is approximately 16 ft on 10 m, 22 ft on 15 m, and 33 ft on 20 m; all of these antennas will perform well at a height of 30 ft or more if they are in the clear.

Because HF beams are seldom mounted high enough above ground or the roof of a support building to be treated as being in free space, the length that results in a resonant antenna (i.e., the feedpoint reactance is zero) will differ slightly from the free-space value. For this reason, the dipole length need not be cut to the exact value given by a formula.

Figure 14.3 shows a generic rotatable dipole design for use on 15, 12, and 10 m. On 12 and 10, the radiator half-elements are made from 10-ft lengths of 3/4-in aluminum tubing cut to the lengths listed in Table 14.1, which are based on the dipole being mounted 1λ over average soil. For these bands, the tubing is mounted on standoff insulators, which in turn are mounted on a 4-ft length of

2×2

wood. The wood should be protected with varnish or exterior polyurethane. Direct mounting to wood can be done, but precipitation can create an RF leakage path. The two ends of the element halves near the centerline of the dipole should be an inch or two apart. Connect one conductor of a coaxial cable or hardline to either half-element directly or through a 1:1 balun. The mast is attached to the wood supports using a pipe flange or similar hardware.



F IGURE 14.3 Rotatable dipole antenna. Inset A shows conventional feed; inset B shows transformer feed.

T ABLE 14.1 Rotary Dipole Half-Element Lengths

Band	Design Frequency	Total Length of One Side	3/4-inch Tubing	5/8-inch Tubing
10 m	28.300 Mhz	8′3″ (2.514 m)	8′3″ (2.514 m)	
12 m	24.930	9′5 3/8″ (2.880 m)	9′5 3/8″ (2.88 m)	
15 m	21.150	11′1″ (3.385 m)	10′ (3.048 m)	1′1″ (0.337 m)

For the 15-m version, the 10-ft length of tubing is not long enough to be resonant. It can be used as is with an ATU, a center

loading coil can be added, or extensions can be added to make the element self-resonant at the lower frequency. The first option results in a VSWR on the 75-Ω feedline of roughly 4.0:1 across the entire band. Even though the resistive part of the feedpoint impedance is an excellent match for 52-Ω cable, the high reactance makes the VSWR lower on 75-Ω cable. This option should not be considered if any substantial transmit power (over, say, 100 W for normal cables) is planned. For the second option, a 0.14 μH loading coil is mounted at the center, with its two ends connected to the two half-elements. The dimensions of the coil are four to five turns, 0.5-in diameter, 4-in length. For low power levels, the coil can be made of #10 (or #12) solid wire—and, for higher levels, 1⁄8-in soft copper tubing.

If the center loading coil approach is chosen, there are two basic ways to feed the antenna, as shown in details A and B in Fig. 14.3. One method is to connect the coaxial cable across half the inductor, as shown in detail A. Be sure the center of the inductor is connected to the shield, or grounded, side of the coaxial transmission line. Or, the feed can be link-coupled to the line through a one- to three-turn loop (as needed for impedance matching), as shown in detail B. One way to do this is with a toroidal inductor, which also lifts both sides of the dipole off the shield ground of the cable.

Using a loading coil as a substitute for part of an antenna element reduces the element's effectiveness as a radiator. The difference between an efficient radiator and a discrete (lumped) circuit component is primarily in the physical configuration of the conducting material. When you coil up a portion of an antenna element into a small volume, you convert that portion of the radiator into a nonradiating lumped component. The deleterious effect on the antenna is greatest when this is done within the high-current section of the element. When possible, loading of elements close to their centers should be avoided. Often, other forms of loading can attain the same reduction in length with far less decrease in radiated signal strength.

Frequencies below 21 MHz can be accommodated by increasing the dimensions of the coil or adding turns until the antenna is usable on 17 m or 20 m in addition to 15 m. Or, eliminating the need for a coil by making the

$\lambda/4$

element halves the proper length may be easier. We can do this by extending each tube with one of a smaller diameter. The added tube should extend back into the original tube 12 in or more to provide strength for the cantilevered portion of the outer tube and to leave some extra material for final adjustments to element length. A 24-in-long 5⁄8-in OD tube with 0.058-in wall thickness will nest inside the 3⁄4-in element half to bring the total length of the half element to the required 11 ft 1 in.

Depending on anticipated wind and ice loading, reinforcement of the 3⁄4-in tubing in the area of the insulated standoffs and for some distance beyond may be needed. With standard 0.058-in wall aluminum tubing, the wind survival speed for the 15-m dipole is in the 70 to 80 mph range. Use any leftover 5⁄8-in stock to slide or drive a few feet of it into the feedpoint end of each 3⁄4-in tube that will be supported.

When mating aluminum tubing, ensure that the surfaces are clean and smooth. An antioxidizing compound can also be used. Nested tubes can be held together easily with stainless steel sheet metal screws. Another way to hold nested tubes is to slot the end of the larger tube and tighten a hose clamp over the slotted region after the smaller-diameter tube has been inserted the correct distance, although this approach may not be secure enough for some installations. Pop rivets can be used as well. The antenna can be initially assembled with hose clamps, raised in the air, and the VSWR measured before permanent installation.

Plastic caps supplied for the element tips by antenna manufacturers have positives and negatives. While keeping insects out, they can trap moisture inside the element. Natural element sag is perfectly capable of draining moisture on its own, so a small piece of wire mesh can be used to block insects as well. Alternatively, drill small holes in the stock plastic caps and install them with the holes

toward ground. The plastic caps should be held in place with electrical tape or hose clamps because they are easily dislodged during the antenna raising process. Caps also tend to crack and lose their tight grip on the element when exposed to the sun over long periods.

If the rotatable dipole is constructed of full-length elements so that no center loading coil is employed, feeding it is simply a matter of connecting a 75-Ω coaxial cable (RG-6, RG-59, RG-11, or equivalent) to its center. The center conductor of the cable goes to one side of the element and the shield or braid goes to the other, using attachment techniques detailed in Chap. 31. The two element halves should be separated about 2 in by the supporting hardware. A 2-ft length of hardwood dowel turned down to the inside diameter (ID) of the element tubing and treated with preservative such as an exterior polyurethane can be used as the center insulator. Strain relief for the cable must also be provided.

14.4.2 Multielement Yagi Construction

Each element of a rotatable multielement Yagi antenna can be constructed in a manner similar to that used for the rotatable dipole described earlier. The only significant difference is that the parasitic elements (reflector and directors) do not have a feedline attached to them and the metal is continuous across the centerline of the element.

Adding parasitic elements to the rotatable dipole described previously introduces a number of issues that must be dealt with if the advantages of the parasitic array are to be realized. Most commercially available rotatable dipoles and multielement Yagis for the HF bands use nested tubing—sometimes relying on five or six different diameters that become progressively smaller the farther out on the element one gets from the boom. These are known as tapered or stepped-diameter elements. Tapering affects the electrical length of the element and corrections must be made to achieve maximum gain or front-to-back ratio. Some modeling software packages such as EZNEC include stepped-diameter

correction algorithms. For a single dipole, tapering may not matter much, but with parasitics, the bandwidth of the antenna is narrower and its sensitivity to dimensional changes increases. The error created by ignoring changes in element diameter over the length of the element can be important.

A rotatable multielement beam requires a boom to support the elements in a common plane while holding them in fixed positions and orientations. In the case of a metal boom, the driven element is frequently insulated from the boom, depending on the specific feed system chosen. The parasitic elements are often mounted directly to it. Wood is a readily available material that is suitable for the smaller beams of 24 MHz and above. Below that frequency, the longer element lengths and increased spacings usually require an aluminum boom. Aluminum tubing suitable for use as a boom as well as brackets and other needed parts can be harvested from used antennas or purchased from antenna suppliers. Various alloys are available with different strengths. In the book *Physical Design of Yagi Antennas*, David B. Leeson, W6NL warns readers that 6061-T6 and 6063-T832 are better choices for antenna builders than the more commonly stocked 6063-T5.

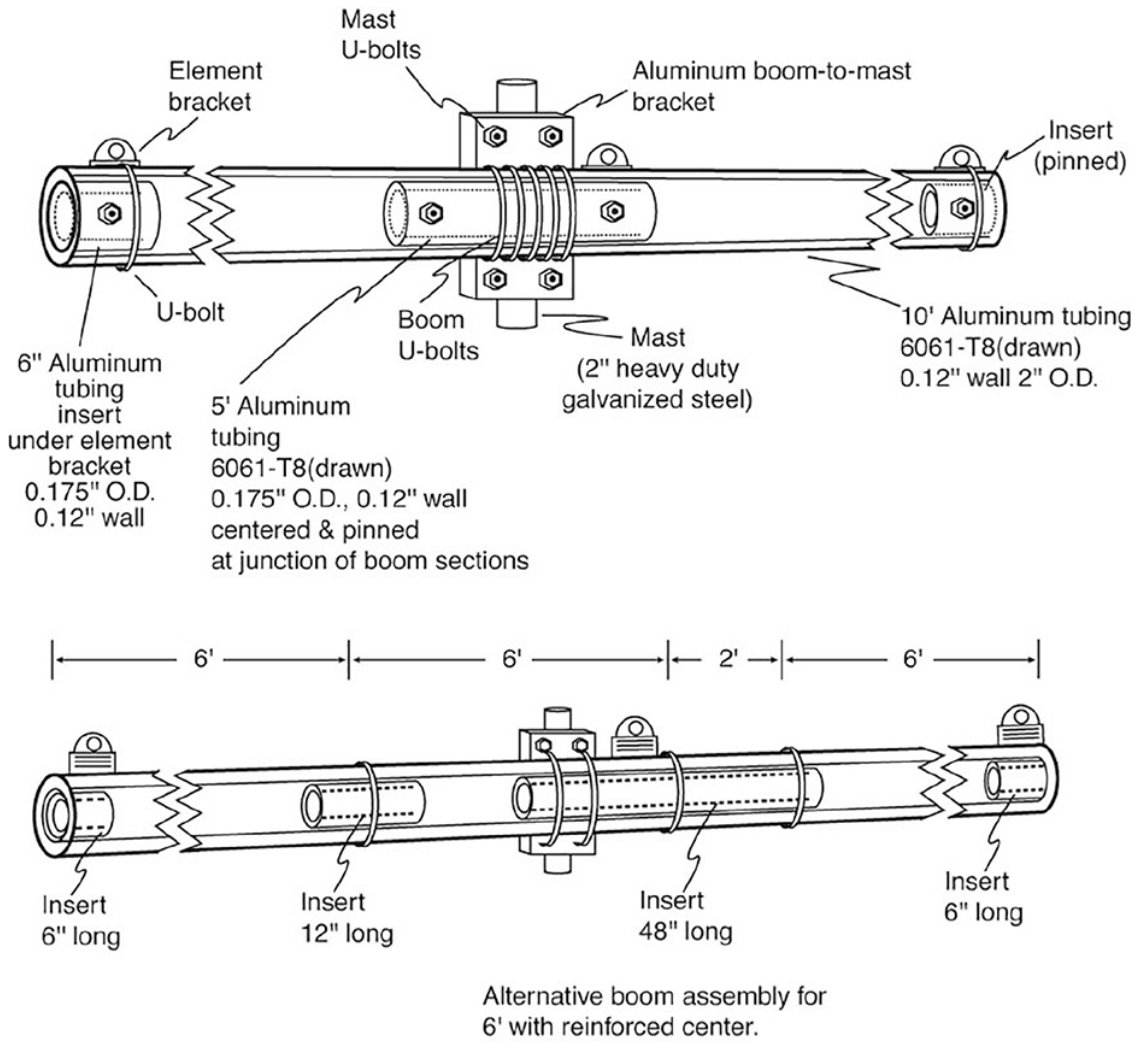
Total boom length and the spacing of elements along the boom are important parameters in the design of any multielement Yagi. Boom lengths for three-element Yagis are typically between 0.15 λ and 0.5 λ. Longer boom lengths often mean wider SWR bandwidths and make it easier to achieve maximum gain and high F/B ratios across the same frequency range. If narrower bandwidths and a slight reduction in maximum forward gain are acceptable to the owner, excellent performance can be obtained from beams with shorter booms.

14.4.3 Rotatable Three-Element 20-Meter Yagi

At frequencies below about 24 MHz, the elements of rotatable Yagis are usually formed from sections of tubing with diameters that get smaller with distance from the boom or center of the element. Such a design approach can provide wind and ice survivability with

minimum weight elements. For 15 m, each half-element will likely consist of a largest diameter center section and perhaps three progressively smaller diameter sections. On 20 m, a center section and four to six progressively smaller diameters are common. Below 20 m, survivability issues usually lead to design approaches involving substantially more complicated methods such as trusses for strengthening booms and elements. Twenty-meter represents roughly the lowest frequency for which a simple design like the one presented here is mechanically adequate for the majority of wind environments.

Stepped-diameter correction of each element's length can be done manually using correction formulas such as those published by David Leeson or using antenna modeling software (such as W7EL's EZNEC 5+) that incorporates the needed corrections. Once an element taper schedule such as that of Fig. 14.4 has been created based on mechanical considerations, the lengths and diameters of the various element sections needed for the correct element length are calculated.



F IGURE 14.4 Stepped-diameter 20-m Yagi reflector half-element.

A convenient boom length for a rotatable three-element 20-m Yagi is 20 ft. The boom can be a single length of common 3-in tubing or a pair of 2-in

$$\text{OD} \times 10\text{-ft}$$

aluminum tubes butted together with a short internal or external reinforcing tube and a boom-to-mast bracket of suitable dimensions, as depicted in Fig. 14.5. Also shown is an alternative approach to constructing the boom from multiple sections of aluminum tubing if the longest that can be obtained conveniently are 6-ft sections. The tubing should be reinforced with properly

nested reinforcing inserts at each point where U-bolts encircle the boom. For 2-in OD boom material with 0.120-in wall thickness, 0.175-in OD tubing makes excellent inserts. Inserts should be pinned to the outer tubes with two stainless steel bolts at right angles to each other.

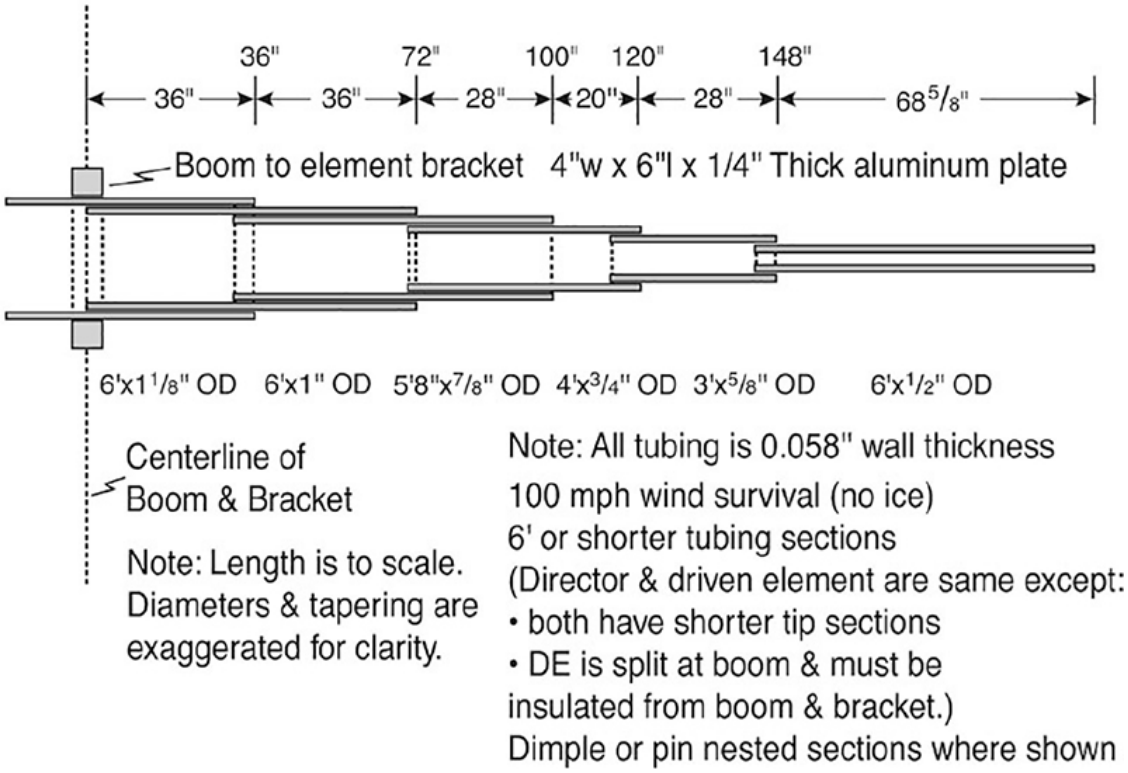


FIGURE 14.5 20-ft Yagi boom with reinforced center section.

After specifying a boom length, there are many possible designs for a three-element Yagi. The designs of Table 14.2 are intended to achieve these goals:

- Free-space forward gain of 8 dBi or greater across the design bandwidth
- Front-to-back ratios of 25 dB or more across the design bandwidth
- Driven element offset slightly from the boom-to-mast bracket location but still well within reach from the tower

top or mast

- No need for boom or element truss supports
- Well-behaved input impedance compatible with simple matching schemes
- Element tubing less than 6 ft long for easy transportability and reduced scrap
- Lightweight low-profile elements with 100 mph wind survival (60 mph with 1/2-in ice)

TABLE 14.2 Element Dimensions and Spacings (in inches) for Three-Element 20-Meter Yagi, Optimized for Three Different Frequency Ranges

	CW	**Mid**	**Phone**
Reflector–DE spacing	100 7/8	100 7/8	100 7/8
DE–director spacing	134 3/8	134 3/8	134 3/8
Reflector tip length	216 5/8	214 3/8	213 1/2
Driven element tip length	210	207 5/8	206 3/4
Director tip length	199 5/8	197 3/4	197

Although it would have been possible to design this Yagi with fewer different diameters of tubing for elements—i.e., with more pronounced steps between nested tubes— the resulting design would have required special techniques for implementing a double diameter step at the junction of certain pairs of tubes. Some commercial antenna manufacturers accomplish this by swaging one end of each tube. Another approach is to insert a few inches of tubing with the right diameter at the junction of the two tubing segments. But if we're going to have to obtain tubing to avoid having to swage joints, we might as well use it for element segments instead.

Aluminum tubing of various diameters—all with 0.058-in wall thickness—is typically used in the construction of HF Yagis so that tubing in increments of 1⁄8-in OD can be nested (telescoped). Determining the length of each section requires an understanding of mechanical design considerations and material strength. Leeson's *Physical Design of Yagi Antennas* does a superb job of relating those matters to the design of HF Yagi antennas and provides specific design examples.

The half-element taper schedule of Fig. 14.4 has good wind and ice survival characteristics because critical regions of the half-element are double thickness by virtue of inserting the next smaller diameter inside the larger tube for at least the distances shown in the figure. If you attempt to scrimp on tubing material and do not overlap tubes at least as much as shown in the drawing, your Yagi will fail at low wind speeds. Telescoping joints and doubled thickness regions should be pinned to prevent slippage between any two tubes. Popular techniques for pinning include dimpling reinforcing tubes with a punch, stainless steel sheet metal screws, and pop rivets. Using hose clamps on telescoping sections is not recommended and requires special care.

An important design decision is whether to shunt-feed the driven element. If so, the element can be continuous metal tubing across its center point as it passes over or under the boom, but the added complexity of a gamma match or other shunt-feeding system may make the design more challenging.

Instead, this beam employs a split driven element, allowing the use of a simple feed technique, the 4:1 balun. The difficulty with this approach is that an 18-ft length of tubing supported at one end is a long cantilevered beam. The boom-to-element bracket for the driven element must be long enough to allow proper spacing of four U-bolts, two on each side of the driven element, along its length and the U-bolts and bracket must be strong enough to withstand the moment arm created by the force of the wind over the entire length of the element half. Insulating tubes must surround the element halves where they are grasped by the U-bolts, and the insulation must extend beyond the outer edges of the bracket.

For the director and reflector brackets, flat aluminum plates at least 1/4-in thick, 3.5 in wide (the edge dimension that is parallel to the boom length) and 6 to 8 in long (the edge dimension parallel to the element length) are adequate for normal wind speeds. Alternatively, aluminum rectangular (bathtub) channel of roughly the same dimensions provides greater strength.

Table 14.2 is used in conjunction with the element taper drawing of Fig. 14.4 to assemble each element of the 20-m Yagi. The table includes three separate sets of dimensions, each having a distinct focus:

- The CW column emphasizes performance across the 14.000 to 14.100 MHz end of the 20-m band. When fed with 50-Ω transmission line, the SWR is less than 1.5:1 across that 100-kHz range and the F/B ratio exceeds 30 dB over the bottom 80 kHz or so.
- The Mid column is a design that maximizes SWR bandwidth while maintaining reasonably flat gain across the 20-m band. Maximum gain at the design frequency is a fraction of a decibel less than in either of the other two designs, and the F/B ratio can drop from a maximum of 30+ dB at 14.200 MHz to 15 to 20 dB at the frequency extremes. The 50-Ω SWR is under 3.0:1 between 14.000 and 14.350 MHz.
- The Phone column is a design that keeps the SWR under 2.0:1 across the 14.150 to 14.350 MHz range that corresponds to the full Amateur Extra license privileges in the United States. Front-to-back ratio is in excess of 30 dB at 14.250 and is above 20 dB throughout the full 200-kHz segment.

The element tapering of Fig. 14.4 is used for this three-element Yagi. Assembly of the six half-elements is identical except for the final setting of the tips. The half-elements in Table 14.2 use identical taper schedules except for the length of the tip (1/2-in OD) sections. The longest tip, the half-element for the "CW" Reflector, is

shown in Fig. 14.4. Where tubes are telescoped, the length of internal overlap should be 3 in or greater. For a detailed discussion of techniques for securing telescoping sections, refer to the earlier section on rotatable dipoles.

For all three designs, the input impedance is close to 75 Ω at the design frequency after passing through a 1:4 step-up balun or other transformer. The SWR bandwidths in systems employing 75-Ω feedlines are wider than the 50-Ω bandwidths. Free-space gain at the design frequency is approximately 8.5 dBi, typically varying ±0.25 dB throughout the useable bandwidth. The designs have been adjusted for mounting 70 ft (approximately one wavelength) above flat ground consisting of average or medium earth.

Multimode DXers might choose the CW design if their voice-mode activity occurs below 14.220 MHz. Adding an adjustable ATU or fixed-component matching network at the equipment end of the transmission line should reduce the SWR at the transmitter or amplifier to under 2:1 across the 20-m band. One setting of the ATU or matching network is sufficient. A coaxial bypass relay or two-position selector can be used to switch the matching network in and out of the circuit for the two halves of the band with the "CW" or "Phone" design. With this approach, the main reason to choose one design over the other is to place the best F/B ratios in either the low or the high end of the 20-m band.

14.4.4 Wire Yagis

Instead of tubing, Yagis can be constructed from wire. A two-element wire beam is shown in Fig. 14.6. The wire beam is made as if it were two parallel half-wavelength dipoles separated by about 0.2 to 0.25 wavelengths. This set of dimensions is for a parasitic director, as the unfed wire is shorter than the driven element.

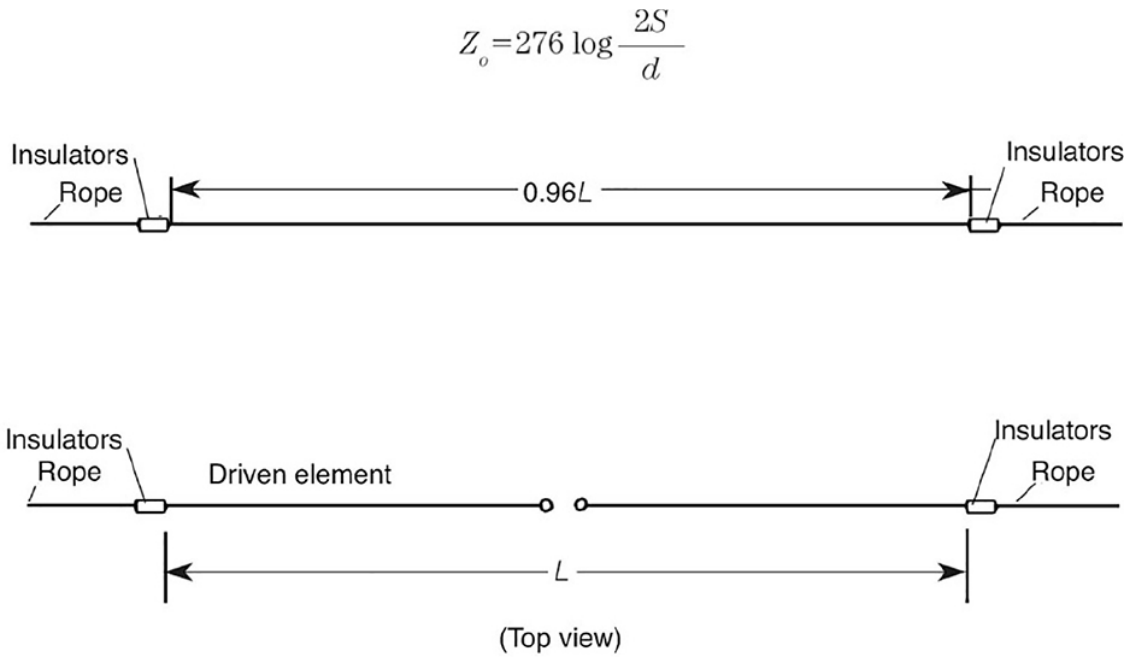


FIGURE 14.6 Wire beam antenna.

Multielement wire beams are also possible. Some leading DXers use wire Yagis of up to seven elements fixed on compass headings. One of the most frequent uses of wire beams is on the lower bands (e.g., 40, 75/80, and 160 m) where rotatable beams are more expensive and difficult to install at useful heights. The small element diameter of wire beams leads to a narrower bandwidth and a slight reduction in the maximum obtainable forward gain compared to Yagis made from larger diameter tubing.

Dimensions for three-element wire Yagis are provided in Table 14.3. #12 uninsulated wire is used and the Yagis are assumed to be about 70 ft above average ground. These beams give up about 0.5 dB of forward gain relative to ones made of aluminum tubing in order to get reasonable F/B and F/S ratios and a good match to either 50-Ω or 75-Ω feed-lines through a 4:1 balun. Since the height chosen is a smaller percentage of a wavelength as the frequency decreases, there is greater ground loss at the lower frequencies and the elevation angle of peak forward gain increases. Using a 4:1 balun to bring the feed-point impedance up to something compatible with commonly available coaxial lines, the 2:1 SWR

bandwidth for a 75-Ω system is about 1.5 percent of the center frequency. The dimensions in this table are incorrect for Yagi elements made of aluminum tubing and may need adjustment for different heights above ground.

TABLE 14.3 Three-Element Copper Wire Yagi Dimensions

Band	**80**	**40**	**20**
Design frequency	3.600 MHz	7.100 MHz	14.125 MHz
Reflector—driven element spacing	$38' \ 3''$	$19' \ 5''$	$9' \ 9''$
Driven element—director spacing	$35' \ 6''$	$18' \ 0''$	$9' \ 1/2''$
Reflector length	$139' \ 4''$	$69' \ 3''$	$34' \ 7''$
Driven element length	$134' \ 5''$	$67' \ 10''$	$34' \ 1/2''$
Director length	$129' \ 9''$	$65' \ 9 \ 1/2''$	$33' \ 1''$
Forward gain at design frequency	9.6 dBi	11.8 dBi	12.7 dBi
F/B ratio at design frequency	20 dB	25 dB	32 dB
F/S ratio at design frequency	20 dB	25 dB	24 dB
Z_{IN} at design frequency	85 Ω	22 Ω	17 Ω
Matching network	$\lambda/4$ *Q*-section	4:1 balun	4:1 balun
2:1 VSWR bandwidth	75 kHz (75 Ω)	100 kHz (75 Ω)	200 kHz (75 Ω)

14.5 Multiband Yagis

Largely because of the costs and complexities associated with erecting multiple beams high in the air, amateurs and professionals have developed a variety of approaches to covering more than one frequency from a single boom:

- Interlaced elements
- Trapped elements
- The log-periodic beam
- Adjustable (motorized) elements

While the shared-boom approach has advantages, the electrical interactions and mechanical issues with multiple elements on a shared boom add to the design complexity and cost and almost always lead to compromises in performance. Antenna modeling software can be used to optimize performance of multiband beam antennas.

Interlaced beams allow separate feedlines for each band. A disadvantage of interlaced beams is that they can be mechanically difficult to construct, erect, and maintain. One popular multiband antenna is the trap tribander (sometimes a five-bander if provision has been made for the 12-m and 17-m WARC bands). Although most tribanders are for amateur use, trapped tribanders designed for key shortwave broadcast bands or long-haul military HF comm links are also available.

Traps are lumped-component parallel-resonant LC circuits inserted partway along both sides of each element, similar to the ones used in trap dipoles and shown schematically in Fig. 10.1. In a 20-15-10 tribander there are two parallel-resonant circuits on each half element. The one closer to the boom is parallel resonant in or near the highest band of operation (10 m). On the next lower band (15 m), the increased reactance of the 10-m capacitor C1 essentially removes it from the circuit, while the 10-m inductor, L1, functions as a loading coil, shortening the element relative to a half wavelength. The second trap in each half-element is parallel

resonant near 15 m. In some tribanders this circuit is housed in the same assembly as the 10-m circuit, but in others it is 1 or 2 ft farther from the boom, in a completely separate enclosure. Either way, it serves to decouple, or disconnect, the outermost part of the element from operation on 15 m. On the lowest band (20 m), both C1 and C2 are essentially out of the circuit because of their high reactances, and the traps are electrically equivalent to two lumped-component inductors.

Because of these lumped-component inductors, each element half of a trapped beam is shorter than its counterpart in a conventional beam for the same frequencies. The inductors may add weight to each half-element compared to the weight of the "missing" aluminum tubing in an equivalent untrapped, longer element.

One disadvantage of a multiband Yagi employing traps is that the spacing between adjacent elements is likely to be less than optimum on at least one of the bands covered by the beam. A good antenna designer can adjust the effective element lengths on each band to counteract the lack of freedom in choosing unique interelement spacings for each band.

Traps cause RF energy loss in the antenna system. The signal strength at a distant point is reduced by the contribution that would have come from each portion of the radiating element that has been replaced by a lumped-component, nonradiating, inductor. Resistive losses in the coils can be quantified if the *Q*-factor of the coils is known. High-*Q* coils will minimize losses, but they may be heavier and more expensive.

While trapped beams are convenient, they may give up between 0.5 and 2.0 dB of gain compared to a trapless monobander of the same number of elements. The additional losses may be insignificant when weighed against the added real estate and dollar costs associated with multiple monobanders and separate supports, or the mechanical complexity and potential for interactions between beams mounted on the same support. Popular tribanders include the Hy-Gain TH-6 and TH-7 family of trapped beams, the Mosley TA-33 and PRO-series beams, and the Cushcraft A-3S and A-4S.

Yagis are available with adjustable elements. Motors at element centers lengthen or shorten the elements, allowing fixed-mounted elements to be optimized over a wide range of frequencies. Two- and three-element Yagis can almost instantly switch the direction of maximum forward gain by 180 degrees simply by commanding the parasitic elements to swap lengths. Since the antenna dimensions do not have to be optimized for more than a single frequency at any one time, element spacing along the boom is relatively unimportant.

14.6 Yagis with Loaded Elements

One popular Yagi is the monoband beam with shortened elements, especially on 30, 40, and 80 m, where beams with full-size elements pose mechanical challenges to the rotor and support system. Elements are shortened with lumped loading coils or linear loading wires. While lumped loading coils may appear very similar to traps, they differ in that they are nonresonant. A loading coil simply replaces part of the radiating element with a lumped component inductor.

Like traps, loading coils have loss. Coils on beams with shortened elements—notably the Cushcraft 40-2CD and the XM240—are not high-*Q* coils. By comparison, the Force 12 Delta series of "shorty 40" loaded beams uses extremely high *Q* coils that are bulky and heavy. Despite the losses, Yagis with loaded elements may still be outstanding performers.

Two 40-m beams with linear loading are the MFJ/Hy-Gain Discoverer series and the

M^2

40M2LL and 40M3LL. The Discoverer DIS-72 2-element beam superseded the venerable 402BA. These antennas have an insulator midway on each shortened half-element. A loading wire runs from one side of the insulator back toward the boom and back out to the other side of the insulator, effectively lengthening the element by folding a portion back on itself. Linear loading wires are inherently low-*Q* components because their inductance is spread out over a large area and they have more loss resistance than high-*Q* coils for

an equivalent shortening of the elements. They may also be subject to a slight shift in inductance with icing.

14.7 Feeding the Yagi

Adding parasitic elements adjusted for maximum forward gain and front-to-back ratio reduces the impedance of the basic half-wave dipole driven element. Often the feedpoint impedance of the antenna between 15 and 30 Ω and too low to be directly fed with coaxial cable. Impedance matching is usually needed. This can be done with a balun, a gamma match or T-match, or a hairpin match.

14.7.1 Balun Feed

A Yagi can be fed with an impedance-matching balun transformer. For most three-element direct-feed Yagis, this requires a step-down configuration, dropping the 50- or 70-Ω impedance of the coaxial cable or hard line to one-fourth of either value, depending on the approximate input impedance of the specific antenna involved. For more on matching, see Chap. 6.

The primary purpose of a balun—and the origin of its name—is to drive balanced feedpoints, such as those found on most horizontal antennas, including dipoles and the driven elements of Yagis, with unbalanced feedlines, especially coaxial cables. The balun should be as close to the feedpoint as possible, such that the wires from the balun terminals to the feedpoint are electrically short. Otherwise, the wire lengths need to be included in the calculations associated with the length of the driven element itself. To minimize the impact of lightning-induced surges on the transmission line, the ground terminal on the unbalanced side of the balun should be directly grounded to the boom of the Yagi at the balun.

14.7.2 Gamma Match

Our earlier discussion of the rotatable dipole and the driven element of multielement Yagi beams was based on a $\lambda/2$

dipole split at the center to allow a direct series connection to a feedline, but that is not the only way to feed the dipole. A popular example of a shunt-feed system attaches the feedline in parallel across a portion of the element that remains electrically continuous across its centerline. This is the gamma match shown in Fig. 14.7.

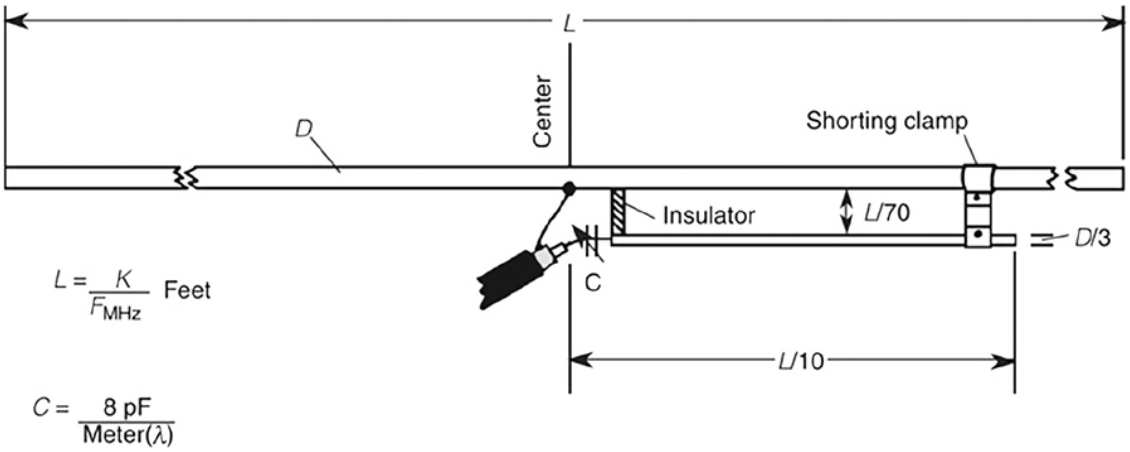


FIGURE 14.7 The gamma match.

The operating principle is that we are connecting the transmission line at a point (or tap) on the element where the impedance of the transmission line is comparable to the impedance of the standing wave of RF energy on the driven element (or dipole). We can best understand this by visualizing each half-element as a tappable resistor that is

$\lambda/4$

in length. The gamma match transforms the impedance at some tap point along the element half to a resistive part that matches the line impedance and cancels any leftover antenna reactance.

Unlike the series fed dipole, the driven element of the gamma-matched Yagi is not broken in the center. The outer conductor or shield of the coaxial cable is connected to the center point of the driven element, which is often directly grounded to the boom. The center conductor of the feedline is connected to one end of a capacitor that is in series with a component known as the gamma rod. The gamma rod is like a link coupling to a zero-turn coil (the element). Early designs used a moveable link to vary the coupling

between the link and the primary resonating inductance in antenna tuners and plate tank circuits. Later antenna tuners used a fixed link but varied a capacitor in series with the link to ground, a configuration quite analogous to the gamma configuration of Fig. 14.6. Moving the shorting bar is equivalent to changing the number of turns on the link.

Here is one set of dimensions for gamma matching a Yagi to 50-Ω line:

- Distance from centerline of element to gamma rod shorting bracket: *L*/10
- Gamma rod diameter: *D*/3
- Spacing of gamma rod from driven element: *L*/70

where *L* is the length of entire driven element, roughly equal to $\lambda/2$, and *D* is the driven element diameter. When constructing the matching system, it can be helpful to make the gamma rod itself substantially longer than *L*/10, since these dimensions are only approximate, and both the position of the shorting bar and the setting of the capacitor are then varied by trial and error.

The capacitor in series with the center conductor of the coaxial cable has a value of approximately 8 pF per meter of wavelength at the lowest frequency in the band of interest or approximately

$$C(\mathrm{pF}) = \frac{2400}{f(\mathrm{MHz})} \quad (14.1)$$

For transmitter power levels of more than a few watts, the capacitor must be a high-voltage transmitting variable type. Gamma match capacitors are usually air or vacuum variables. Those found on older Cushcraft four-element monobanders for 20, 15, and 10 m consist of a fixed length of RG-8 coaxial cable with the outer jacket and shield braid removed so that only the center conductor and dielectric remain. The insulated wire is inserted into a portion of the hollow aluminum gamma rod, where the center conductor of the

stripped coaxial cable forms one plate of the capacitor and the hollow gamma tube is the other.

Tuning the gamma match is accomplished in two steps:

- Adjust the length of the driven element for resonance using an antenna analyzer or other means.
- Alternately adjust the capacitor and the shorting bar/clamp for best impedance match between the antenna and the transmission line.

Results of this tuning operation are monitored at the antenna with a portable analyzer capable of measuring and displaying SWR or impedance curves over the frequency range of interest (see Chap. 30).

If a good match (as indicated by very low SWR) cannot be obtained with the initial spacing of the gamma rod from the element, increase the distance of the gamma rod from the element and go through the adjustment process again. If the SWR obtained this time is lower, continue to increase the gamma rod spacing until no further improvement is possible. If the SWR increased, reduce the spacing between the rod and the element and perform the adjustment procedure again. In some Yagi designs, it may also be necessary to shorten the driven element to get a match.

14.7.3 T-Match

The T-match is a balanced (or two-sided) gamma match. Design considerations are identical. Some antenna experimenters prefer the T-match because it presumably provides a more symmetrical input drive to the balanced driven element. In addition to requiring twice as many components, pattern skewing resulting from unbalanced drive to a dipole or driven element is hardly noticeable in practical usage. Whatever advantage may accrue to minimizing pattern skew is probably negated by the loss in the lumped-component balun used to convert the unbalanced feed from coaxial cable to the balanced set of gamma circuits.

14.7.4 Hairpin Match

In some multielement Yagi designs, the feedpoint impedance at the terminals of the driven element lends itself to matching via a stub, which is a short piece of transmission line attached to the feedpoint terminals of the driven element and short-circuited at the far end. The driven element is split and insulated from ground on both sides and the feedline attached somewhere along the stub between the feedpoint terminals and the short-circuit. Figure 14.8 is a sketch of a typical hairpin matching system. The point of attachment is one of the adjustments made when tuning the system for minimum SWR. Use of a stub is appropriate when the driven element impedance has a negative reactive (capacitive) component.

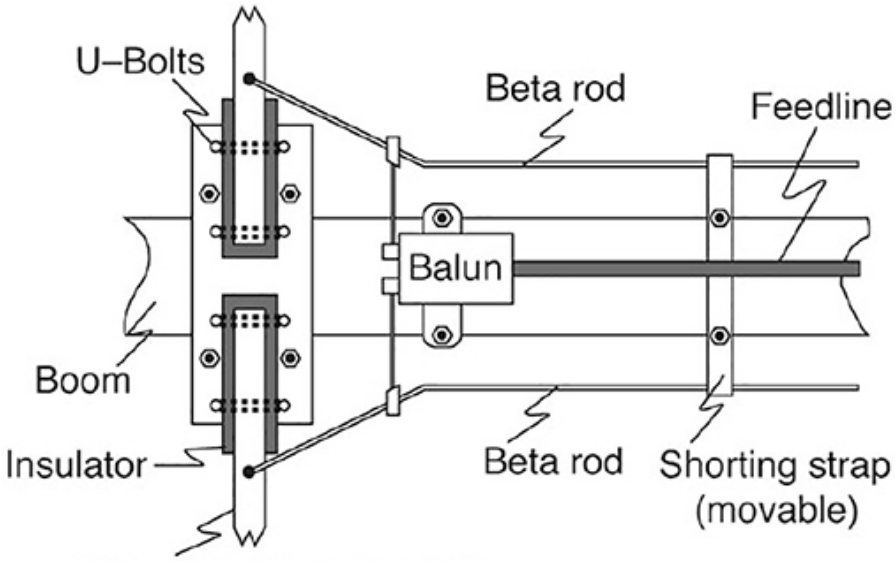


FIGURE 14.8 Beta match example.

The hairpin has appeared under the trade name "Beta Match" on a number of MFJ/Hy-Gain beams over the years, including their line of *Long John* monobanders and the TH-6DX and DXX trap triband beams. On a commercially produced beam such as the Hy-Gain, the only adjustment necessary is to set the element lengths, the position of the feedline connection, and the position of the shorting

bar to the dimensions shown in the instruction sheet for the desired center frequency.

Tuning the T-match is a two-step process similar to that of the gamma match:

- Adjust the length of the driven element (or dipole) for resonance, using an antenna analyzer or other means. Shorten each side of the element to make it exhibit some capacitive reactance. A reasonable starting point for HF hairpins is an initial reduction of approximately 1 in for every 3 m of operating frequency. Example: For a 15-m driven element, shorten each side by 4 or 5 in.
- Alternately, adjust the distance of the feedline tap from the driven element and the shorting bar/clamp for best impedance match between the antenna and the transmission line. If you cannot obtain a good match, it may be necessary to go back to the first step and try another length for each side of the driven element.

The M^2 40M2LLA two-element linear loaded monobander for 40 m uses a hairpin with the transmission line feeding a balun whose secondary leads connect directly to the driven element feedpoint terminals, so one degree of freedom (position of feedline tap along the shorted stub) is lost. The manufacturer's instructions effectively combine both of the preceding steps by having the user alternately adjust the position of the shorting bar and the lengths of the driven element tips in 1-in increments.

When designing a hairpin match, the Smith chart (Sec. 6.10) can be a great help. With a measurement of the input impedance of the driven element, the stub-matching procedure provides the hairpin dimensions and the point at which the feedline should be attached.

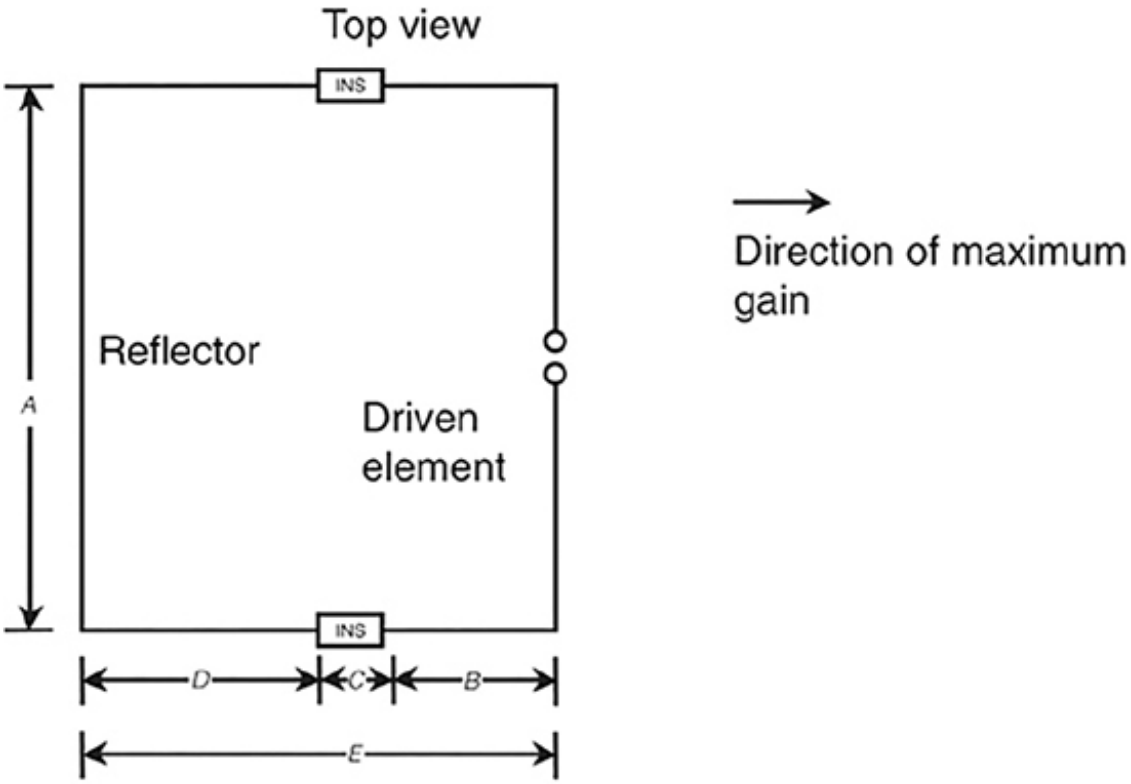
14.8 Moxon Beam

The field radiated by a dipole results mainly from the high-current portions of the antenna. The standing wave of current in the outermost $\lambda/8$-section of each end of the dipole contributes little. While the outer ends provide little radiation, they make the dipole resonant and minimize the reactive part of the input impedance. The low-current ends also allow dipoles to be connected to form loops and other phased arrays of dipoles exploiting the principle that the standing wave of current reverses direction in each half-wave wire segment. With 80- and 160-m dipoles, to reduce the overall antenna size, the outermost $\lambda/8$-section of each side can drop toward the ground. The same principle can be used to create two- and three-element Yagis for limited spaces by constructing the center half of each element out of rigid aluminum tubing and then dropping weighted lengths of #10 or #12 wire from the ends of each.

The two-element Moxon beam (Fig. 14.9), named for L. A. Moxon, is similar to an NVIS horizontal loop where the drooping ends of the two dipoles are pulled up to lie in the same plane as the centers of the elements. Unlike a loop antenna, the dipole ends are connected together mechanically but not electrically. Rather than being square, the outline of the Moxon beam as viewed from above is rectangular because the centers of the two bent dipoles are spaced less than

$$\lambda/4$$

to optimize the antenna's performance as a parasitic array having one driven element and one reflector. The bent ends of the two elements must be less than $\lambda/8$ long and the center section of each element is perhaps $5\lambda/16$ or more in length.



To cover the entire 20-m band with a VSWR less than 1.6:1 into 50–Ω feedline:

A: 296.0"
B: 46.4"
C: 6.0"
D: 56.6"
E: 109.0"

FIGURE 14.9 Moxon beam.

Like all other two-element Yagis employing a reflector, forward gain, F/B, and SWR change more rapidly with frequency when approaching the design center of the beam from lower frequencies. Dimensions and matching circuits should be optimized for the lower end of the band, since any of these performance characteristics will usually be adequate at higher frequencies. A Moxon beam constructed of tubing will have a wider bandwidth than one using wire.

Some implementations of the Moxon utilize wires with spreaders, some use aluminum tubing for the element centers (the portions at right angles to the boom) and wires for the element ends, and others use aluminum tubing for the entire radiating structure. A properly designed Moxon also boasts higher F/B ratio than the typical two-element Yagi.

The dimensions shown in Fig. 14.9 for a 20-m Moxon rectangle constructed with #12 wire yield a free-space forward gain of 6.1 dBi in conjunction with F/B and F/S ratios of nearly 24 dB at the design frequency of 14.1 MHz. The input impedance is

$51 + j0.4\ \Omega$

at the design frequency and the SWR stays below 1.6:1 across the entire band. If the segments of the 20-m design in the figure are scaled according to wavelength, the design translates to other bands while retaining the 50-Ω resistive input impedance.

14.9 Stacking Yagi Antennas

To increase gain and pattern flexibility by connecting two antennas for the same frequency or allow for multiple bands with one support, antennas can be stacked on a single mast or tower.

14.9.1 Stacking for Gain

As we saw in Chap. 7, splitting the available transmitter power between two identical antennas can lead to increased signal strength in certain directions far from the antennas, at the expense of signal in other directions. Optimum HF or VHF stack spacings for a pair of identical beams mounted one above the other are around one wavelength. Stacking two 3-element Yagis can provide the same peak forward gain as a single 6- or 7-element Yagi on a much longer boom-.

Improvement from stacking will be impaired if the upper and lower beams are too close together or if the lower beam is too close to ground. Stacking can fail if the lower beam is lower than

$\lambda/2$

above ground. The higher the frequency, the more flexibility is possible on a support of a given height. A 70-ft tower is needed for a two-Yagi, 20-m stack, while a 10-m stack can be mounted at different heights and spacings on the same tower.

To rotate the beam in azimuth, a rotating tower can be used. Coverage of a full 360 degrees may be unnecessary. The lower beam in the stack can attain as much as 300 degrees rotation by mounting its rotator on a short outrigger mast supported 1 or 2 ft from the nearest tower leg at both ends of the mast. Some rotator controllers can be interconnected in a master/slave configuration so that the lower beam follows the commands for the upper through the more limited range. Or, one or more lower beams can be fixed in position by being mounted directly to sides of the tower at the appropriate height(s). A DXer in the northeastern United States might attach beams aimed at 45 degrees (to maximize coverage into Europe) and 135 degrees (for South America and the southern half of Africa) directly to the tower at points roughly a wavelength below the top beam.

Stack spacings on a guyed tower will be determined as much by guy wire attachment heights and element clearance distances as by theoretical projections of stacking gain versus spacing. Exact stack spacing is usually not critical.

Because the stacking gain curve exhibits a broad peak as a function of spacing, multi-band beams covering a 2:1 frequency span can be spaced as much as 3

$\lambda/4$

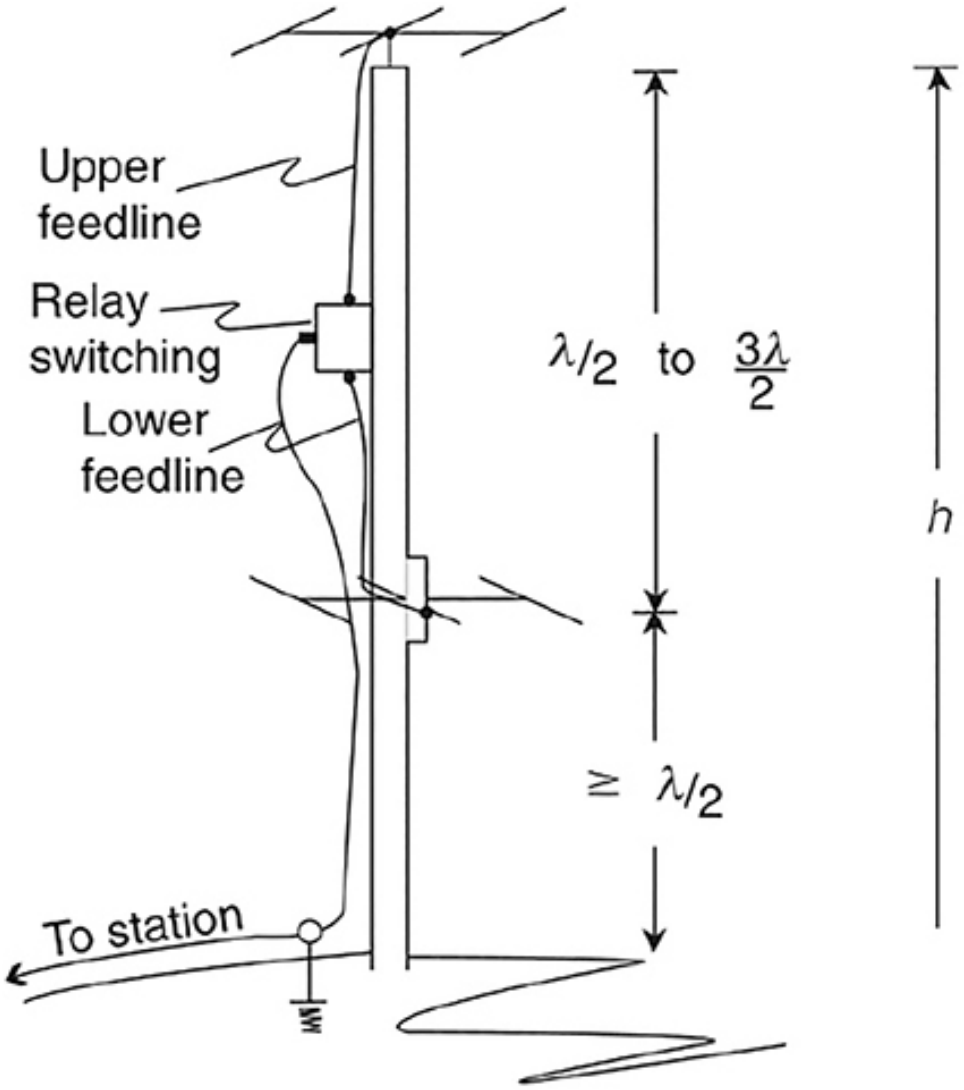
apart at the lowest frequency of operation, corresponding to a spacing of

$3\lambda/2$

at the highest frequency. HF tribanders or five-banders (when the 12- and 17-m WARC bands are included) covering a 2:1 frequency span can be stacked with good results.

The feedline configuration to each beam in the stack is normally identical in length and impedance—a “T” connection, in other words, as shown in Fig. 14.10. This provides roughly equal

transmitter power in phase to the two Yagis. A useful adjunct to the two-beam stack is a coaxial cable relay control box with a four-position switch and associated transmission line phasing sections that allow the user to select beams in phase (BIP), beams (180 degrees) out of phase (BOP), upper beam only, or lower beam only. Each of these choices is useful at one time or another not only because we are usually interested in communicating over a wide range of distances but also because the optimum elevation angle for HF ionospheric communications is always varying.



Upper & lower feedlines should be exactly the same length, measured from antenna feedpoint to switch box.

FIGURE 14.10 Feeding stacked Yagi antennas.

One advantage to using one or more fixed lower beams or having the upper and lower beams independently rotatable is that transmitter power and receiving capability can be split between two different directions. The upper beam might be rotated in the direction of one continent while the lower beam is aimed at another.

Having two antennas available at two different heights, selected singly or in combination, ensures a reasonable pattern gain at virtually any elevation angle useful for communicating over ionospheric "skip" distances. This, rather than any specific stacking gain, may be the real advantage of having higher and lower beams for the same band at HF.

A final consideration is maintainability. The added benefit of providing full 360-degree rotation for the lower beam in a stack compared to, say, 300-degree coverage with one or more fixed or side-mounted beams may not be worth the extra cost and complexity. The added gain of a four- or five-element Yagi compared to a simpler three-element beam may be lost if there is a failure in the driven element or matching circuit.

14.9.2 Stacking for Limited Space

Stacking is also used when unrelated antennas for different frequency ranges or bands are mounted on the same mast or support structure. The hope is that unrelated beams can be mounted near each other without adversely affecting performance. This can be accomplished with modeling software, experimentation, or manufacture recommendations for minimum stacking distance, although recommended minimum stacking distances can be ridiculously large for practical HF installations.

Some general guidelines can be given. Lower-frequency beams usually have a greater effect on higher-frequency beams than the reverse. Avoid stacking antennas with full-size or trapped elements for harmonically related frequencies (such as 10 and 20 m), particularly third harmonic combinations such as 40 m and 15 m. Combine 15- and 20-m beams, or 10 and 15, or 12 and 17. This prohibition is often not necessary for a beam with loading coils because a lumped inductance will make the element nonresonant on harmonics.

If it is necessary to stack two beams that interact, consider rotating one of the beams 90 degrees with respect to the other. The boom of one beam may be resonant in the frequency range of the

other. A current breaker or detuning element can help. Choose the overall height of one or both beams, if possible, to create a ground reflection null in the elevation pattern(s) at 90-degree elevation (straight up). Put the lower-frequency beam above the higher-frequency one whenever the tower and mast strength allow it. Lastly, consider using a multiband beam (having traps or interlaced elements on a shared boom) in place of separate beams.

14.10 Custom Yagi Designs and Further Reading

There are many excellent resources for designing Yagi antennas, including modeling software. For electrical design and performance, see *Yagi Antenna Design*, by James L. Lawson, ex-W2PV, ARRL, 1986. *Low-Band DXing*, by John Devoldere, ON4UN, ARRL, fifth edition, 2011, treats Yagis at 80 and 40 m as well as hairpin matching.

The designs and construction techniques of this chapter do not take into account severe environments. High winds, icing, and saltwater impose extra burdens on antennas and their supports. A resource on strengthening beams is *Physical Design of Yagi Antennas*, by David B. Leeson, W6NL (ex-W6QHS), ARRL.

CHAPTER 15
Cubical Quads and Delta Loops

Two popular multielement antennas employ wire loops having a total length of approximately one wavelength. The cubical quad employs square loops and the delta loop is built with triangular loops. Both antennas are found in single-element and multielement varieties, and both can be analyzed in terms of the equivalent half-wave dipoles that make them up. The single-element version of each is a form of 1λ loop that was considered in Chap. 9.

15.1 Cubical Quad

The cubical quad was designed in the mid-1940s at radio station HCJB in Quito, Ecuador. According to the story, HCJB originally used Yagi antennas to provide consistent signal strength in selected directions. In the thin air of Quito, the high voltages at the ends of the elements caused arcing that resulted in voltage standing wave ratio (VSWR) changes capable of shutting down or damaging the high-power transmitters and, over time, destroying the element tips. Station engineer Clarence Moore designed the cubical quad antenna to solve this problem.

Figure 15.1 shows the dimensions and construction techniques for a typical quad loop. Because it is a full-wavelength antenna, with each side a quarter-wavelength long and fed at a current loop in the center of one side, the voltage maxima occur

$\lambda/4$

from the feedpoint in either direction. In contrast to the open-circuited element tips of a dipole or Yagi, at the voltage maxima, the wire loop is continuous, which reduces the tendency to arc.

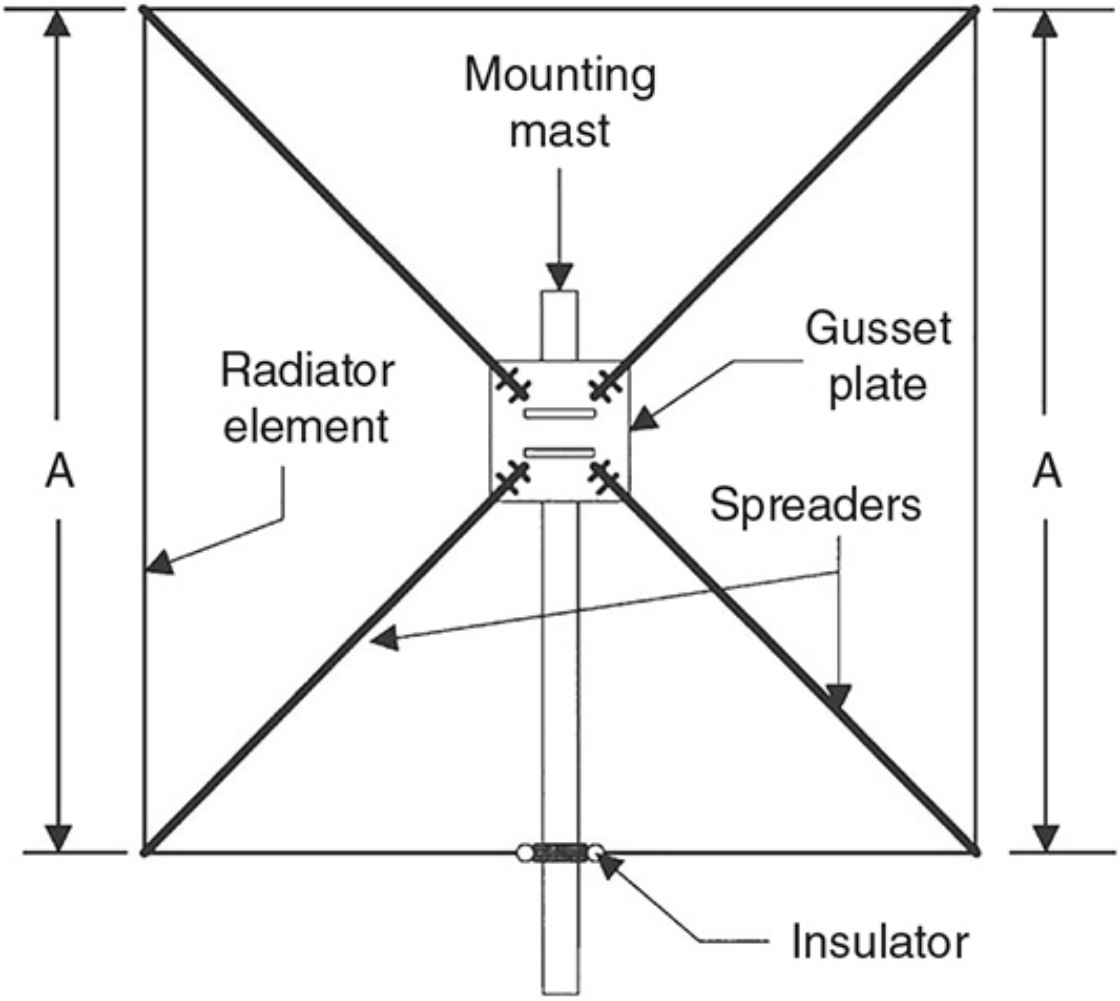


FIGURE 15.1 Quad loop antenna.

In addition to the reduction in corona discharge, the quad loop antenna requires a smaller footprint for a given frequency or band because it is only a quarter-wavelength long across the top or bottom (dimension A in Fig. 15.1). The loop may also be less susceptible to local electromagnetic interference (EMI) and precipitation static.

The quad loop of Fig. 15.1 can be used by itself or as a building block for a multi-element cubical quad. When fed at the center of a horizontal side, the single-element loop will exhibit a horizontally polarized figure eight azimuthal radiation pattern similar to that of a dipole, but the elevation pattern will be squashed in amplitude at

high angles. The quad loop can be thought of as an array antenna. The loop consists of two bent dipoles whose centers are spaced $\lambda/4$ apart. Compared to a $\lambda/2$ dipole having the same orientation as the fed leg of the loop, the quad loop exhibits about 1 dB greater broadside gain in free space. Factors keeping the quad from realizing full stacking gain of a two-element array are the closeness of the two bent dipoles and mutual coupling effects between the two.

Just as a Yagi is formed from a $\lambda/2$ driven element in conjunction with reflectors and directors, a multielement cubical quad consists of a driven quad loop and one or more parasitic loops whose actual or electrical lengths have been adjusted slightly from a nominal $\lambda/2$ so that each acts as a reflector or a director. Two-, three-, and four-element cubical quads are common.

15.1.1 Quad Versus Yagi

In the years since Clarence Moore's article there has been a running controversy regarding the performance of quads as compared to other beam antennas, particularly the Yagi. Some users claim that the two-element cubical quad has a forward gain of about 1.5 to 2 dB more than a two-element Yagi with a comparable boom length between the two elements. The quad may also have a lower angle of radiation and work better at low heights above the earth's surface.

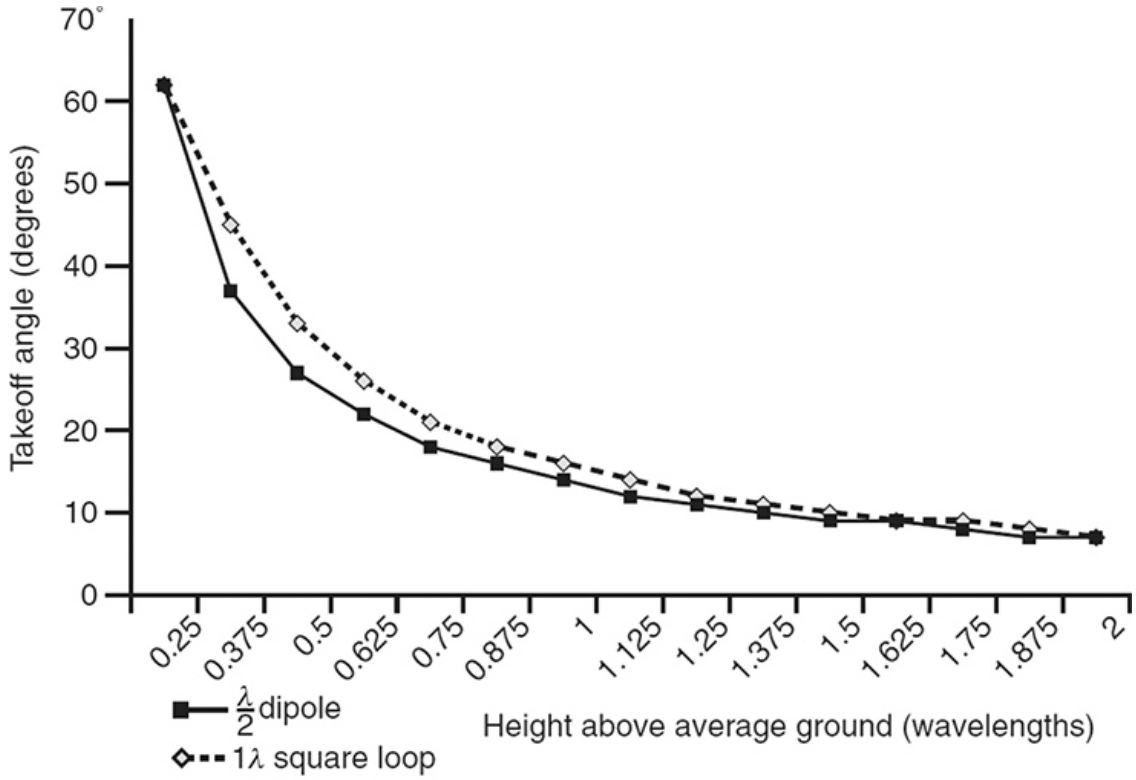
One difficulty is defining the height of a quad. When comparing quads to Yagis, do we put the height of the top wire on the quad loop at the same height as the Yagi or set the equivalent boom heights (i.e., the center of each array in the vertical plane) to be the same? Another difficulty is deciding what performance measures are most important. Forward gain? Front-to-back (F/B) or

front-to-side (F/S) ratio? VSWR bandwidth? Gain bandwidth? Total weight? Wind load? Boom length? Ease of installation? Further complicating the comparison, with both Yagis and quads, maximum forward gain, F/B ratio, F/S ratio, and bandwidth generally occur at different frequencies and design dimensions.

Modeling software can be used to compare the two antenna types. Results using EZNEC 5+ are shown in Fig. 15.2 for a quad loop, Yagi, and dipole at the 15 m band. Compared to a conventional horizontal

$\lambda/2$

dipole at the same height above average ground as the top (horizontal) leg of a loop fed at the center of either the top or bottom leg, the loop exhibits lower broadside gain until a height above ground of about 0.7 λ. Above that height the loop has higher gain. The differences throughout that range are on the order of 0.5 to 1 dB. The dipole has a lower elevation angle or angle of peak broadside radiation. The higher elevation angle of the quad is caused by radiation from the lower bent dipole.


70°
60
50
40
30
20
10
0
Takeoff angle (degrees)
0.25
0.375
0.5
0.625
0.75
0.875
1
1.125
1.25
1.375
1.5
1.625
1.75
1.875
2
λ/2 dipole
1λ square loop
Height above average ground (wavelengths)

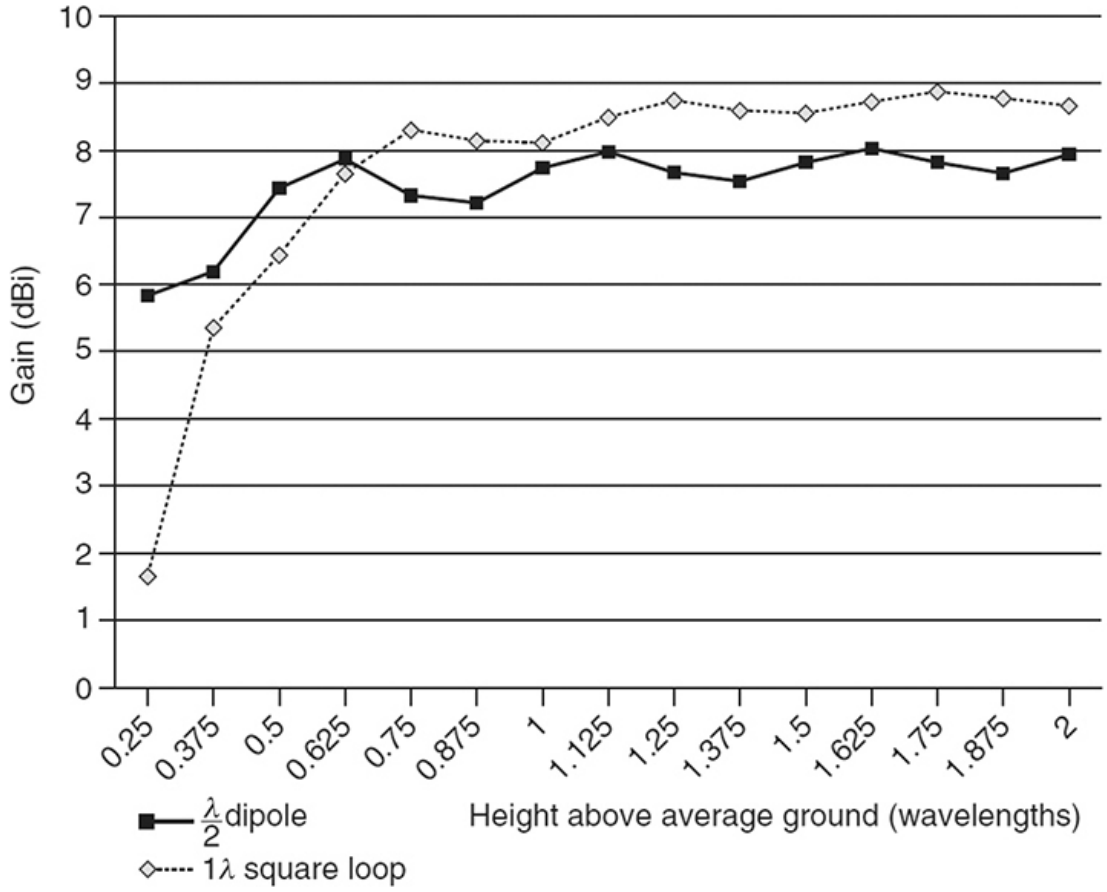

10
9
8
7
6
5
4
3
2
1
0
Gain (dBi)
0.25
0.375
0.5
0.625
0.75
0.875
1
1.125
1.25
1.375
1.5
1.625
1.75
1.875
2
λ/2 dipole
1λ square loop
Height above average ground (wavelengths)

FIGURE 15.2 Modeled gain (bottom) and takeoff angle (top) of 1 λ quad loop versus half-wave dipole over average earth.

The two-element quad enjoys about 1.2 dB forward gain advantage over a two-element Yagi at the same design frequency (in this case, 21.2 MHz) when both are modeled with #12 wire for their elements. But a rotatable Yagi will have elements of aluminum tubing instead of #12 wire supported by insulated spreader arms, which improves the bandwidth of the Yagi for a given boom length. The two-element wire quad shows a greater variation in input impedance (hence, VSWR bandwidth) and F/B or F/S ratios with frequency than does the two-element Yagi constructed of either copper wire or aluminum tubing.

With the two-element quad, it is easier to maintain F/B and F/S ratios in excess of 10 dB. These ratios show much greater variation across the band than do those of the Yagi. It is somewhat easier to adjust the Yagi driven element by lengthening or shortening the element. Similar adjustments to the quad driven element affect all four legs unless the supporting frame is distorted.

Proximity to the earth can hurt the performance of the quad. With the top leg of each element at 1λ (e.g., 46 ft on 15 m), the two-element quad exhibits an advantage over the Yagi at the same height over average ground. As both antennas are brought closer to ground, the quad begins to lose its gain advantage, and by a height of

$\lambda/2$

, the Yagi actually enjoys a slight advantage in forward gain.

Like the basic 1λ loop, one advantage of the cubical quad is a smaller footprint. The span of each quad element is roughly half that of each corresponding element of a full-size Yagi, so the turning radius of a two- or three-element quad will almost always be somewhat smaller than that of a Yagi of the same number of elements. The advantage is somewhat less when the quad is

compared with a Yagi employing traps, loading coils, or linear loading wires to shorten its elements.

15.1.2 Feeding the Quad

The driven element of a cubical quad can be fed in the center of a horizontal side (Figs. 15.1A and 15.3A), in the center of a vertical side (Fig. 15.3B), or at a corner (Fig. 15.3C). In all cases, the currents on opposing sides of the quad geometry are in phase, so maximum gain is along the axis of the loop. For all the quads in Fig. 15.3, maximum gain is into and out of the page.

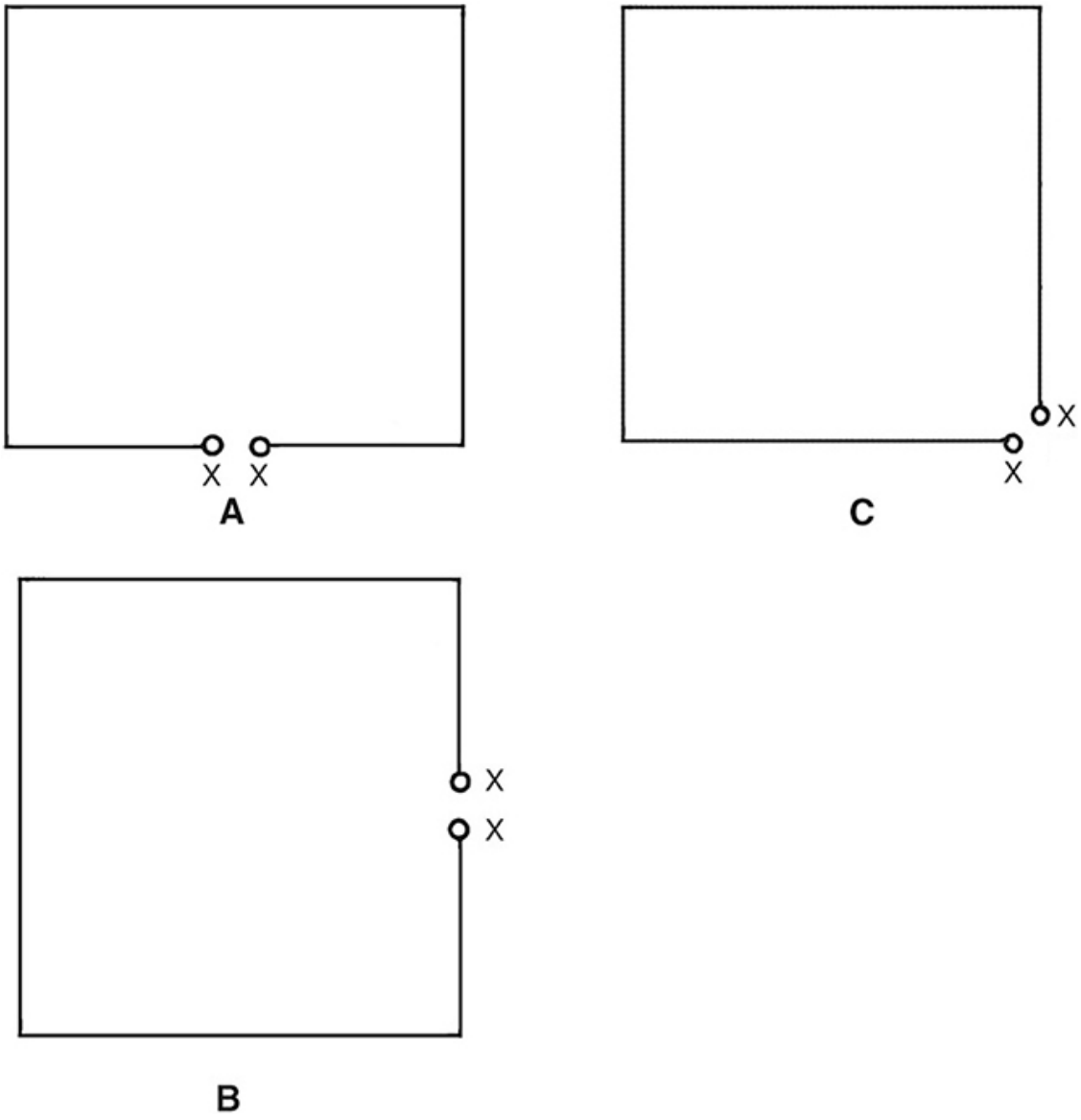


FIGURE 15.3 Feed options for the quad.

A loop located within a few wavelengths of ground or another reflecting surface can be viewed as a four-element array consisting of two bent dipoles and the two image antennas below the reflecting surface. When the bottom of the loop can be lifted to a height of at least a half-wavelength above the earth or other reflecting surface, it begins to make sense to use horizontal polarization by feeding it at the midpoint of the top or bottom side. In this manner, the ground reflection can provide up to 6 dB

additional gain. For most amateurs, this suggests horizontal polarization at 14 MHz and above. At support heights below $\lambda/2$, the ground reflection factor for horizontally polarized antennas tends to put most of the radiation at high elevation angles, so if long-distance communication is the user's objective there is merit to switching to vertical polarization by feeding the loop on one side or the other. As a compromise, the quad loop can be fed at a corner to obtain equal parts horizontal and vertical polarization.

There is little difference in the pattern of a quad loop when feeding the bottom leg instead of the top. The choice should be determined by the specifics of your installation. Feeding at the bottom usually results in a shorter feedline.

Boundary conditions for the two bent dipoles that form a loop element limit its operation to the band for which it is designed. Favorable broadside gain of two bent dipoles having their ends connected requires the currents in opposing legs of the loop to be in phase. Multiband cubical quad antennas are typically by suspending wire loops for higher-frequency bands inside the loop for the lowest frequency, thus giving rise to a "spiderweb" appearance. All bands share a common boom but the parasitic elements of one band may not be coplanar with those of any other band. Typically the driven element for each band is independently fed; separate feedlines can be run all the way back to the transmitter room or a remote mast-mounted coaxial switch capable of handling the expected transmit power can be employed at considerable savings in cable costs.

15.1.3 Construction Ideas

In its simplest form, a single-element quad loop antenna is mounted to spreaders connected to a square gusset plate. Quads are often built with standoff insulators on aluminum supports or with fiberglass spreaders.

Details for a one-element gusset plate are shown in Fig. 15.4. The gusset plate can be made of aluminum like a Yagi boom-mast

clamp or it can be a strong insulating material such as fiberglass or 3⁄4-in marine-grade plywood. The plate is mounted to the support mast using two or three large U-bolts (stainless steel or galvanized, to prevent corrosion). Spreaders are mounted to the gusset plate using smaller U-bolts.

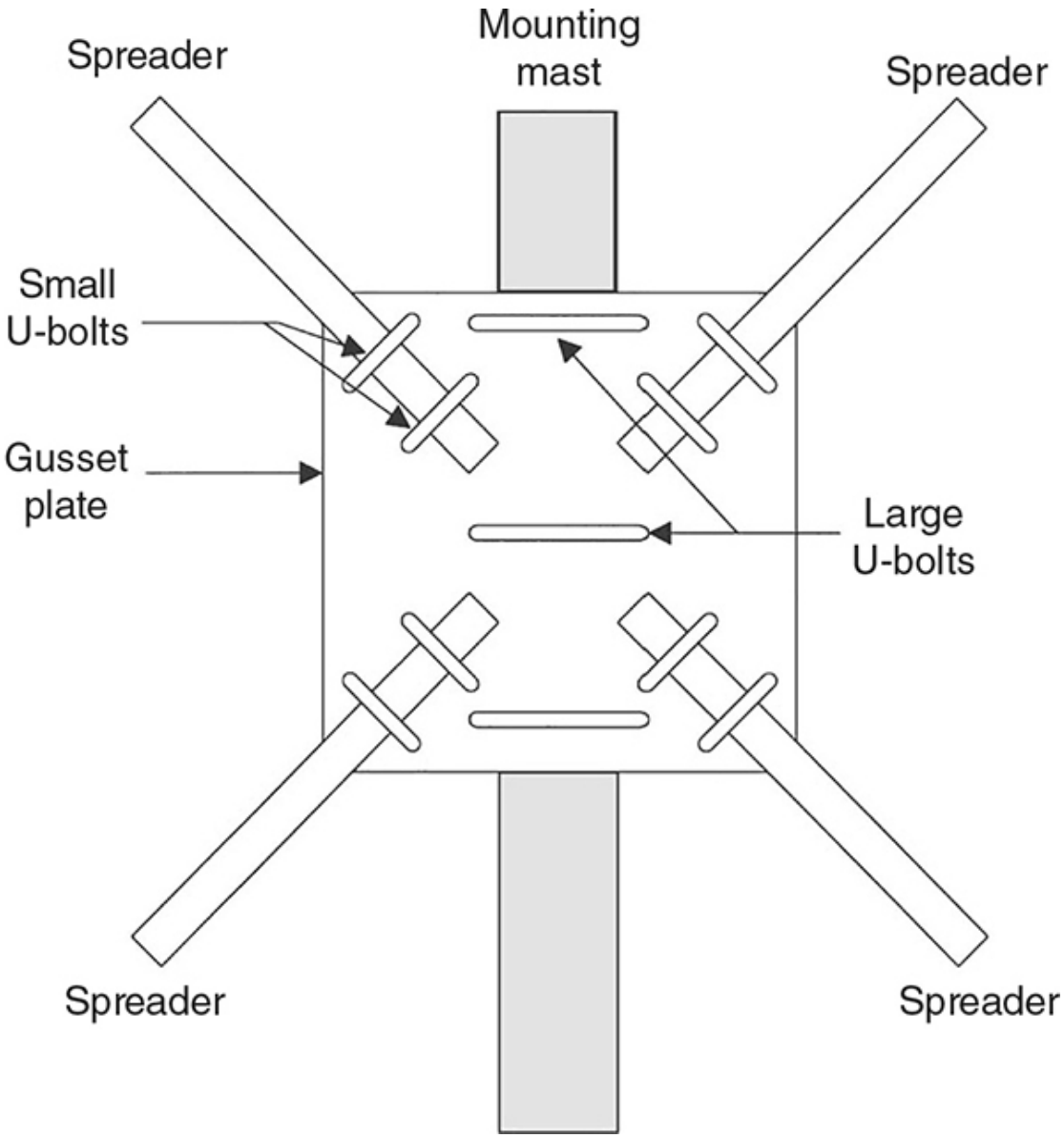


FIGURE 15.4 Quad loop gusset plate.

Element spacing in a multiband quad poses a design problem analogous to that found in multiband trap Yagis. Element spacings

for a 10-m quad should be roughly half the equivalent spacings for a 20-m quad. Clever mechanical design, exemplified by the boom-spreader device—known as a spider—of Fig. 15.5, has solved this problem for the two- and three-element multiband quads.



FIGURE 15.5 Multiband quad "spider."

A four-element quad is a bit more complex. Some four-element quad designs employ the gadget of Fig. 15.5 in conjunction with

three separate single-loop boom-spreader brackets spaced along the far end of the boom to support and properly position a fourth element (usually a second director) on each band. Alternatively, a "half-spider" that secures the spreaders for all second directors at some angle other than 45 degrees might work, depending on the spacings called for within a given design.

Figure 15.6 shows a schematic representation of a horizontally polarized two-element quad with elements spaced 0.11 to 0.22 wavelengths apart. The driven element is connected to the coaxial-cable feedline directly or through a

$\lambda/4$

matching section of 75-Ω line. The other element is a reflector, so the total wire length of the loop is a few percent longer than that of the driven element. A tuning stub can be used to adjust the reflector loop for maximum gain or F/B ratio.

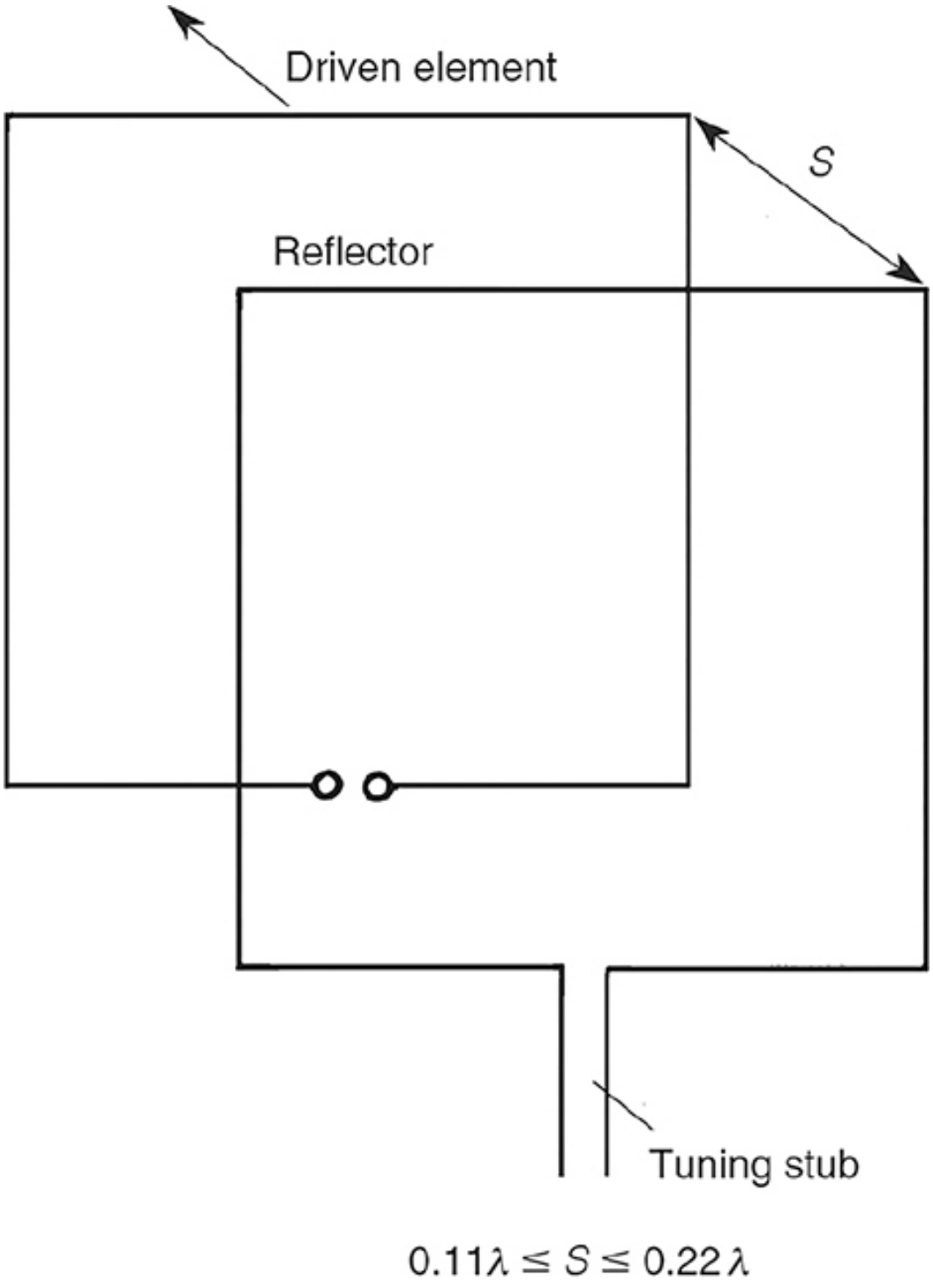


FIGURE 15.6 Quad beam antenna.

Because each element is a square loop, the actual length varies from the naturally resonant length by about 4 percent. The forward gain and F/B ratios are more sensitive to the reflector dimensions

than they are to the driven element length. The overall lengths of the wire elements are given by the following formulas:

$$\text{Driven element}: \qquad L\ (\text{ft}) = \frac{1005}{f(\text{MHz})} \quad (15.1)$$

$$\text{Reflector}: \qquad L\ (\text{ft}) = \frac{1030}{f(\text{MHz})} \quad (15.2)$$

$$\text{Director}: \qquad L\ (\text{ft}) = \frac{975}{f(\text{MHz})} \quad (15.3)$$

One method for constructing a single-band, two-element quad is shown in Fig. 15.7. This particular scheme uses a $12\text{-} \times 12\text{-in}$ wooden plate at the center, fiberglass spreaders, and a wooden or metal boom. The construction must be heavy-duty in order to survive wind loads. A quad kit consisting of spreaders and spider for either a two- or a three-element quad could be used instead.

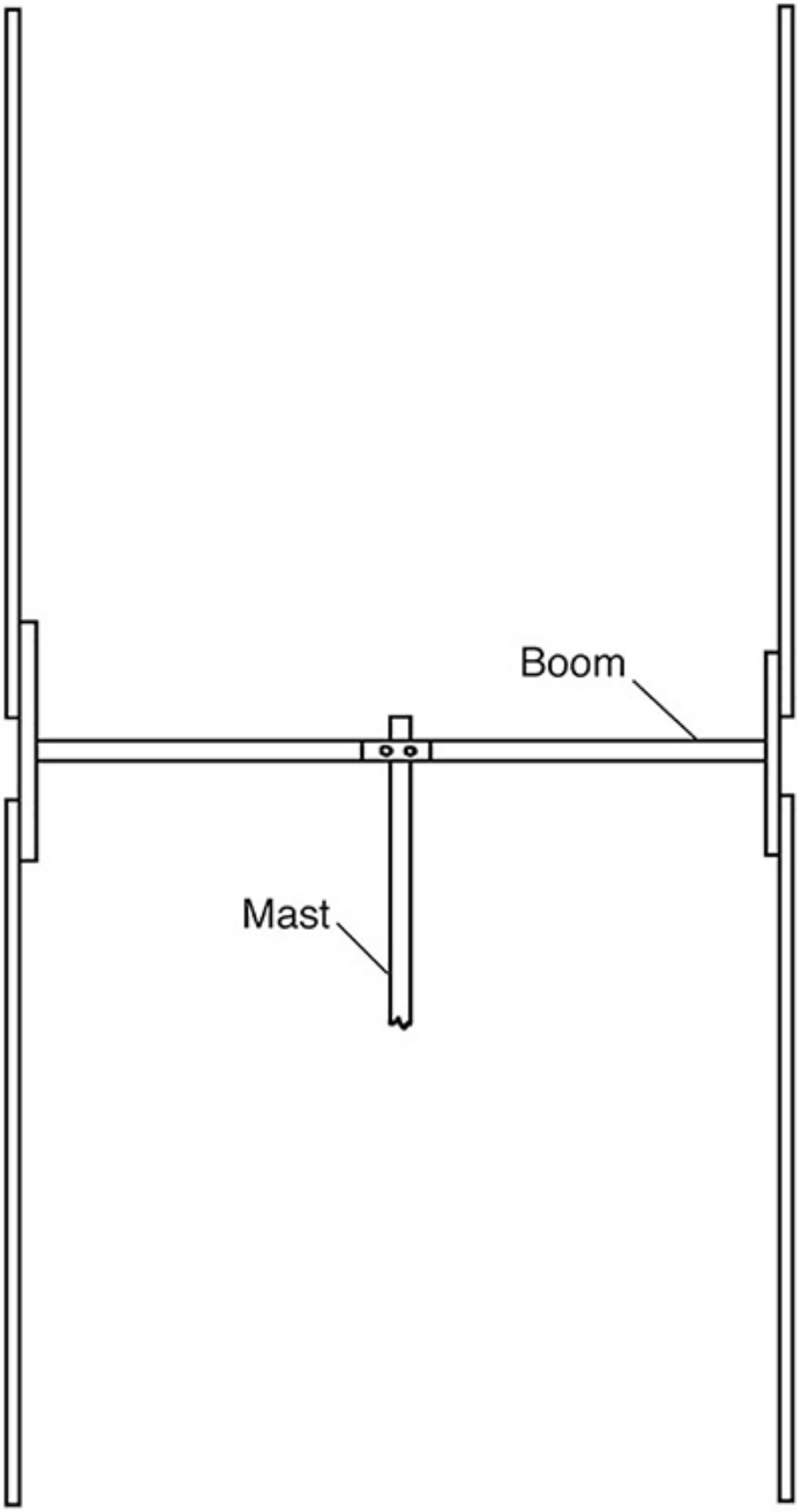

Boom
Mast

FIGURE 15.7 Single-band two-element quad on a boom.

With most kits, more than one band can be installed on a single set of spreaders. The size of the spreaders is set by the lowest band of operation, so higher frequency bands can be accommodated with shorter loops on the same set of spreaders. While the use of a spider allows element spacing proportional to wavelength on multiple bands, there is no shortage of two-, three-, four-, and five-element quads that forsake proportional spacing and simply nest concentric loops for three, four, and five bands on vertical spreaders. The construction technique of Fig. 15.7 can be used to support loops for multiple bands as long as it is understood that VSWR bandwidth and other key performance parameters such as forward gain, F/B and F/S ratios across the operating band will be different.

15.2 Delta Loop

At frequencies below 14 MHz, spreaders and supports for fixed or rotatable multielement quads are challenging to build. The fixed delta loop is a more practical form of multielement parasitic array for the 30-, 40-, and 80-m bands.

The delta loop is a triangular or three-sided 1λ loop. The ideal configuration for a delta loop is with the horizontal side of the triangle at the top, but this is usually practical only at VHF. Turning the loop upside down allows it to be suspended from a single high support in return for a slight reduction in performance. Rope or wire catenaries between tall trees are a popular way of supporting the delta loop, but a two-element 40-m delta loop can be suspended from the ends of a 20-ft aluminum boom attached near the top of a tower that is 70 ft or more tall. The approach has been used on 80 m, but the boom length and minimum tower height requirements for that band typically lead to greater mechanical complexity, such as the use of trusses to strengthen the boom. The

two lower corners of the delta loop are usually tied off with nonconducting cordage to nearby trees or dedicated anchors.

Like the quad, the delta loop can be fed anywhere along the loop, with the ratio of horizontally to vertically polarized radiation changing accordingly. Feeding at the top or at the midpoint of the bottom wire results in a horizontally polarized loop. Moving the feedpoint halfway (or

$\lambda/4$

) along the total wire distance between those two points results in a vertically polarized loop. The feedpoint for vertical polarization is therefore

$\lambda/4$

down from the top on either side; if the loop is an equilateral triangle, this is equivalent to $\lambda/12$ above either bottom corner, or 1/4 the length of a side above one of the two bottom corners, as shown in Fig. 15.8.

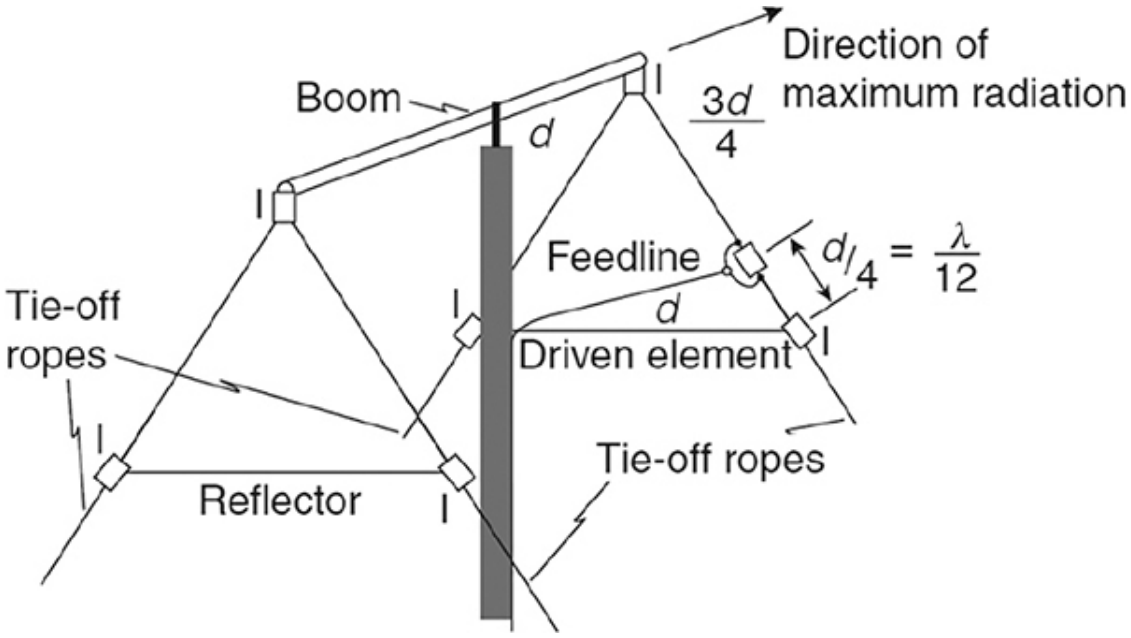


FIGURE 15.8 Vertically polarized two-element delta loop.

If the delta loop is an equilateral triangle, its total height is approximately $\lambda/(3.5)$, or 41 ft at 7 Mhz and 81 ft at 3.5 Mhz. For the loop to have reasonable gain, the bottom should be no closer to

the earth than around $\lambda/8$, so in practice the top support for a delta loop should be at least 70 ft high for 40 m and at least 120 ft high for 80 m.

Like Yagis and cubical quads, most users of a two-element delta loop prefer the parasitic element to be a reflector rather than a director. As with the quad, the reflector loop should be 3 to 4 percent longer than the resonant length for the desired center frequency. At HF, the earth will be close to the bottom of the loop, so experimentation to find the exact length that works best likely will be required. Some users use relays in protective enclosures at each loop to switch extra wire in and out of either or both of the two elements as a way of reversing the pattern or of having optimum lengths at two widely separated frequencies in the 3.5-MHz band.

Because one wavelength of wire is split among three sides instead of four, typically the bottom leg of an HF delta loop will be closer to the ground than the bottom leg of a quad for the same support height. In many cases, the delta loop is favored because the user has a single support that is higher than any other support, especially in the case of multielement delta loops, where the elements are often suspended from a long boom fixed-mounted near the top of the tallest tower in an antenna installation.

Because of its triangular shape and acute interior angles, the delta loop is a compromise compared to the square quad loop, as the square loop typically has higher gain. Nonetheless, the delta loop can do a passable job at providing low-angle gain on the lower HF bands.

15.3 Diamond Loop

The diamond loop is a square loop rotated by 45 degrees. When fed at a top or bottom corner, as shown in Fig. 15.9, it delivers 1.5 dB more forward gain in free space than a delta loop designed for the same frequency. But since the height of the diamond loop is twice that of the delta loop, the diamond loop is not usually a

candidate for the lower HF bands. Single-element and multielement versions might warrant consideration as fixed arrays at 20 m and above if the user has a single tall support or can stretch a catenary line between two such supports.

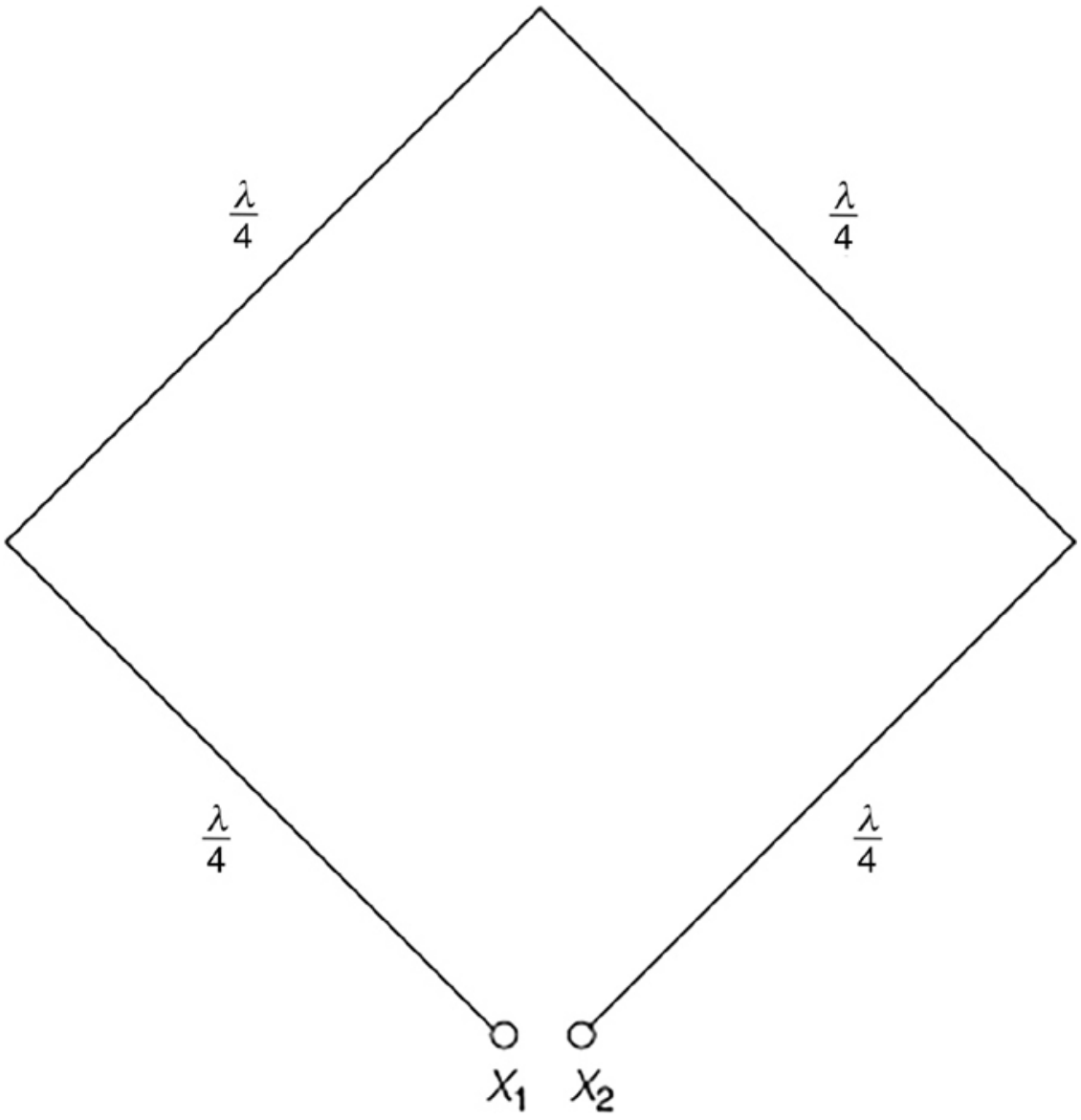


FIGURE 15.9 Horizontally polarized diamond loop.

15.4 Concluding Thoughts

HF antennas are usually low enough to the ground that ground reflections increase the gain. Theoretically, the reflection can

provide a 6 dB gain increase, but in practice, the reflection gain from real earth is less than 6 dB or the antenna is so low that the increased signal goes mostly into the ionosphere at high angles, where it may be of no use. Similarly, the theoretical 3-dB gain from the second radiating element (the other half of the loop) is seldom achieved in the real world. For 1*λ* loops and quad elements, the spacing of the opposing sides is only
$\lambda/4$
—leading to quadrature phasing and partial, rather than full, cancellation of the signal in unused directions. Equally important, the radiated field from the lower leg of a square loop is subject to the ground reflection pattern of a much lower height, so in general, a loop's maximum gain is at a somewhat higher elevation angle than it is for the top leg alone. Unless one has some other reason for choosing the quad—such as the high-power corona discharge problem at HCJB—the mechanical complexity of the quad may outweigh its benefits. Often, more can be realized with a two- or three-element Yagi at a height that yields the desired elevation pattern.

PART IV
Antennas for Specialized Applications

CHAPTER 16
Receiving Antennas
CHAPTER 17
Hidden and Limited-Space Antennas
CHAPTER 18
Mobile and Marine Antennas
CHAPTER 19
Emergency and Portable Antennas
CHAPTER 20
Antennas for 160 Meters
CHAPTER 21
VHF and UHF Antennas

CHAPTER 16
Receiving Antennas

Because antennas are almost always reciprocal devices, they can be used for transmitting and receiving. In many cases, the best antenna for a transmitter is also ideal as a receiver. But occasionally a separate receiving antenna can lead to better reception and an overall improvement in communications reliability. This can occur for the very low frequency (VLF), low frequency (LF), medium frequency (MF), and lower high-frequency (HF) ranges of the radio spectrum.

At low frequencies, the dimensions and construction costs of efficient transmitting structures (including the radials beneath them) are relatively large. Most transmitting antennas in use below 10 MHz consist of a single radiating element fixed in one position with a nearly omnidirectional pattern around one axis. As receivers, these antennas provide limited rejection of interfering signals from directions other than that of the desired signal.

Further, at low frequencies atmospheric and galactic noise is many orders of magnitude above the noise floor of most receivers. High antenna efficiency is relatively unimportant and electrically smaller antennas can be used. Multielement directive receiving array designs capable of being steered to multiple headings can be built on smaller land parcels at a lower cost compared to a larger and more efficient transmitting antenna.

The ability to copy weak signals is often limited by atmospherics (lightning-generated noise, or QRN) or local noise sources that are stronger at lower frequencies and often have different arrival angles or wave polarization than the desired signal. It is often easier and more important to null out the noise with an easily steerable receiving antenna than it is to boost the incoming signal enough to override the noise. Phase-locked receivers each connected to its

own receiving antenna or one connected to a receiving antenna and the other to the transmit antenna can provide enhanced ability to copy weak signals in the presence of signal fading (QSB), ionospheric skew paths, and/or polarization shift.

As outdated long range navigation (LORAN) systems were decommissioned and 160-m band frequencies became available to amateurs, receiving antennas for that band have been developed. The low-band receiving enthusiast—whether radio amateur, broadcast band listener, or shortwave listener—has available an arsenal of antenna types, including the beverage and snake, multiturn loop, EWE, K9AY loop, pennant, flag, short verticals, and longwires.

Many of these antenna types are employed not only singly but also in phased configurations with duplicates of themselves. Due to low radiation resistance and radiation efficiency and poor power handling, all except the longwire are generally unsuitable for transmitting purposes. Because of their inefficiency, these antennas generally present a signal level to the receiver input terminals that is anywhere from 10 to 40 dB below the levels that would be delivered by an efficient transmitting antenna used on receive. But because the atmospheric noise level at MF and lower HF frequencies is so high, receiver sensitivity is seldom an issue, and preamplifiers are only occasionally required.

16.1 Beverage or "Wave" Antenna

The Beverage or wave antenna is considered by many to be the best receiving antenna available for very low frequency (VLF), AM broadcast band (BCB), medium-wave (MW or MF), or tropical band (low HF region) DXing.

In 1921, Paul Godley, 1ZE, journeyed to Scotland to listen for amateur radio signals from North America. After World War I, hams had been consigned to the supposedly useless shortwaves (

$\lambda < 200\ \mathrm{m}$

), and it was not clear that reliable international communication was possible. He was successful—23 North American amateur stations were heard. Godley reportedly used a wave antenna for the task. The antenna was described by Beverage in *QST* for November 1922, in an article entitled "The Wave Antenna for 200-Meter Reception." It then virtually disappeared from popular view for decades as amateurs, military, and commercial services moved higher and higher in frequency, until removal of LORAN restrictions on the 160-m band by the United States and other countries in the 1960s and 1970s caused an upswing of interest in that band.

The Beverage is a terminated longwire antenna, usually greater than one wavelength (Fig. 16.1). The Beverage provides good directivity but is not very efficient, as much of the supplied power is dissipated in a termination resistor at one end. As a result, it is seldom used for transmitting. Unlike the longwire, the Beverage is intended to be mounted close to the earth's surface (typically $< 0.1\lambda$); heights of 8 to 10 ft are most commonly employed, but lower heights can reduce stray pickup from the vertical drop wire at either end. At greater heights, the Beverage loses its relationship to the (lossy) ground beneath it and its pattern becomes that of a traditional end-fed terminated wire.

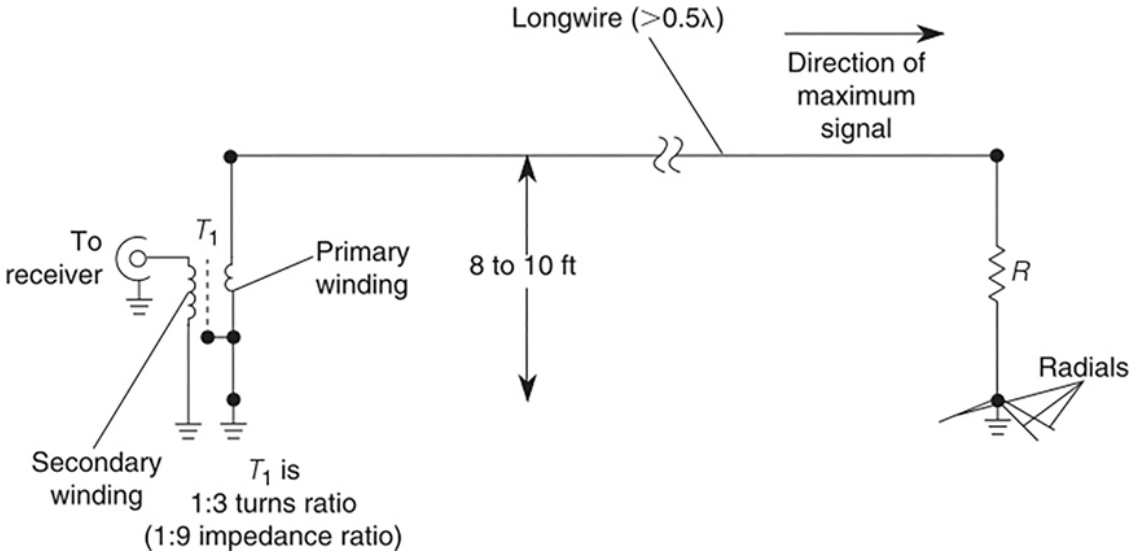


FIGURE 16.1 Beverage antenna.

When deployed over normal ground, the Beverage antenna responds to vertically polarized waves arriving at low angles of incidence. These conditions are representative of the AM broadcast band, where nearly all transmitting antennas are vertically polarized, and of incoming long-haul DX signals on the amateur bands.

The Beverage works best in the low-frequency bands (VLF through 40 m or so). Beverages are rarely used at higher frequencies because high-gain rotatable directional arrays have a much smaller cost and footprint than a set of six or so Beverages would require and can also be used for transmitting. At higher frequencies, the Beverage can lose the purity of its vertically polarized pattern due to its higher electrical height above ground.

The Beverage performs best when erected over moderately or poorly conductive soil. Figure 6.14 shows why poorly conductive soil is needed. The incoming ground wave (broadcast band) or low-angle (long-haul DX) E-field vectors arrive essentially perpendicular to the earth's surface. Over imperfectly conducting soil, the field lines tend to bend near the point of contact with the ground. As shown in the inset of Fig. 16.2, the incoming wave delivers a horizontal component of the E-field vector along the length of the

wire, thus generating an RF current in the conductor. Signals arriving from the unterminated end of the wire build up a voltage on the wire as they travel down the wire, but that voltage is absorbed, or dissipated, in the termination resistor. Signals arriving from the termination end travel to the feedline end, where they are collected by a step-down transformer or the feedline itself. If there is a mismatch at the junction of the Beverage wire and the feedline or at the receiver itself, the reflected signal arrives at the termination resistor and is absorbed.

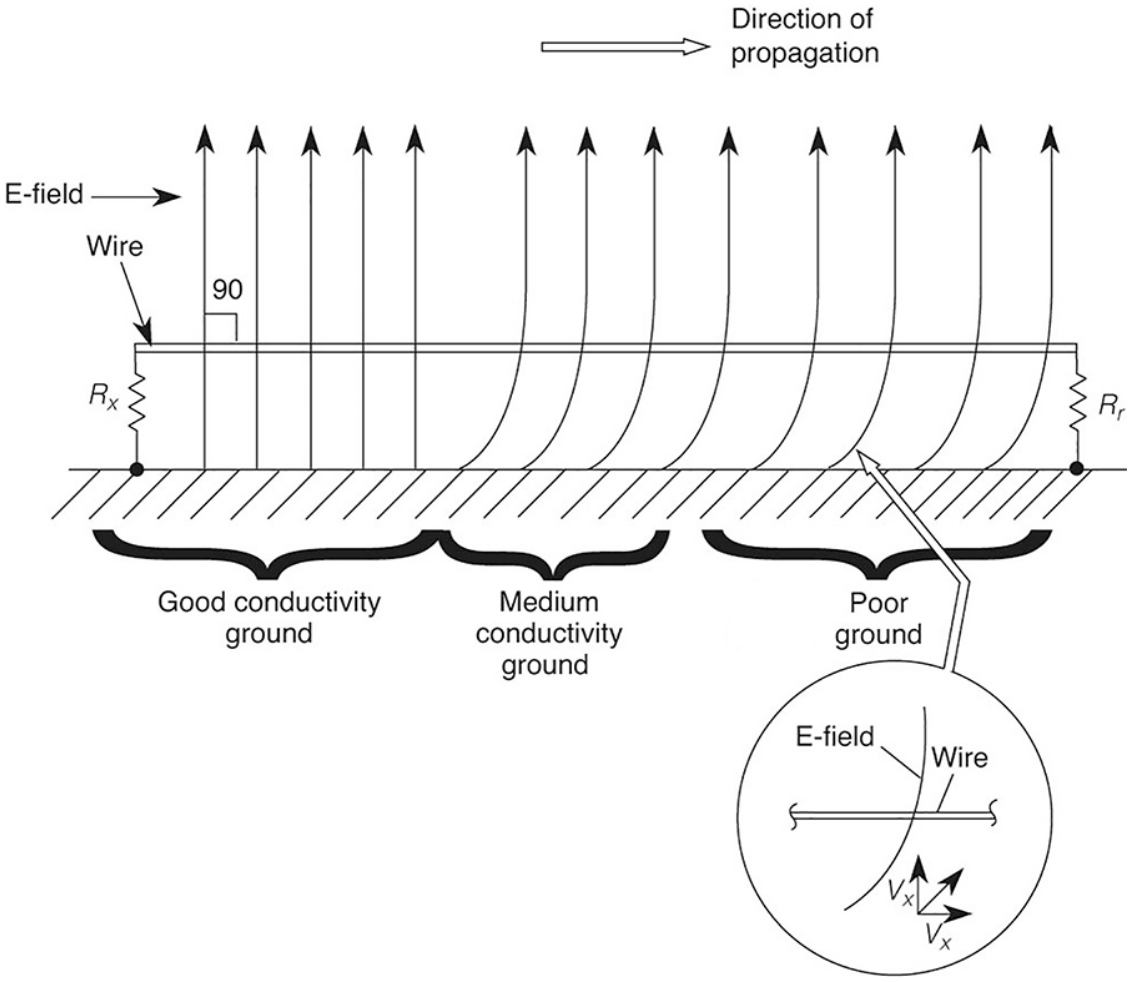


FIGURE 16.2 Wave tilt versus ground conductance.

There has been some debate over the optimal length for the Beverage antenna. But generally the exact length is not critical. It can range from around a half wavelength to nearly two

wavelengths, depending on the band. There is little to recommend in lengths beyond 500 to 550 ft (1λ on 160 m) for use of frequencies between 1 and 7 MHz. Some Beverages are in the 250- to 300-ft range. They have made it possible to pull weak stations out of the noise on 160 and 80 m on so many occasions, it is hard to believe there is much improvement left to be gained by lengthening them. The longer a Beverage is, the narrower the main lobe and more would be required to cover a large range of azimuth angles.

Like the nonresonant longwire antenna, the Beverage needs a termination resistor that is connected to a good ground. This requirement is somewhat in conflict with the assumption that Beverage antennas work best over lossy ground, but the solution is to make an artificial RF ground with radial wires at the terminated end. Insulated or bare wires at least 0.2λ long on the lowest frequency of interest make the best radials. A substantial improvement in the ground is possible using wires only 15 to 20 ft in length atop or just below the surface of the soil. Some experienced Beverage users recommend just one or two radials at right angles to the Beverage wire itself. Wires in close proximity to ground should be thought of as providing nonresonant capacitive coupling rather than resonance, so their exact length is not important. In addition to the radials and ground rod, a wire connection between the ground connection at the termination resistor and the ground connection at the receiver transformer can be used to stabilize impedance variations at higher frequencies.

The basic single-wire Beverage antenna of Fig. 16.1 consists of a single conductor (#18 to #8 wire gauges are most common) erected about 8 to 10 ft above ground. Because the RF power in a receiving Beverage is small, the size and metallic composition of the wire are primarily set by mechanical and corrosion lifetime considerations. The greatest threat to Beverage longevity comes from dead trees and limbs falling on them or animals running into them. Copper wires are often used but may need retightening over time. Some have used copper-clad steel or electric fence wire, but alternative metals can cause other problems.

Some Beverages are unterminated and bidirectional, but most are terminated at one end in a resistance *R* equal to the antenna's characteristic impedance, Z_0. For the heights discussed, Z_0 is likely to be between 300 and 800 Ω. The receiver end should also be terminated in its characteristic impedance. This usually requires a matching transformer to reduce the antenna impedance to a 50-Ω or 75-Ω standard impedance.

For safety, the antenna should be installed at a height of 8 to 10 ft above ground. The wire can be run using a compass to align with the desired beam heading. Typical Beverage lengths range from 250 to 1200 ft. If space permits, choose your Beverage location to maximize the distance between it and other low-band antennas, towers, power lines, buildings, and nearby sources of electrical noise. Support the wire as necessary (every 25 to 100 ft, depending on wire type). Beverages in the open can be mounted on treated 4×4 posts, but a potentially less expensive approach is to drive a few feet of an inexpensive steel rod (rebar, for instance) partway into the ground and drop an 8-ft length of PVC pipe over the rod. Drill or slot the top of the PVC to hold the wire in place.

In wooded areas insulated wire can be draped over tree limbs of the appropriate height, cord between trees, or electric fence insulators nailed into the tree trunk. Keep the wire a relatively constant height above ground following the slope of the terrain. Drop the vertical end sections straight down or bring them down gradually through some horizontal distance. Drive a 6- or 8-ft ground rod at the termination end and attach one or two radials, 50 to 200 ft long, at right angles to the Beverage. The termination end is the end from which signals will be received. Connect a termination resistor from the bottom end of the drop wire at the termination end of the Beverage to a clamp on the ground rod there. Drive another ground rod and lay out a separate set of

radials at the opposite end of the wire. This is the feedpoint of the Beverage.

For the terminating resistor, a 470-Ω 2W carbon composition resistor makes a great Beverage termination. If you have only larger resistance values, connect two or more in parallel; the greater the installed wattage, the less susceptibility to damage from lightning-induced surges. The exact value of resistance is not critical, but the type of resistor is. Some wire-wound resistors are not suitable for this application because they are intended for use in DC or audio circuits and exhibit inductive reactance at RF. A suitable alternative to carbon composition is a thick-film resistor, such as Ohmite's TCH line.

Once the Beverage is installed, a good preliminary test is to measure the DC resistance between the unterminated drop wire and the ground rod at that end of the Beverage. With average ground characteristics,

R_{TOTAL}

(the sum of the termination resistance plus the ground path between the meter and the grounded end of the termination) will probably range from 1.5

R_{TERM}

to

$4R_{\mathrm{TERM}}$

, depending on the length of the Beverage and the characteristics of the ground beneath it. The primary value of this test is simply to prove continuity through the Beverage wire and termination resistor into the ground. This test should be performed when the installation is new, and the value of

R_{TOTAL}

can be recorded for future reference.

If you have the necessary test equipment, consider sweeping the impedance of the Beverage at its feedpoint over the frequency range of interest while you adjust the value of the terminating resistor for flattest response over the useful range using a noninductive potentiometer in place of the permanent terminating

resistor. Terminate or match the feedpoint end of the Beverage according to the instructions with the test equipment. This is an excellent use of a vector network analyzer (VNA) as discussed in Chap. 30.

An inexpensive Beverage feedline is 75-Ω RG-6 coaxial cable purchased by the 500- or 1000-ft spool from a building supplies store. With either 50-Ω or 75-Ω coaxial cable for the feedline, a 9:1 step-down transformer at the Beverage feedpoint can be used. Do not put the transformer at the receiver end of the feedline. Figure 16.3 shows how to build a transformer using a ferrite core and a trifilar winding. Type 31 or type 43 materials should work well, but many mixes will work nearly as well. Ideally, the ground for the shield of the coaxial feedline at the transformer needs to be kept isolated from the grounded end (winding terminal a_2 in Fig. 16.3) of the transformer winding that connects to the Beverage, which would require a second, completely isolated coaxial cable output winding identical to winding $a_1 - a_2$ on the same core and a second ground terminal.

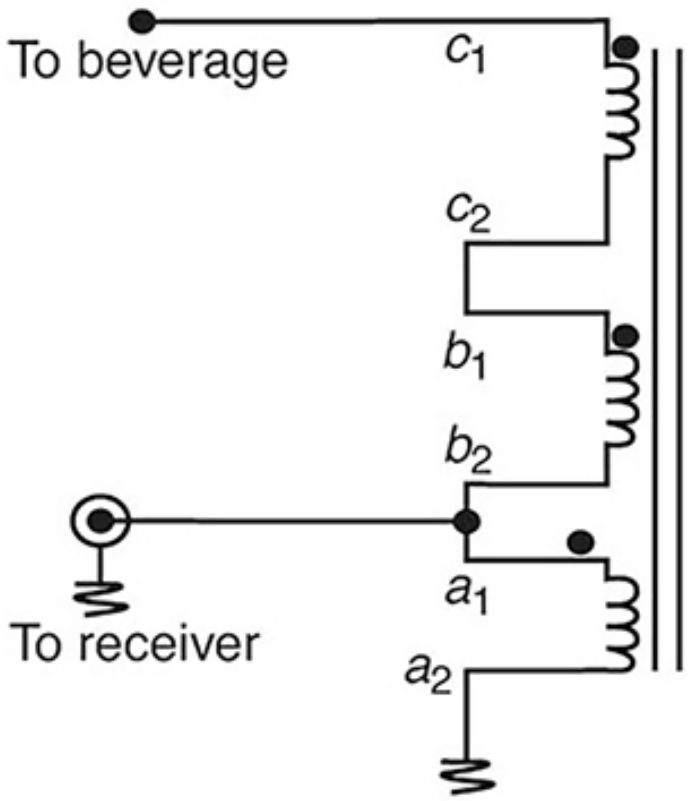


FIGURE 16.3 Trifilar-wound 9:1 step-down Beverage transformer.

After connecting the feedline to the step-down transformer at the unterminated end of the Beverage, perform the same DC continuity test with an ohmmeter located at the receiver end of the feedline, thus checking the integrity of that portion of the signal path as well. You should see a resistance of no more than a few ohms. At this point, if a VNA or other suitable piece of test apparatus is available, a sweep of the entire antenna system from the receiver end of the feedline to the termination resistance at the far end of the Beverage wire can be useful.

Much has been learned and disseminated about the Beverage over the years. But even imperfectly constructed Beverages can make a whole new layer of weak DX signals on the AM broadcast band or lower frequency ham bands jump out of the background noise. Short Beverages may need no more than two or three hours to install and connect and they work on 160 m, 80 m, 40 m, and

the AM broadcast band. The antennas could be fine tuned but perform well as is and they are cheap and easy to install. In short, the Beverage is the epitome of a practical antenna.

16.1.1 The Bidirectional Beverage

Two Beverages, consisting of two parallel wires and capable of receiving from two separate compass headings 180 degrees apart, can share a common set of supports and a common feedline attached to the end of the wire pair that is nearer the receiver.

The bidirectional beverage requires a fairly complex transformer in the feed network. Figure 16.4 shows the circuit configuration that allows switching of the wires and the feedlines for two opposing directions. The parallel wires can be implemented with separate wires about a foot apart, supported by cross-arms atop wood or PVC supports, open-wire transmission line, jacketed house wiring, or surplus multiconductor cable.

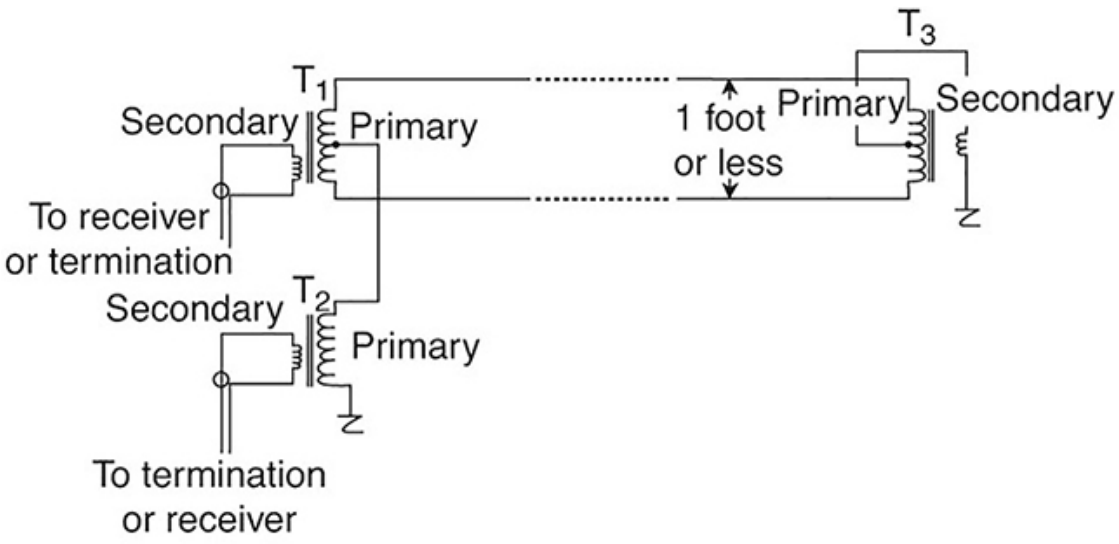


FIGURE 16.4 Two-wire reversible Beverage.

Transformers T_1 and T_2

convert the Beverage impedance to the characteristic impedance, Z_0, of the coaxial feedline to the receiver. One of the two coaxial cables feeds the receiver while the other cable must be terminated in Z_0. This can be done at the receiver end, or to avoid running two transmission lines to the receiver, the resistive load switching and swapping of transformers feeding the receiver can be performed in a protective enclosure near the end of the Beverage that is attached to transformer T_1. In the latter configuration, the DC control signal to switch a multipole double-throw relay in the enclosure can be multiplexed on the coaxial cable.

To analyze the feeding network, we assume first that the feedline attached to transformer T_1 is terminated in Z_0. Signals coming from the right side of the page build in voltage with respect to ground along both Beverage wires equally. At the left end of the wires, this common mode voltage appears across the entire primary winding of T_1, including its centertap, and drives the top end of the primary winding of T_2, whose secondary feeds the active coaxial cable to the receiver. Signals coming from the left side of the page build up equally on both Beverage wires from left to right, appearing on the entire primary of T_3, including its centertap, which drives one end of a secondary winding. The signal across the secondary of transformer

T_3
now drives the two Beverage wires in push-pull or differential mode. That signal propagates from right to left along the wires until it reaches the primary of
T_1
, where it is coupled to the secondary and dissipated in that winding's
Z_0
termination. To receive signals from the opposite direction, the receiver is connected to the secondary of
T_1
, and the secondary of
T_2
is terminated in
Z_0
.

The number of turns required on each transformer is dependent upon the specifics of each individual Beverage installation, including wire diameter and length, spacing between the two wires, height above ground, and ground conductivity. If the termination and cable switching is done near
T_1
and
T_2
, care must be taken to isolate the ground end of the termination from the coaxial cable ground (through a separate pole on the relay), or the ground connecting the two must be virtually perfect.

16.1.2 Phased Beverages

Additional directivity and signal amplitude can be obtained by combining two or more Beverages with a controlled phase relationship. Two identical Beverages, parallel to each other, spaced
$\lambda/4$
or more apart, with no offset relative to the desired receiving direction. They can be thought of as two opposing long sides of a

rectangle. The outputs of the two wires are combined in phase before reaching the receiver. Alternately, two Beverages can be located parallel to each other and closely spaced (i.e., within a few feet) but offset somewhat in the desired receiving direction (picture the long sides of a very skinny parallelogram).

A Beverage erected with two wires parallel to each other, at the same height, spaced about 12 in apart (Fig. 16.5), with a length that is a multiple of a half-wavelength is capable of null steering. The rear null in the pattern can be steered over a range of 40 to 60 degrees. This feature allows strong, off-axis signals to be reduced in amplitude so that weaker signals in the main lobe of the pattern can be received.

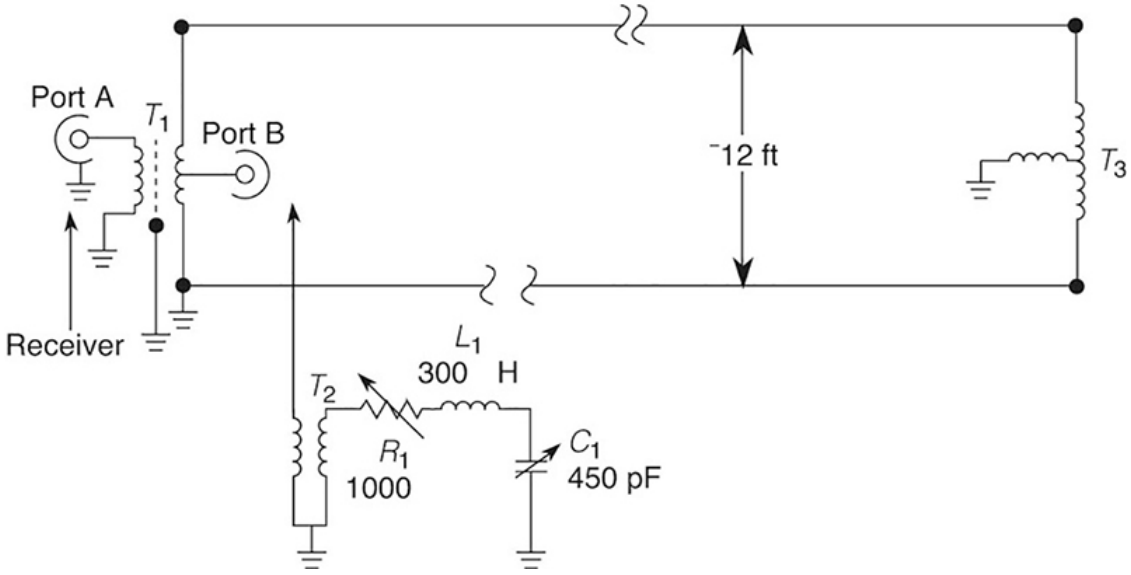


FIGURE 16.5 Beverage antenna with a steerable null.

If null steering behavior is desired, then a phase control circuit (PCC) is required. This can be accomplished with a potentiometer, an inductance, and a variable capacitor connected in series. Varying both the pot and the capacitor will steer the null. The direction of reception and the direction of the null can be selected by using a switch to swap the receiver and the PCC between port A and port B.

16.1.3 Snake or BOG Antenna

The height above ground of a Beverage is the result of a compromise that includes a need to locate the antenna high enough to avoid damage to or from humans and animals. As a result, the rejection of signals from undesired directions and signals with horizontal polarization is lessened.

The Beverage on ground (BOG) or snake antenna attempts to minimize unwanted signal ingress by eliminating the vertical downlead at each end of the Beverage. A side benefit is the elimination of all supports. A disadvantage is a reduction in received signal level stemming from the close coupling of the antenna to the ground beneath it. Typically, a preamplifier must be added to a BOG or snake in order to bring signal levels above the receiver noise floor. Ideally the preamp should be located at the junction of the Beverage and its feedline, rather than at the receiver end of the feedline.

16.1.4 Limitations of the Beverage Antenna

Even in their smallest implementations, the preceding antennas require a fair amount of real estate for their proper operation. In addition to the space taken up by the antenna itself, there is a hidden requirement of another
$\lambda/2$
or more distance between the receiving antenna and any transmitting antennas resonant on the same band(s). Further, it is wise to locate and orient Beverages so as to keep sources of RF noise behind the antenna (i.e., at the receiver, or unterminated, end). A broadcast-band listener (BCL) or shortwave listener (SWL) contemplating the installation of Beverages for multiple compass headings will usually need a 10-acre or larger parcel.

16.2 Receiving Loops

For users with limited space, two entirely different classes of antenna have been of great use. The first of these is the small,

multiturn loop that has been around for years. The second is a class of small, single-turn loops that emulate the directivity of large transmitting loops while exploiting size reduction design techniques uniquely available to receiving antennas that do not need to exhibit high RF efficiency.

16.2.1 Small Loop Receiving Antennas

Radio direction finders and people who listen to the AM broadcasting bands, VLF, medium-wave, or the so-called low-frequency tropical bands are candidates for a small loop antenna. These antennas are different from the large (transmitting) loops of Chap. 9 and the multielement versions of Chap. 15. Large loop antennas typically have a total wire length of 0.5λ or greater. Small loop antennas, on the other hand, have an overall total conductor length that is less than about 0.1λ.

An electrically small coil of wire is an inductor and in most cases radiation from a coil is so small that we can ignore it. The radiation efficiency is so low that small loops are rarely used as transmitters. For receiving applications, particularly at bands below around 10 MHz where sky noise is especially high, a small coil can be used as an antenna in spite of the low radiation efficiency. For receivers, low radiation efficiency translates into reduced signal level and additional thermal noise added by the antenna, but if the signal level is adequate and other noise sources dominate, a small loop antenna with low efficiency may be used in place of a much larger dipole-type antenna.

In a large loop, the current is a standing wave that varies sinusoidally from one point in the conductor to another, much like a dipole, and we analyze its operation in much the same way we have examined dipoles. In the small loop antenna, the current is the roughly constant throughout the entire loop.

Since the electric and magnetic fields of a radiated wave in free space are at right angles to each other and to the direction the wave propagates through space, and we are primarily listening for

vertically polarized signals with loops, the magnetic field is oriented horizontally.

The differences between small loops and large loops show up in some interesting ways, but perhaps the most striking difference is found in the main lobes and the directions of the nulls. Both types of loops produce figure eight patterns but at right angles with respect to each other. The large loop antenna produces main lobes that are broadside to the plane of the loop. Nulls are off the sides of the loop. The small loop, however, is exactly the opposite: the main lobes are off the sides of the loop (in the plane of the loop), and the nulls are broadside to the loop plane (Fig. 16.6).

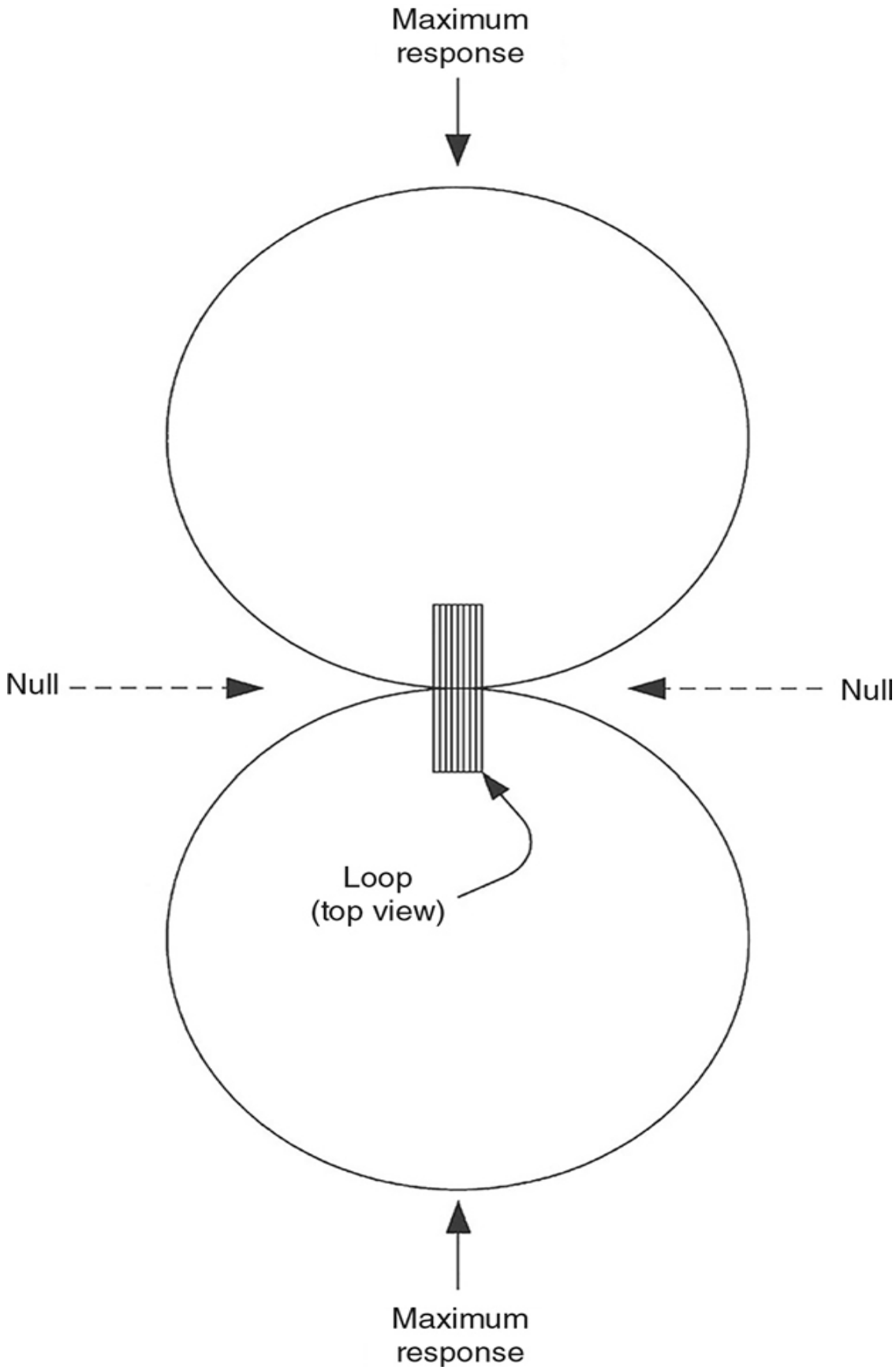

Maximum
response
Null
Null
Loop
(top view)
Maximum
response

FIGURE 16.6 Small loop antenna.

The small loop should be contrasted with the loopstick antenna. Loopstick antennas are made of coils of wire wound on a ferrite or powdered-iron rod. The direction of maximum response for the loopstick antenna is broadside to the rod, with deep nulls off the ends (Fig. 16.7). Both loopsticks and small wire loops are used for radio direction-finding and for shortwave, low-frequency medium-wave, AM broadcast band, and VLF listening.

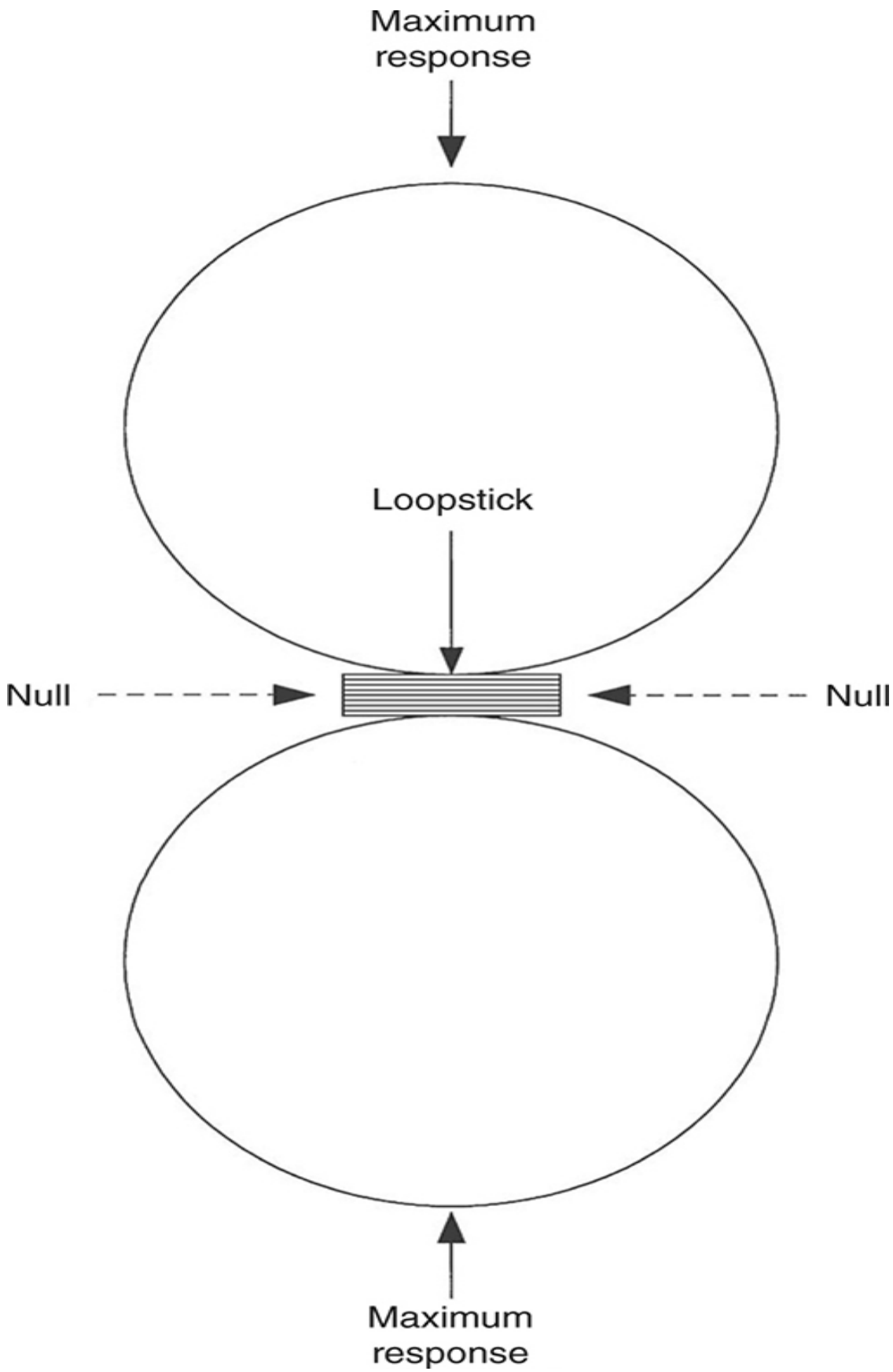

Maximum response
Loopstick
Null
Null
Maximum response

FIGURE 16.7 Loopstick antenna.

The nulls of a loop antenna are very sharp and very deep. If you point a loop antenna so that its null is aimed at a strong station, the received signal strength drops dramatically at the center of the notch. Turn the antenna a few degrees one way or the other and the signal strength increases sharply. The depth of the null can reach 10 to 15 dB on sloppily constructed loops and 30 to 40 dB on well-built units (30 dB is a common value). The sharpness and depth of the null is largely determined by the construction and uniformity of the loop. This feature is useful for radio direction-finding (RDF) as discussed in Chap. 26.

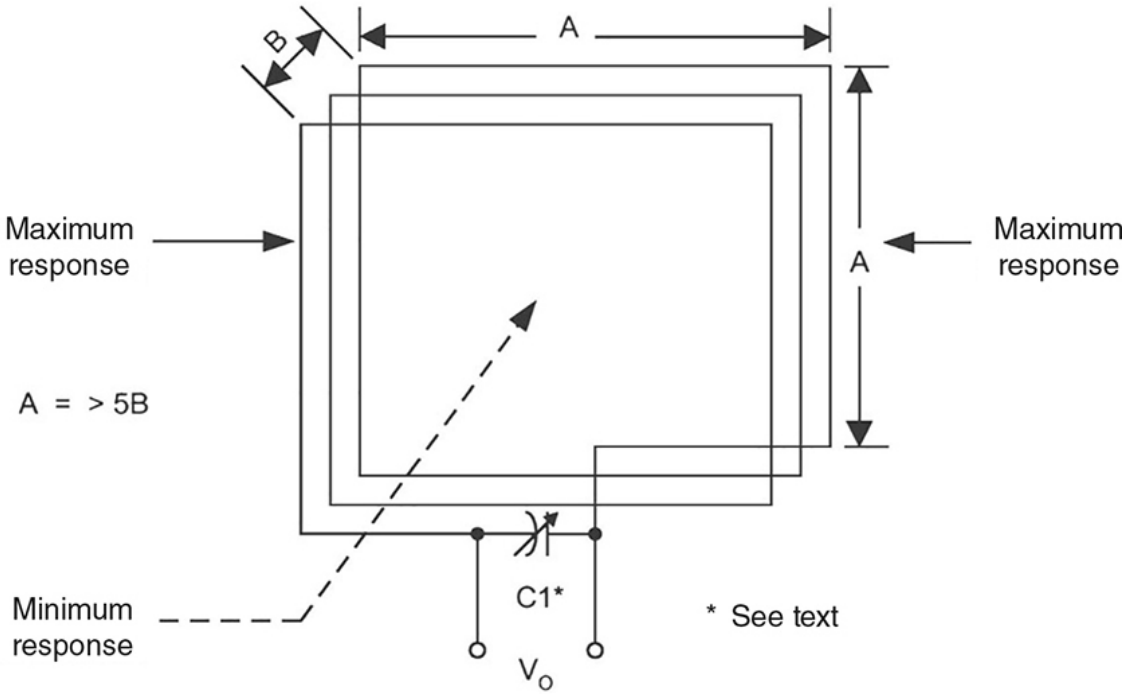
When small loops are used for receiving at low frequencies (VLF through lower HF range), even though the loop is electrically small, it can still be physically rather large and require many turns of wire to develop the necessary signal amplitudes and notch depth. Receiving loops can also be used at VHF and are commonly used in the 10-m ham band for hidden transmitter hunts.

16.2.2 Air Core Frame ("Box") Loops

A wire loop antenna is made by winding a large coil of wire, consisting of one or more turns, on some sort of frame. The sloop can be circular, square, triangular, hexagonal, or octagonal, but for practical construction reasons the square loop is most popular. With one exception, the loops considered in this section will be square, so you can easily duplicate them.

The basic form of loop, as shown in Fig. 16.8, is square, with sides of length *A*. The width of the loop (*B*) is the distance from the first turn to the last turn in the loop or, alternatively, the diameter of the wire if only one turn is used. The turns of the loop in Fig. 16.8 are depth-wound, meaning that each turn of the loop is spaced in a slightly different parallel plane, and spaced evenly across distance *B*. Alternatively, the loop can be wound such that the turns are in the same plane (known as planar winding). In

either case, the sides of the loop (*A*) should be at least five times the width (*B*). There seems to be little difference in performance between depth- and planar-wound loops. The far-field patterns of the different shape loops are nearly the same if the respective cross-sectional areas (
πr^2
for circular loops and
A^2
for square loops) are less than
$\lambda^2/100$
.



F IGURE 16.8 Simple loop antenna.

The reason a small loop has a null when its broadest aspect is facing the signal is simple, even though it seems counterintuitive at first glance. In Fig. 16.9, we have two identical small loop antennas at right angles to each other. Antenna A is in line with the advancing radio wave, whereas antenna *B* is broadside to the wave. The magnetic field associated with the wave is orthogonal to the

direction of propagation (horizontal in the figure) and passes through loop A, inducing a voltage around the loop. For antenna B, the magnetic field does not pass through the loop and no voltage is induced.

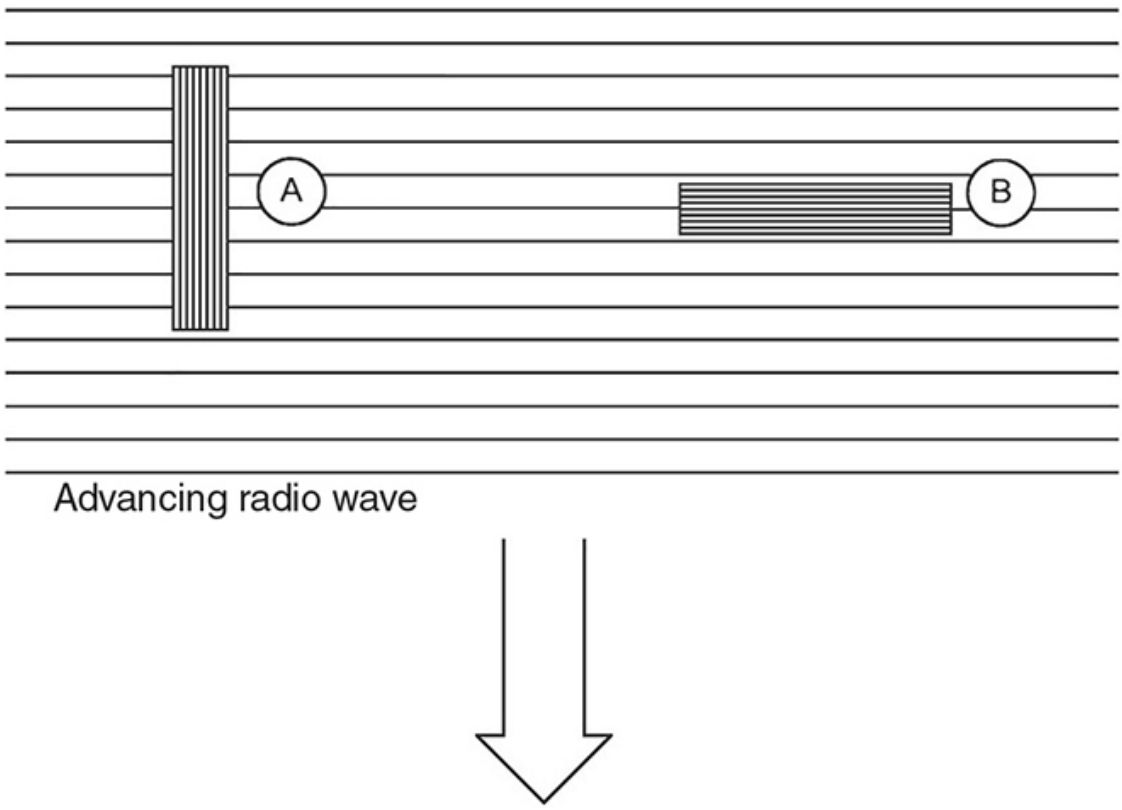


FIGURE 16.9 Two small loop antennas at right angles to each other.

The actual voltage across the output terminals of an untuned loop is a function of the angle of arrival of the signal *α* (Fig. 16.10), as well as the strength of the signal and the design of the loop. The voltage is given by

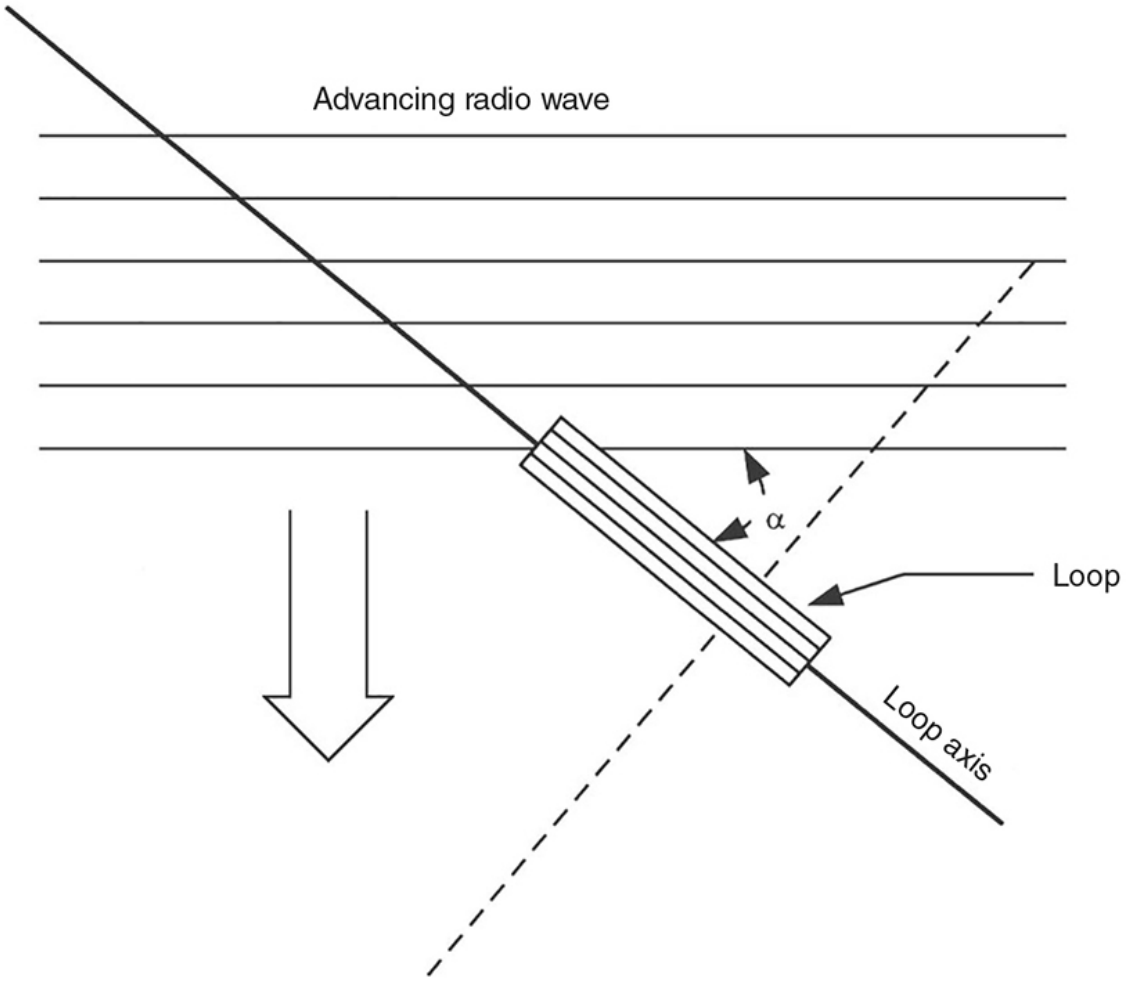
$$V_o = \frac{2\pi N A E_f \cos\alpha}{\lambda} \quad (16.1)$$

where *A* is the area of loop, *N* is the number of turns, E_f

is the incident field strength in V/m, *α* is the angle of arrival of the signal and *λ* is the wavelength. Loops are sometimes specified in terms of the effective height of the antenna, that has a height of

$$H_{\text{eff}} = \frac{2\pi NA}{\lambda} \quad (16.2)$$

This is the length of a wire antenna with the same output as the loop.



F IGURE 16.10 Untuned loop antenna at an angle to received wave.

If a capacitor (such as C_1

in Fig. 16.8) is used to tune the loop, then the output voltage V_o will rise substantially. The output voltage is multiplied by the loaded Q of the tuned circuit,

$$V_o = \frac{2\pi N A E_f Q \cos\alpha}{\lambda} \quad (16.3)$$

where the Q may range from 50 to 100 for typical small loops. Even though the output signal voltage of tuned loops is higher than that of untuned loops, it is nonetheless low compared with other forms of antenna. A preamplifier usually is needed for best performance.

16.2.3 Transformer Loops

It is common practice to make a small loop antenna with two loops rather than just one. Figure 16.11 shows such a transformer loop antenna. The main loop consists of several turns of wire on a large frame with a tuning capacitor to resonate it at the frequency of choice. The other loop is a one- or two-turn coupling loop. This loop is installed close to the main loop, usually on the inside edge a few centimeters away. The second loop couples signal induced from the main loop to the receiver at a more reasonable impedance match.

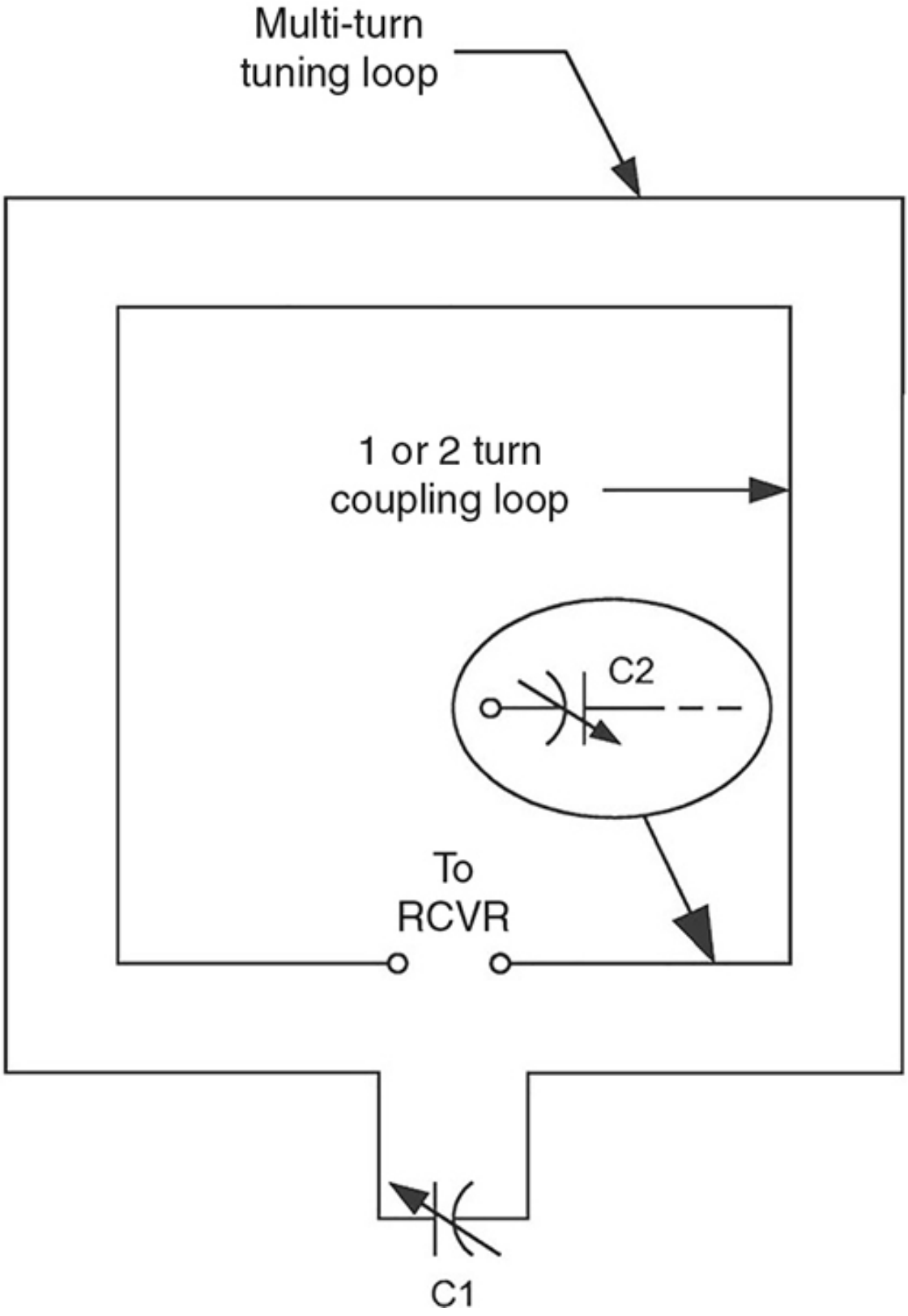


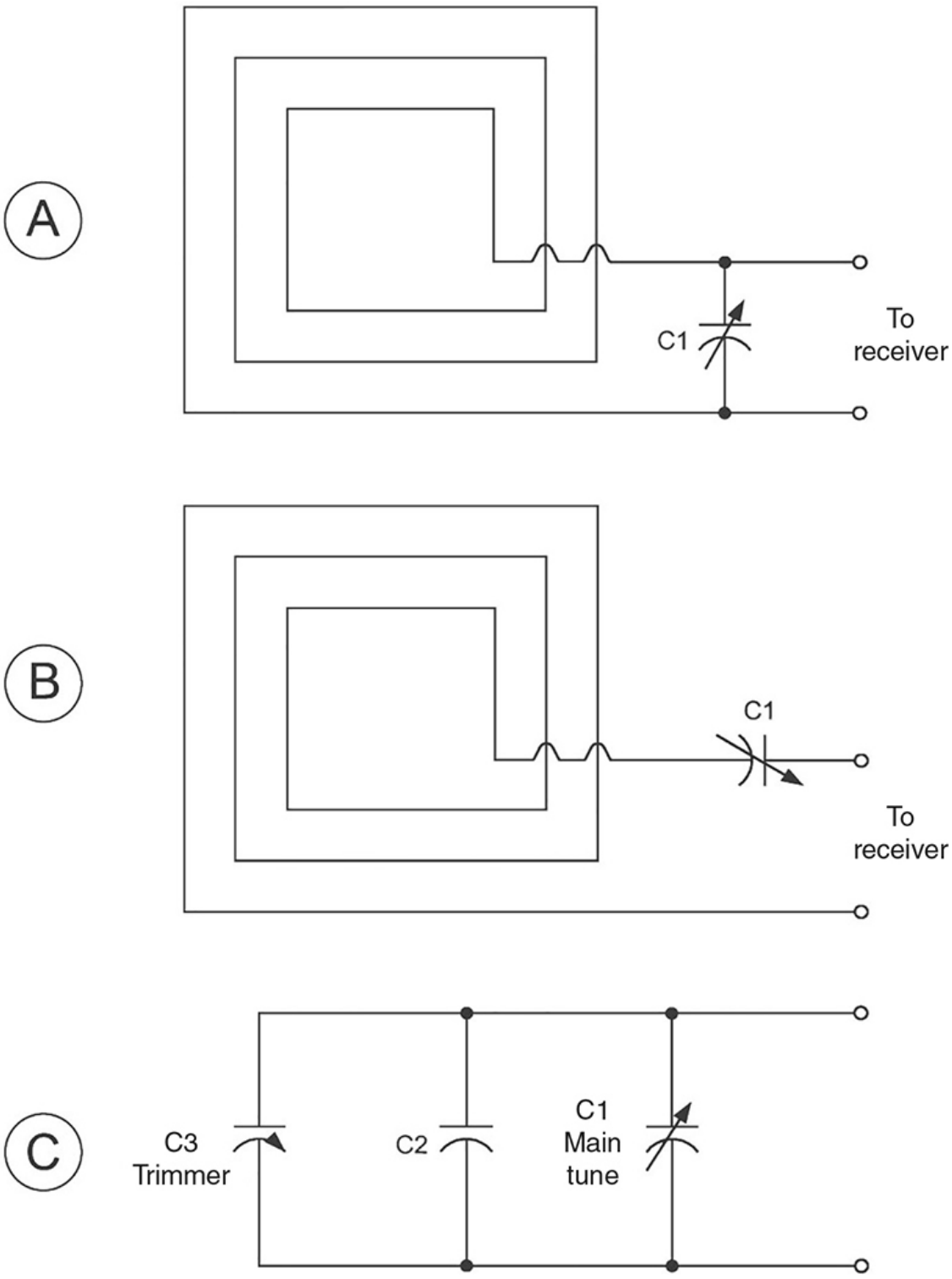
FIGURE 16.11 Transformer loop antenna.

The coupling loop is usually untuned, but in some designs a tuning capacitor (C_2) is placed in series with the coupling loop. Because there are fewer turns on the coupling loop than on the main loop, its inductance is smaller. As a result, the capacitance to resonate is usually much larger. A 15-turn main loop may resonate in the AM band with a standard 365-pF capacitor, for example, whereas the two-turn coupling loop requires three sections of a ganged $3 \times 365\text{-pF}$ capacitor connected in parallel to resonate at the same frequency. The coupling loop can be constructed in a variety of ways, including multiconductor ribbon cable.

16.2.4 Tuning Schemes for Loop Antennas

Loop performance is enhanced by tuning the inductance of the loop to the desired frequency. The bandwidth of the loop is reduced, which reduces front-end overload from strong off-frequency signals. Tuning increases the signal level available to the receiver by a factor of 20 to 100 times. Although tuning can be logistically difficult if the loop is installed remotely from the receiver, the benefits are usually worth the effort.

Various techniques for tuning are detailed in Fig. 16.12. The parallel tuning scheme, which is the most popular, is shown in Fig. 16.12A. The capacitor (C_1) is connected in parallel with the inductor, which in this case is the loop. Parallel-resonant circuits present a high impedance to signals on their resonant frequency and a low impedance to other frequencies. As a result, the voltage level of the signal at resonance is much larger than the voltage level of off-frequency signals.



F IGURE 16.12 Various tuning schemes. (A) Parallel. (B) Series resonant. (C) Series or parallel.

The series-resonant scheme is shown in Fig. 16.12B. Series resonant circuits offer a high impedance to all frequencies except the resonant frequency (exactly the opposite of the case of parallel-resonant circuits). As a result, current from the signal will pass through the series-resonant circuit at the resonant frequency, but off-frequency signals are blocked by the high impedance.

There is a wide margin for error in the inductance of loop antennas. Equations for the capacitance and inductance needed for tuning are only estimates. The geometry of the loop determines the actual inductance. Tuning provided by the capacitor is approximate, so some form of compensation is needed. In some cases, the capacitance required for resonance is not easily available in a standard variable capacitor, and means must be provided for changing the capacitance. Figure 16.12C shows how this is done. The main tuning capacitor can be connected in either series or parallel with other capacitors to fine tune the overall value. If the capacitors are connected in parallel, the capacitances add. If the extra capacitor is connected in series, the total capacitance is reduced. The extra capacitors can be switched in and out of a circuit to change frequency bands.

Hand tuning may be difficult, so remote tuning can be used. Traditional means of tuning called for a low-rpm DC motor or stepper motor to turn the tuning capacitor, or a varactor diode can be used. This is a reverse-biased diode whose junction capacitance is a strong function of the applied DC bias voltage. A high voltage (such as 30 V) drops the capacitance, whereas a low voltage increases it. Varactors are available with maximum capacitances of 22, 33, 60, 100, and 400 pF. The latter are of great interest to us for this application because they have the same range as the tuning capacitors normally used with loops.

Figure 16.13 shows a remote tuning system for loop antennas. The tuning capacitor is a combination of a varactor diode plus two optional capacitors: a fixed capacitor (C_1) and a trimmer (

C_2
). The DC tuning voltage (
V_T
) is provided from a fixed DC power supply (
V_+
) located at the receiver end. A potentiometer (
R_1
) is used to tune the loop by setting the voltage going to the varactor. A DC blocking capacitor (
C_3
) keeps the DC tuning voltage from being shorted out by the receiver input circuitry.

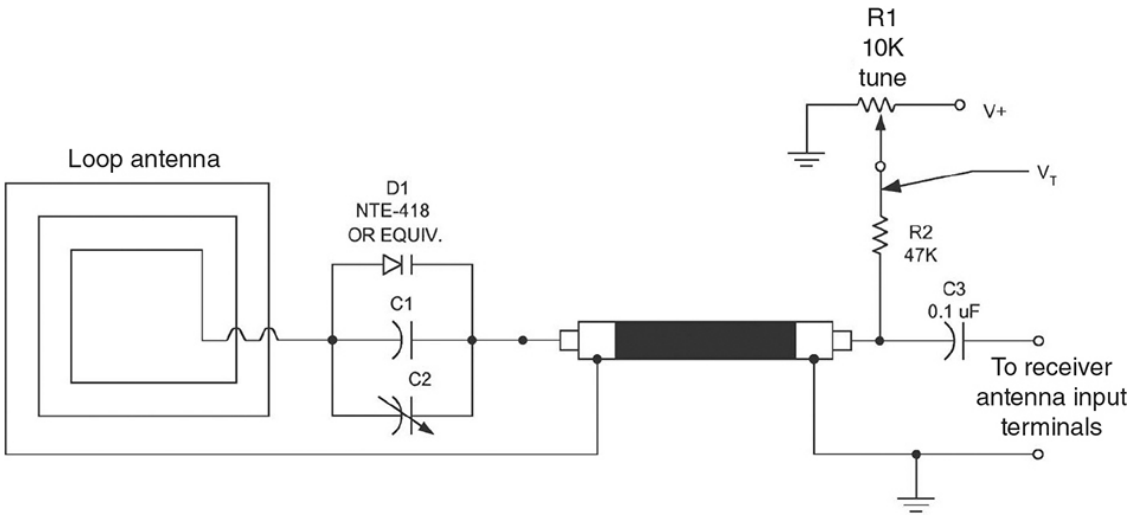


FIGURE 16.13 Remote tuning scheme using varactor diode.

16.2.5 The Sports Fan's Loop

Dating back to the day when the only way to follow a game was to tune into a weak AM broadcast, the signal strength can be improved with the sports fan's loop antenna. For the AM band, this was done with a 16-turn,
60-cm^2
tuned loop and an AM portable radio at the center with its loopstick aimed so that its null end is broadside of the loop. The nulls of both the loop and the loopstick are in the same direction. The signal will

be picked up by the loop and then coupled to the radio's loopstick antenna. Sixteen-conductor ribbon cable can be used for making the loop. A 365-pF tuning capacitor is used to resonate the loop. If you listen to only one frequency, this capacitor can be a trimmer type. By rotating the antenna, a weak signal from a desired direction can be received.

16.2.6 Shielded Loop Antennas

The loop antennas discussed thus far in this chapter have all been unshielded types. Unshielded loops work well under most circumstances, but in some cases their pattern is distorted by interaction with the ground and nearby structures (trees, buildings, etc.). Some interactions are so severe that the pattern is distorted beyond all recognition.

Interaction can be reduced by shielding the loop, as in Fig. 16.14. Loop antennas operate on the magnetic component of the electromagnetic wave, so the loop can be shielded against E-field signals and electrostatic interactions. In order to retain the ability to pick up the magnetic field, a gap is left in the shield at one point. The shield can be constructed from adhesive-backed copper-foil tape or aluminum foil. Another method is to insert the loop inside a copper or aluminum tubing frame.

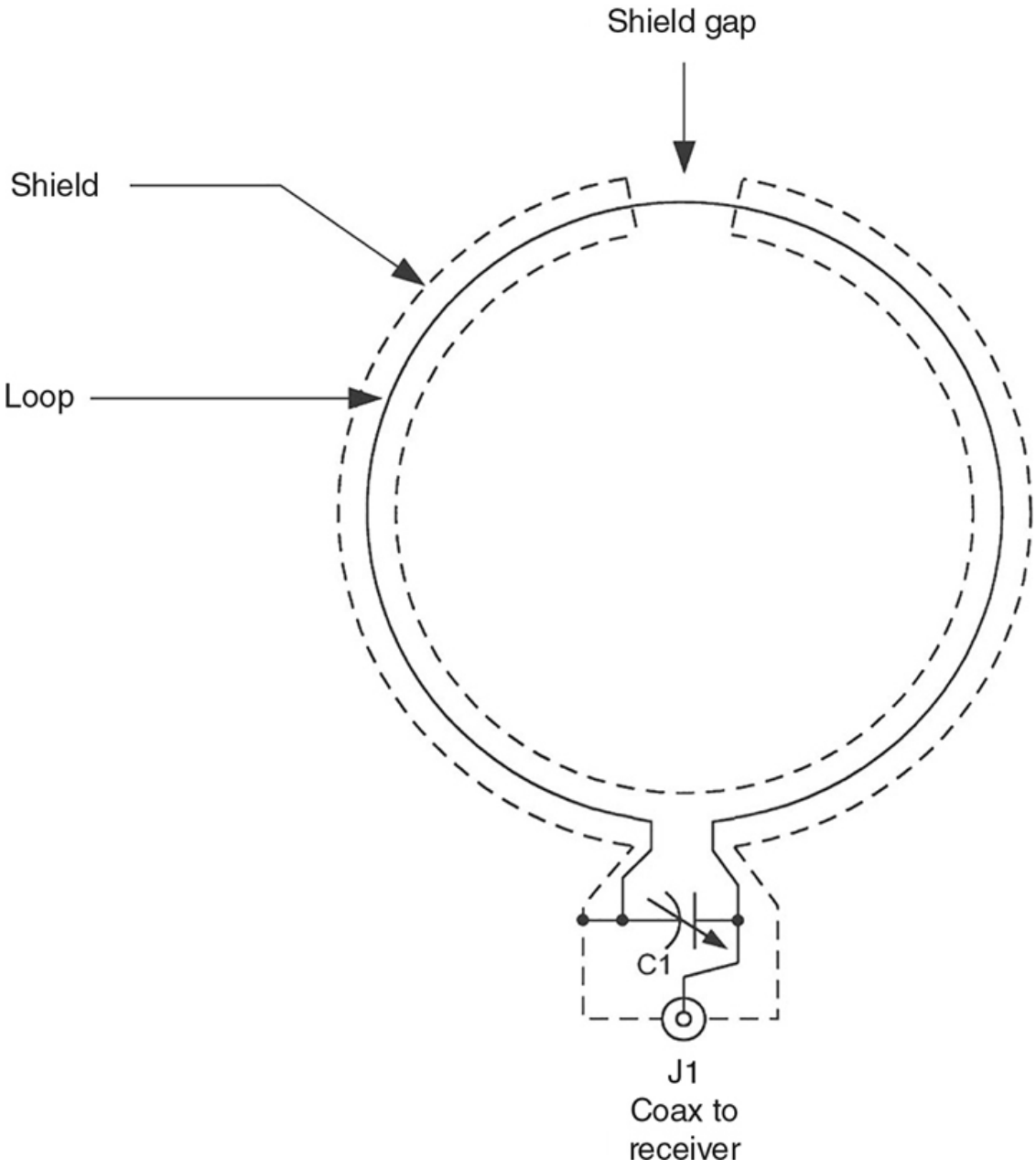


FIGURE 16.14 Shielding a loop antenna.

16.2.7 Sharpening the Loop

Receivers in the past often included a *Q*-multiplier. The *Q*-multiplier improves receiver performance by increasing the sensitivity and reducing the bandwidth of certain receiver stages.

The *Q*-multiplier of Fig. 16.15 is an active electronic circuit placed at the antenna input of a receiver. It is essentially an Armstrong

oscillator that does not quite oscillate. These circuits have a tuned circuit (L_1/C_1) at the input of an amplifier stage and a feedback coupling loop (L_3). The degree of feedback is controlled by the coupling between L_1 and L_3 . The coupling is varied by changing both how close the two coils are and their relative orientation with respect to each other. Other versions use a series potentiometer in the L_3 side that controls the amount of feedback.

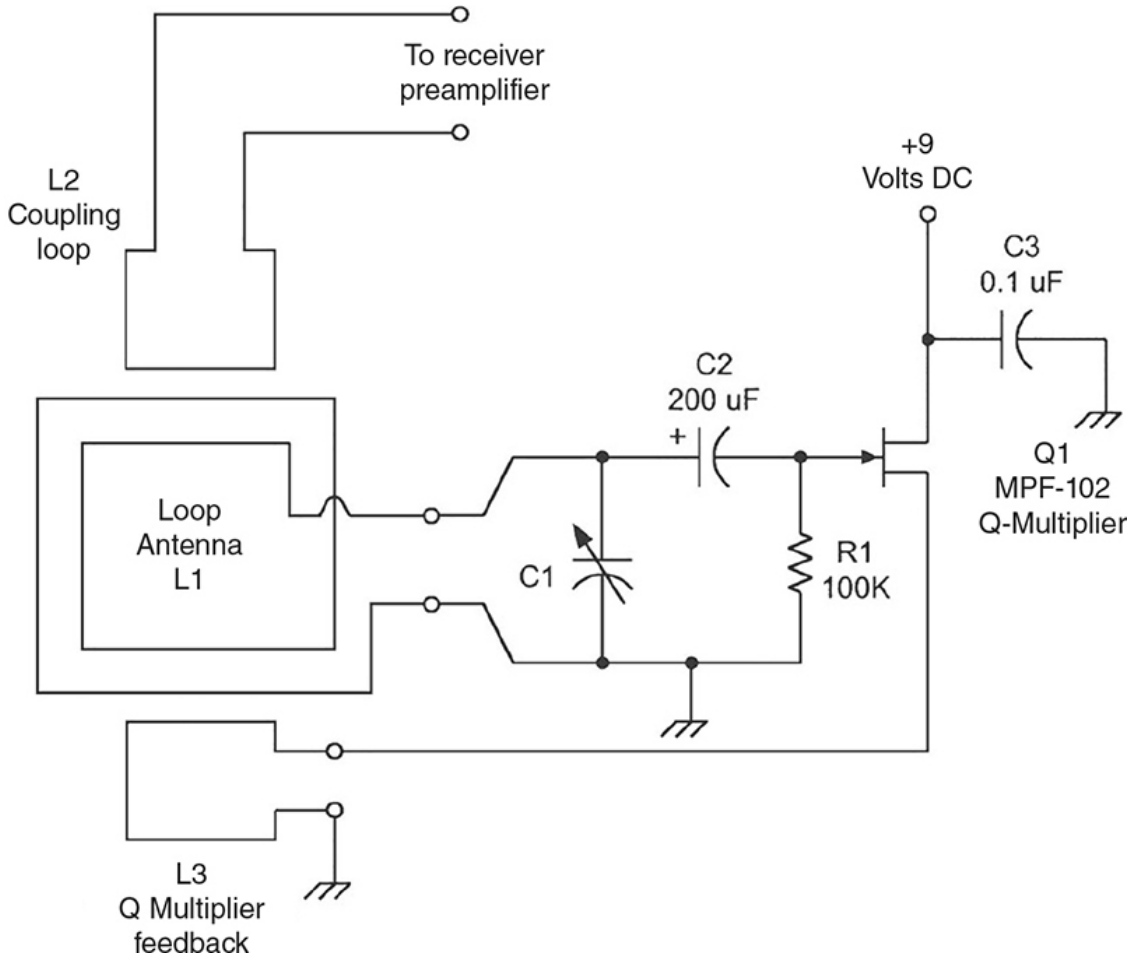


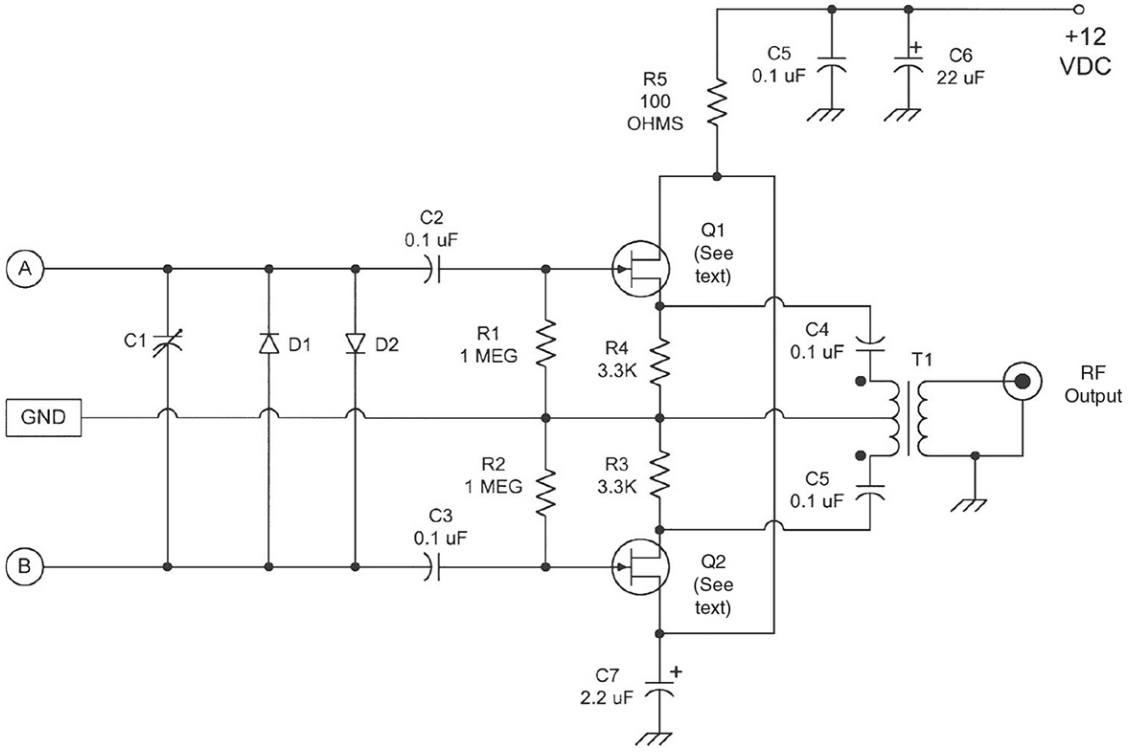
FIGURE 16.15 *Q*-multiplier.

The *Q*-multiplier is adjusted to the point that the circuit is just on the verge of oscillating. As the feedback is backed away from the threshold of oscillation—but not too far—the bandwidth narrows and the sensitivity increases. It takes some skill to operate a *Q*-multiplier, but it is easy to use once you get the hang of it and is a terrific accessory for any loop antenna.

16.2.8 Loop Amplifier

Figure 16.16 shows a loop amplifier that can be used with either shielded or unshielded loop antennas. The design junction field-effect transistors (JFETs) connected in cascade. The standard common-drain configuration is used for each transistor, so the

signals are taken from the source terminals. The drain terminals are connected and powered from the
$+12\ \mathrm{V\ DC}$
power supply. A 2.2-μF bypass capacitor is used to put the drain terminals of
Q_1
and
Q_2
at ground potential for AC signals while keeping the DC voltage from being shorted out.



F IGURE 16.16 Loop amplifier circuit.

The two output signals are applied to the primary of a transformer with grounded center tap. To keep the DC on the source terminals from being shorted through the transformer winding, blocking capacitors (
$\mathrm{C}_4, \mathrm{C}_5$
) are used. Input signals are applied to the gate terminals of

Q_1
and
Q_2
through DC blocking capacitors
C_2
and
C_3
. Diodes
D_1
and
D_2
keeps high-amplitude noise transients from affecting the operation of the amplifier. They are connected back to back in order to snub out both polarities of signal. A tuning capacitor
C_1
is used in lieu of a capacitor in the loop; it resonates the loop at a specific frequency. The transistors used for the push-pull amplifier (
Q_1
,
Q_2
) can be nearly any general-purpose JFET device (MPF-102, MPF-104, etc.).

16.2.9 Self-Resonance for VLF/LF Loops

A capacitance is formed whenever two conductors are side by side. A coil exhibits capacitance as well as inductance because the turns are side by side. With large multi-turn loops, this capacitance can be quite large. The distributed capacitance of the loop self-resonates with the inductance. Loop antennas do not work well at frequencies above their self-resonant frequency, so it is sometimes important to raise the self-resonance to where it does not affect operation at the desired frequencies.

Figure 16.17 shows one way to raise the self-resonant point. The turns on the loop are broken into two or more groups separated by spaces. This method reduces the effective capacitance by placing

the capacitances of each group of wires in series with the others. The effective capacitance of a series string of capacitors is always less than the value of the smallest capacitor.

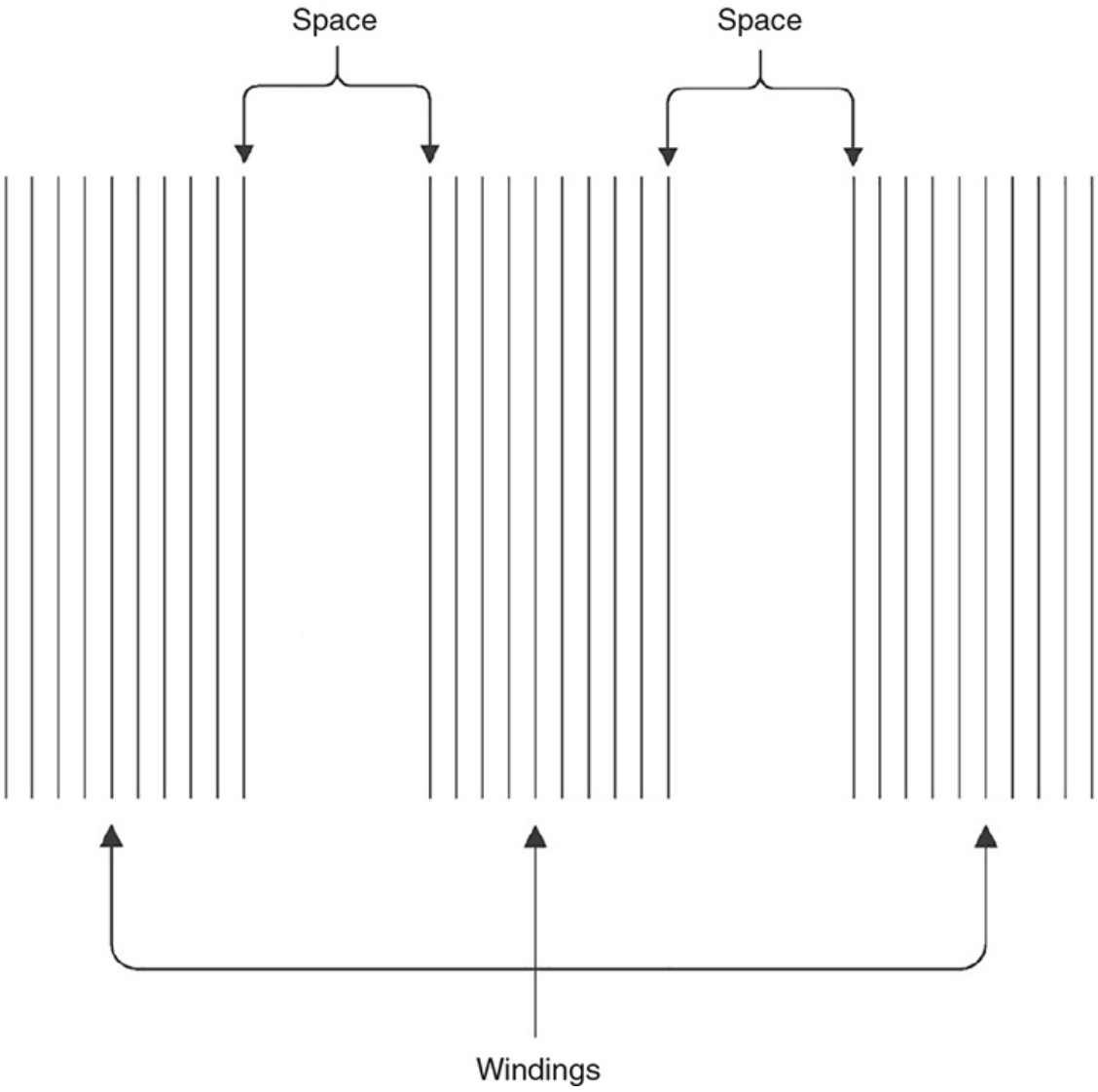


FIGURE 16.17 Raising the self-resonant frequency.

16.2.10 Coaxial-Cable Loop Antennas

One of the more effective ways to make a shielded loop is to use coaxial cable (Fig. 16.18). Although only a single-turn loop is shown, there can be any number of turns. The coaxial cable is connected so that the outer conductor is a Faraday shield. At the

output end, the center conductor of the coaxial cable is connected to the center conductor of the coaxial connector and the cable shield is connected to the connector ground/shield terminal. At the other end of the loop, the shield is left floating. The center conductor is connected to the shield at the connector end, not at any other point.

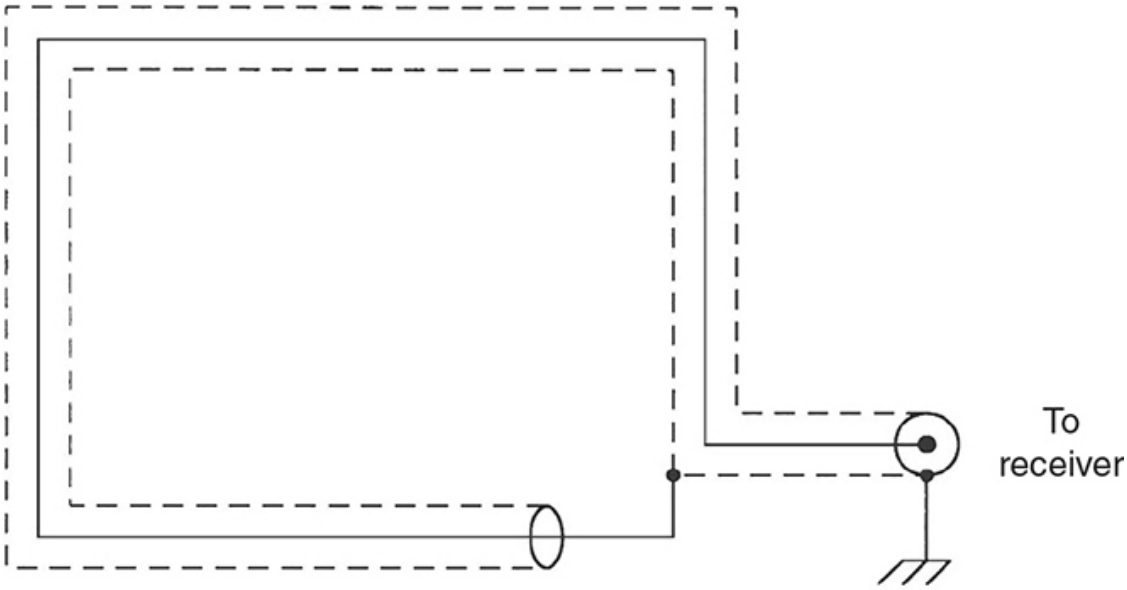


FIGURE 16.18 Coaxial-cable shielded loop.

16.3 The EWE Antenna

In 1995 Floyd Koonce, WA2VWL, published details of the EWE antenna, which provides MF and low HF receiving directivity in a much smaller footprint than a Beverage antenna. The acronym EWE is not only descriptive (Earth-Wire-Earth) but a play on words as well, since the antenna has the shape of an inverted "U."

The EWE of Fig. 16.19 is a two-element phased array of short verticals. Like the Beverage, the EWE is a traveling wave antenna. Directivity is a result of phase and amplitude differences between the two vertical segments. The phase and amplitude differences are determined by the length of the horizontal wire connecting the vertical runs and the value of the termination resistor.

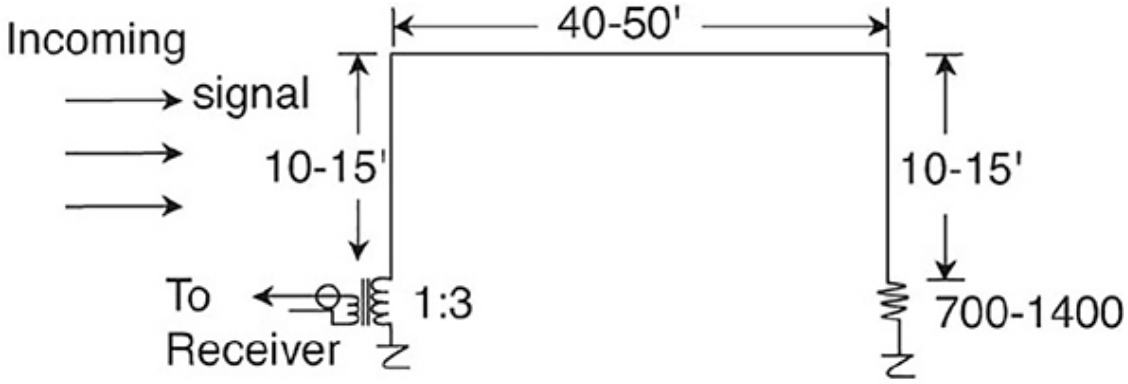


FIGURE 16.19 EWE antenna.

Unlike the Beverage, the EWE receives from the end of the antenna opposite the termination resistor. Because the antenna is not a resonant device, it is relatively broadband in operation. The optimal value of termination resistance varies with frequency and with the exact characteristics of the ground beneath the antenna. The value of the termination is set by tuning for a null off the back of the antenna by listening to stations in the AM broadcast band or on 160 m. Typical values for the termination resistor are in the 700- to 900-Ω range. The input impedance is somewhat lower and is a reasonably good match to 50-Ω coax and receiver inputs via a 9:1 transformer.

Although the EWE requires a much smaller space than a Beverage antenna, to achieve its full potential it needs to be kept away from other conductors. This constraint can substantially increase the actual space required by the EWE.

16.4 Pennants, Flags, and the K9AY Loop

The appearance of the EWE inspired other experimentally inclined radio amateurs to develop additional low-band receiving antennas for small spaces. As is the case with the EWE, for those lacking the necessary space for a set of Beverage antennas, these loops provide an alternative, although those fortunate enough to have the space for both generally report better performance with Beverages.

Directivity of this family of loops is based on constructive and destructive combining of the E- and H-fields that constitute an arriving radio wave. An arriving E-field in the plane of the loop induces a voltage in the loop just as it would in any other short wire. The H-field at right angles to the loop plane induces a current in the loop that becomes a voltage across the terminating resistor. When these fields arrive from one direction, the two voltages sum; when arriving from the opposite direction, they subtract.

Figure 16.20 depicts the K9AY terminated loop. The exact shape of the loop is not critical, but the bottom leg must come near earth for the termination and transformer ground connections. The total length of wire in the loop is about 85 ft, and a single support only 25 ft tall is required. The feedpoint impedance is similar to that of a typical Beverage, so a 9:1 transformer is required. Like the EWE and unlike the Beverage, the direction of maximum signal pickup is opposite the terminated end.

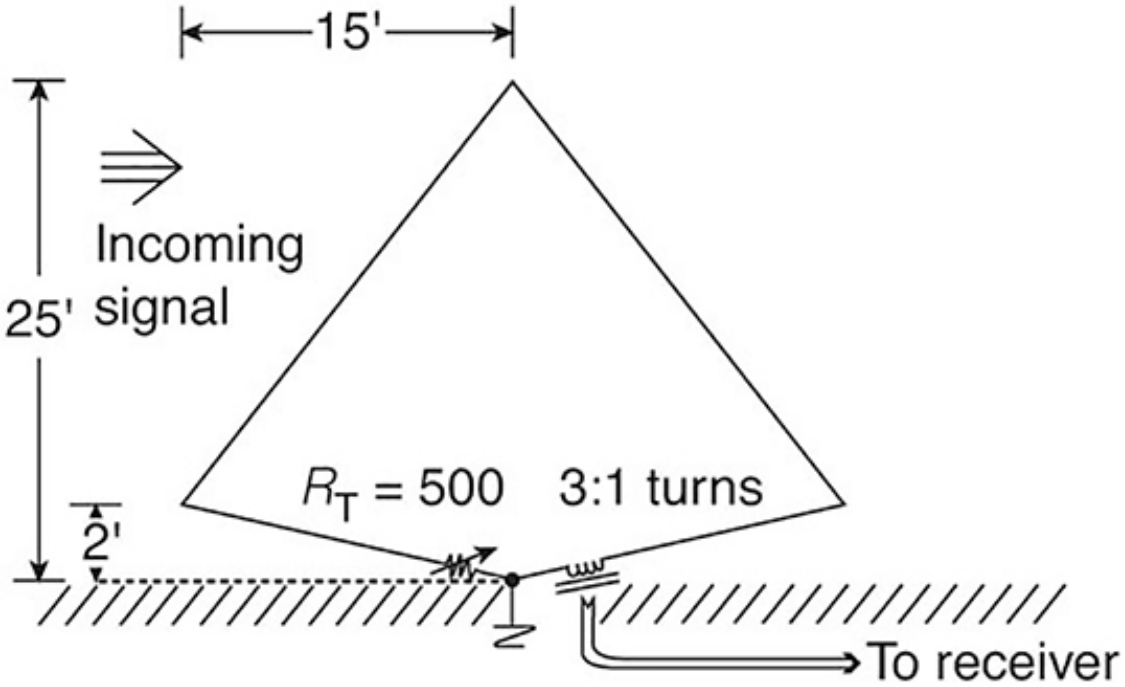


FIGURE 16.20 K9AY loop.

The pattern of the K9AY loop is unidirectional. By using relays to switch the termination and feedpoint, one loop can be switched

between two opposing directions. Two loops, at right angles to each other and similarly switched, can provide four directions from a single fixed mast.

Like the other members of the pennant and flag family, the K9AY loop is sensitive to detuning from nearby conductors. As a result, some of its impressive space savings are not always fully realizable if the user also has transmitting antennas and towers nearby. Proper operation of the K9AY depends on the value of the termination resistor, the operating frequency, and the characteristics of the ground beneath the antenna. Some implementations use a remotely controlled termination resistor to optimize the depth of the null to the rear as the received frequency is changed and to compensate for seasonal variations in soil characteristics.

In many operating scenarios, the value of this antenna is likely to be more for its null off the rear than for any forward directivity it exhibits. This is useful for weak-signal reception on the 160- and 80-m amateur bands or when broadcast band DXing.

The pennant (Fig. 16.21) was developed by Jose Mata, EA3VY, and Earl Cunningham, K6SE, to overcome the EWE's sensitivity to the characteristics of the ground. The resulting pennant design exhibits stability of pattern and performance over a wide variety of grounds and can be raised on a mast to a wide range of heights. The favored design sets the dimensions and the terminating resistor value to arrive at a purely resistive feedpoint impedance for optimal matching with a 9:1 transformer.

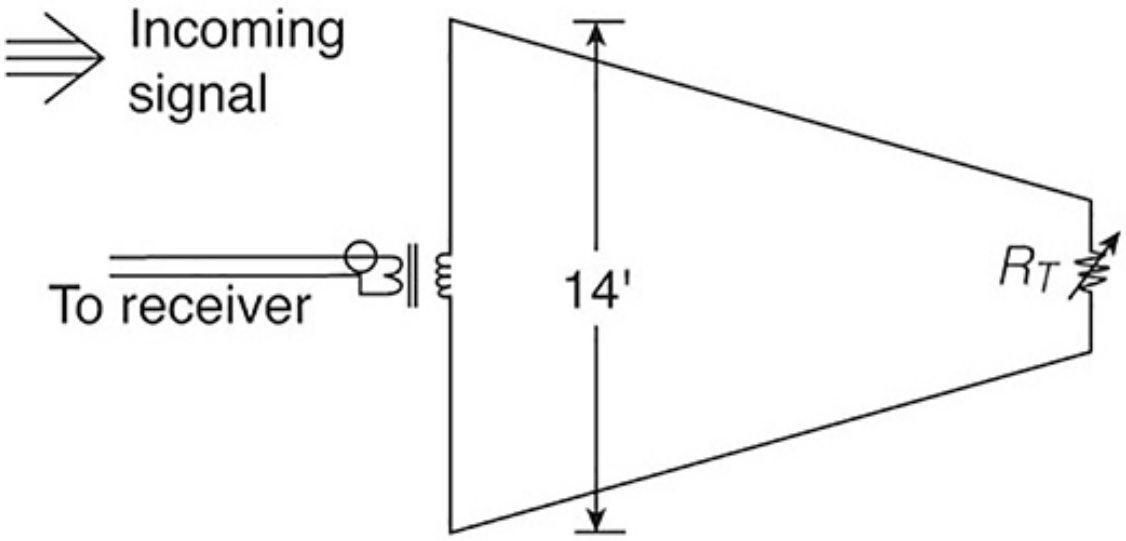


FIGURE 16.21 Pennant antenna.

Just as the pennant is a triangular receiving antenna, the flag (Fig. 16.22) is a rectangular one. Because the capture area of a flag is somewhat larger than that of a pennant, the flag antenna provides somewhat greater output voltage to the feedline for a given signal than does the pennant.

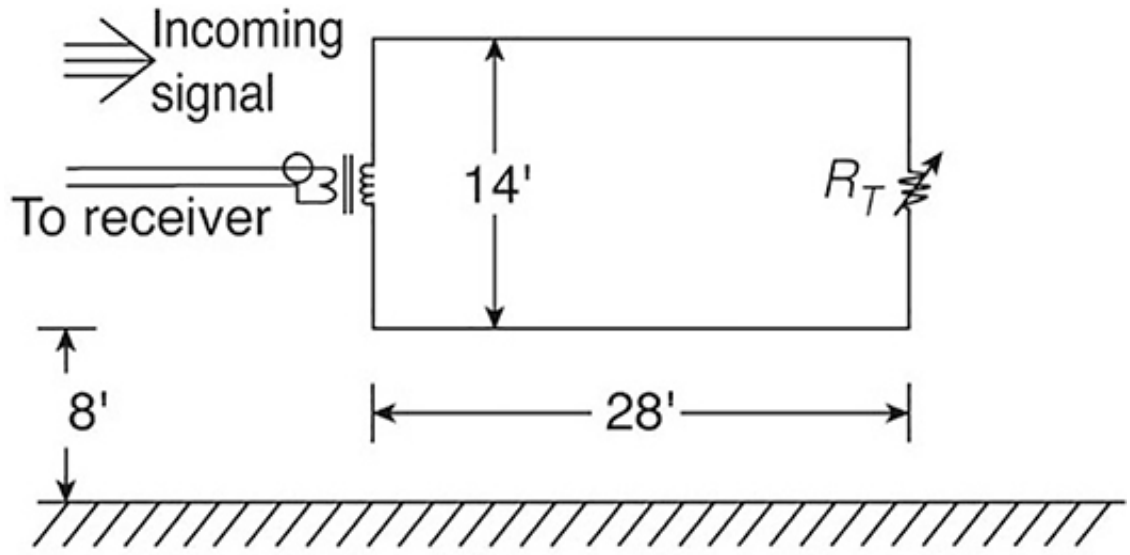


FIGURE 16.22 Flag antenna.

Unlike the Beverage, these antennas receive away from the terminating resistor. They are mounted on or supported by nonmetallic members (such as PVC pipe) and should not be located near metallic structures, especially those having dimensions in the vertical plane that are a sizeable fraction of a wavelength at LF through lower HF frequencies, unless steps are taken to detune those structures at the frequencies of interest while receiving. Many amateurs and BCLs have employed multiple phased pennants and flags to obtain even greater improvements in received signal-to-noise ratio.

16.5 Longwire Antennas

No discussion of HF receiving antennas would be complete without mention of the long-wire antenna. This is often the first external antenna used by a broadcast band DXer or SWL. Its primary attribute is simplicity. Its disadvantage is that it has a pattern that is unpredictable or shifts from broadside to axial with increasing frequency. If it is at least 100 ft long, this antenna can provide hours of enjoyable listening to anything from VLF airport beacons through the AM broadcast band and on up into the amateur and international shortwave broadcast bands. The longwire can be used with receivers that are not designed with an external connection. Simply bring the end of the longwire near the receiver, perhaps wrapping it around the receiver enclosure a few turns.

The longwire has little capability to reject unwanted signals. Strong off-frequency signals can overload a receiver unless some form of selectivity is applied at its front end. Even a high-end receiver will benefit from having a low-power antenna tuner (ATU) inserted in the path between it and a longwire antenna. Because the longwire enters the structure where the receiver is located, it picks up interference from local radiators and electronic devices as well.

The longwire antenna is covered at length in Chap. 10 as a transmitting antenna, but some of the installation considerations

there are unnecessary for a receive-only installation. When used for transmitting a longwire must be operated against a good RF ground system, but if used only for receiving, the grounds associated with the receiver’s power cord are probably sufficient at all but the highest frequencies.

CHAPTER 17
Hidden and Limited-Space Antennas

One of the most significant impediments to amateur radio operators and short-wave listeners is the space available for their antennas. City ordinances and neighborhood covenants may also limit the size and location of antenna installations. Here we will examine some of the alternatives available to those readers who have either a limited-space situation (such as a small city lot) or are unable to move out of a subdivision where there are prohibitions against outdoor antennas. The suggestions contained in this chapter are not universal and can be adapted in creative ways to unique situations.

17.1 Hidden Antennas

A hidden antenna is one that is either shielded from view or disguised as something else. We also include in this category antennas that are in semiopen view but which are not obvious (except to the trained and diligent eye).

The basic dipole antenna can be hidden in various ways. A dipole can be located in an attic, which is normally the highest part of a structure. The ends of the dipole can be bent to fit the antenna in a small space by dangling them or attaching them to angled rafters. A wire dipole can be fed at one end instead of in the middle or modified to be an end-fed Zepp or longwire. A good ground system is needed for best performance. To operate in multiple bands, traps can be used. A full-range antenna tuning unit (ATU) can further extend the operating range of the hidden dipole.

Antennas for the 10-, 13-, 15-, and (possibly) even 18- and 20-m bands will fit entirely inside the typical townhouse attic. This is also

true of the 11-m citizens band antenna. But what about the lower-frequency bands?

Figure 17.1 suggests a solution for the lower frequencies (specifically, 80, 40, and 30 m) in a townhouse or residence where the radio enthusiast has access to the highest interior space. The dipole can be installed with arms bent in conformity with the space available. In the pictured example, only one of many possible methods for accomplishing this job is shown. Each quarter-wavelength section is composed of two legs, *AB* and *CD*. Segments *B* and *C* are ideally the longest dimensions. Another method is to reverse the direction of one end leg—say, for example, *D*—and run it to the other corner of the building over the peak of the roof. Adjustments can be made to fit a uniquely shaped space.

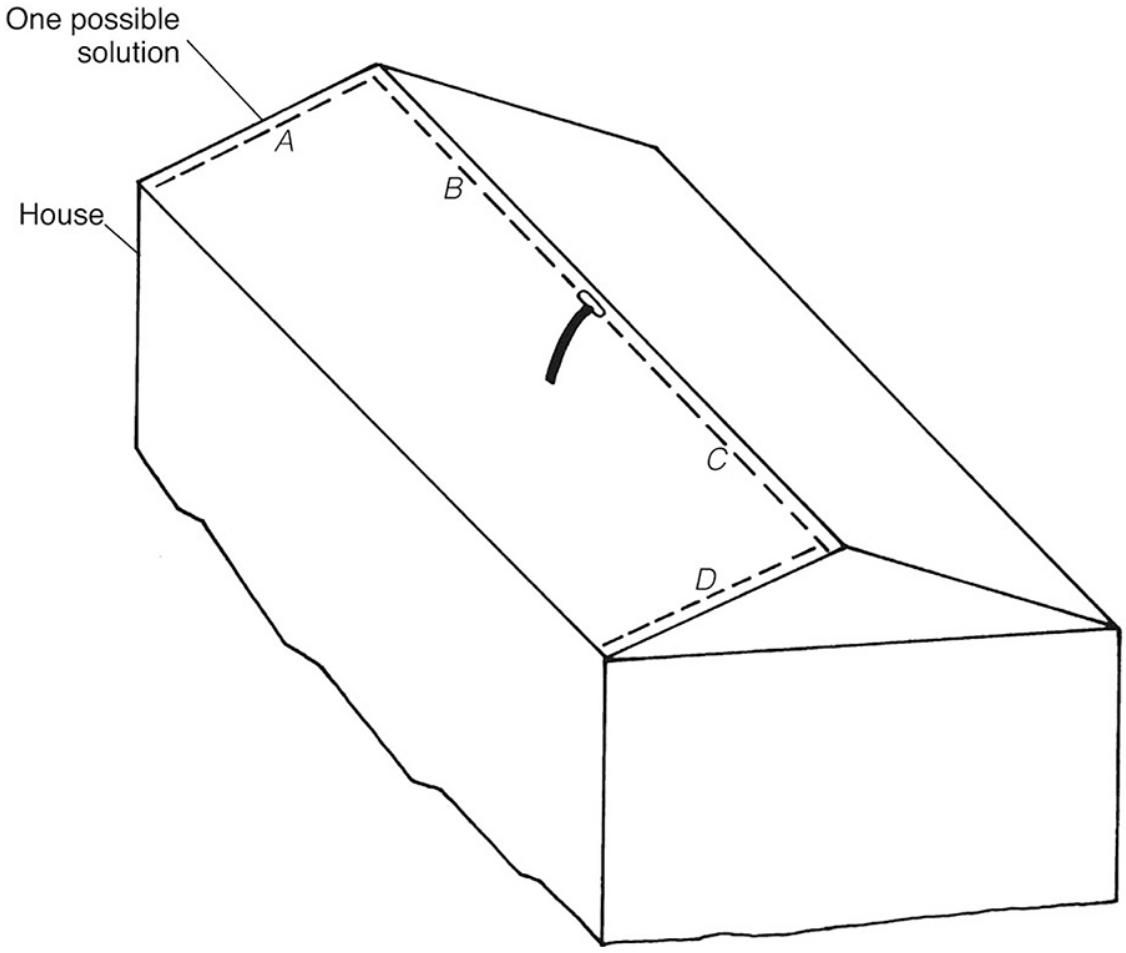


FIGURE 17.1 Installing dipole in attic.

Although wood rafters and studs are insulators, the wire used in the constrained dipole or other indoor antennas should be mounted on TV-type screw-in standoff insulators if any significant output power is employed.

Caution *Do not simply tape or staple the wire to the wooden underside of a roof. In transmitting antennas, voltages can be high enough not only at the ends but at many multiple locations along the wire to produce corona effects. The resulting arcing can easily start a fire.*

The constrained dipole of Fig. 17.1 generally will not perform as well as an ideal dipole. The pattern will be distorted and the feedpoint impedance will usually be different from 50 Ω or 73 Ω, so you will be required to use an antenna tuner of some kind.

An alternative for the attic antenna is the nonresonant loop shown in Fig. 17.2. Although presented as a top view, the loop can be installed in any configuration that is compatible with the available space. The best performance will be bidirectional when the loop is installed vertically. Again, use standoff insulators and insulated wire for the installation. Ideally, the giant loop is fed with parallel line and tuned with a balanced ATU, but it can also be fed with coaxial cable.

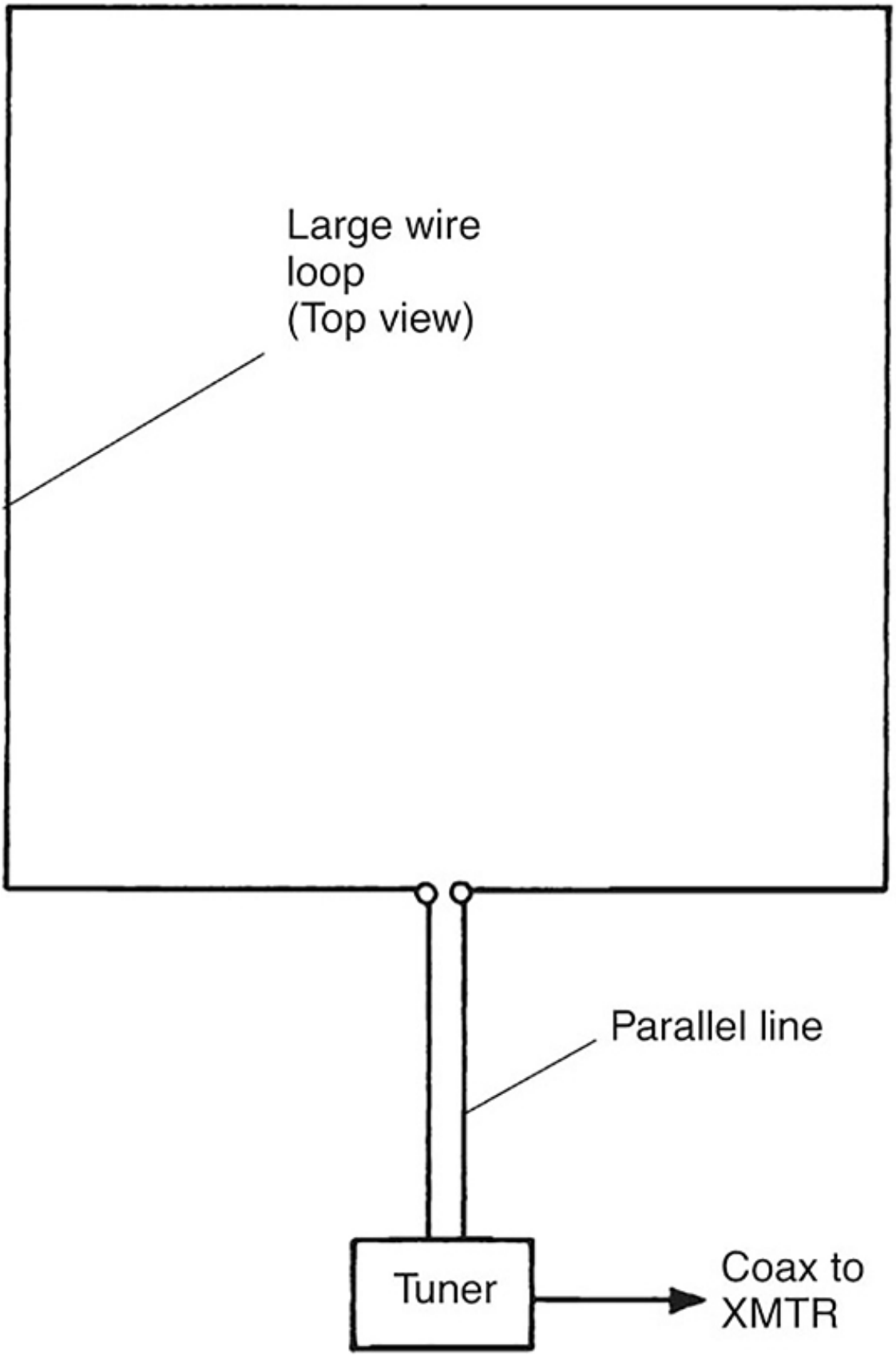

Large wire
loop
(Top view)
Parallel line
Tuner
Coax to
XMTR

FIGURE 17.2 Large loop antenna.

Indoor antennas are usually inferior receiving antennas because of their proximity to noisy power wiring in the walls of the structure and indoor electronics. If you are plagued by interference from a single direction, consider the use of a noise-canceling accessory, such as those made by ANC or MFJ. These devices pick up the noise on a sense antenna that is often nothing more than a short piece of wire and then change the phase of the detected noise signal so it cancels the noise on your main antenna. The noise canceler approach works best when the offending noise comes from a single source or a specific direction.

Hidden outdoor antennas are also possible. A brass or aluminum flagpole can be fed directly from an ATU. For single-band operation, especially on the higher frequencies, you can delta-feed the flagpole unobtrusively and call your flagpole a vertical antenna.

Figure 17.3 shows another method for creating a flagpole antenna. The pole is constructed from PVC plumbing pipe. Heavier grades of PVC pipe are self-supporting to heights of 16 to 20 ft. Figure 17.3 shows the use of a PVC flagpole in which a #12 (or #14) wire is hidden inside. This wire is the antenna radiator. For some frequencies the wire will be resonant, and for others it will be nonresonant. An ATU is used either at the base of the antenna or inside at the transmitter. If the wire is too long for resonance (as might happen in the higher bands), place a capacitor in series with the wire. Multiple settings may be required, so use a vacuum or multisection transmitting variable that has a total capacitance selectable to more than 1000 pF. If the antenna is too short for resonance, as will occur in the lower bands, insert an inductance in series with the line to "lengthen" it. Alternatively, use an L-section tuner at the feedpoint.

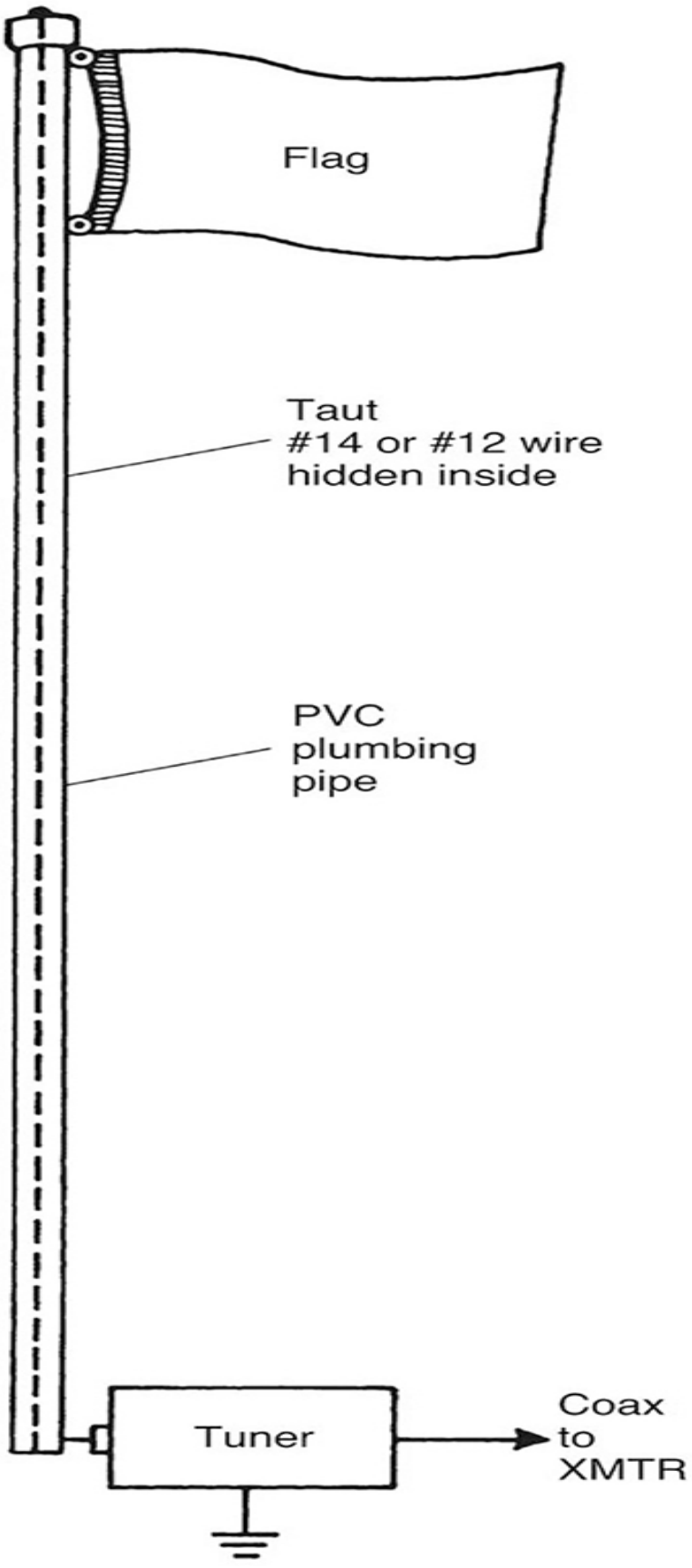

Flag
Taut
#14 or #12 wire
hidden inside
PVC
plumbing
pipe
Tuner
Coax
to
XMTR

FIGURE 17.3 Flagpole antenna made from thick-walled PVC pipe.

A good compromise is the use of a 16-ft length of "flagpole" pipe with a 16-ft wire embedded inside. The 16-ft wire is resonant at 20 m, so it will perform like a quarter-wave grounded vertical monopole antenna on that band. The tuner will accommodate frequencies above and below 20 m. A monopole will not work well at all without an appropriate ground or system of radials beneath it. Radials should be relatively easy to hide.

A further alternative is shown in Fig. 17.4. The wire radiator is replaced with a section of aluminum tubing. A wooden or plastic insert supports the aluminum tubing inside the PVC tubing. One way to make the support is to use a core bit in an electric drill to cut out a disk that fits snugly inside the PVC tubing. Rat out the center hole left by the core bit pilot to the outside diameter of the aluminum tubing. The support can be held in place with screws from the outside or simply glued in place. One advantage of using tubing instead of a wire is a modest increase in the bandwidth of the vertical.

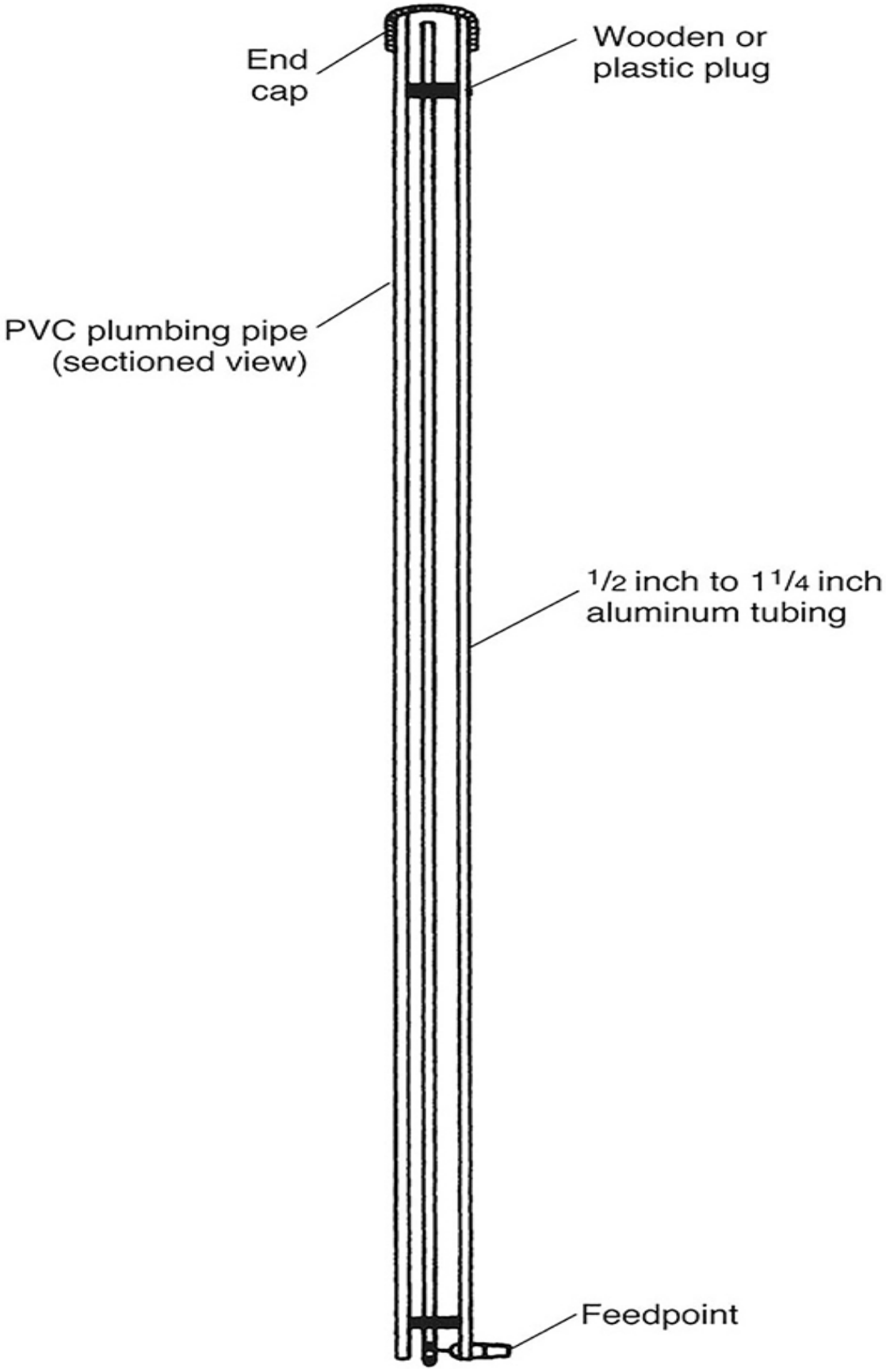

End cap
Wooden or plastic plug
PVC plumbing pipe (sectioned view)
1/2 inch to 1 1/4 inch aluminum tubing
Feedpoint

FIGURE 17.4 PVC flagpole hiding an aluminum tubing vertical.

Coming up with the right concept for operating with a hidden antenna can pose a serious challenge. But with some of these guidelines and a little creativity, you should be able to get on the air and enjoy your hobby.

17.2 Limited-Space Antennas

Many people live in situations where it is permissible to install an outdoor antenna, but it is not practical to install a full-size antenna. Here we examine some of the options open to those with limited space for amateur radio, citizens band (CB), or shortwave listener (SWL) antennas.

Once again we return to the simple dipole as the basis for our discussion. Figures 17.5 to 17.8 sketch several alternatives for installing an outdoor antenna in a limited space. In Fig. 17.5, the slanted dipole antenna uses the standard dipole configuration, but one end is connected to the high point of the building, while the other is anchored near the ground. The coaxial cable is connected to the midpoint of the antenna in the usual manner for regular dipoles. If the end of the dipole is within reach of people on the ground, someone may get a nasty RF burn if the antenna is touched while you are operating. Take precautions to keep people and pets away from any part of your antenna.

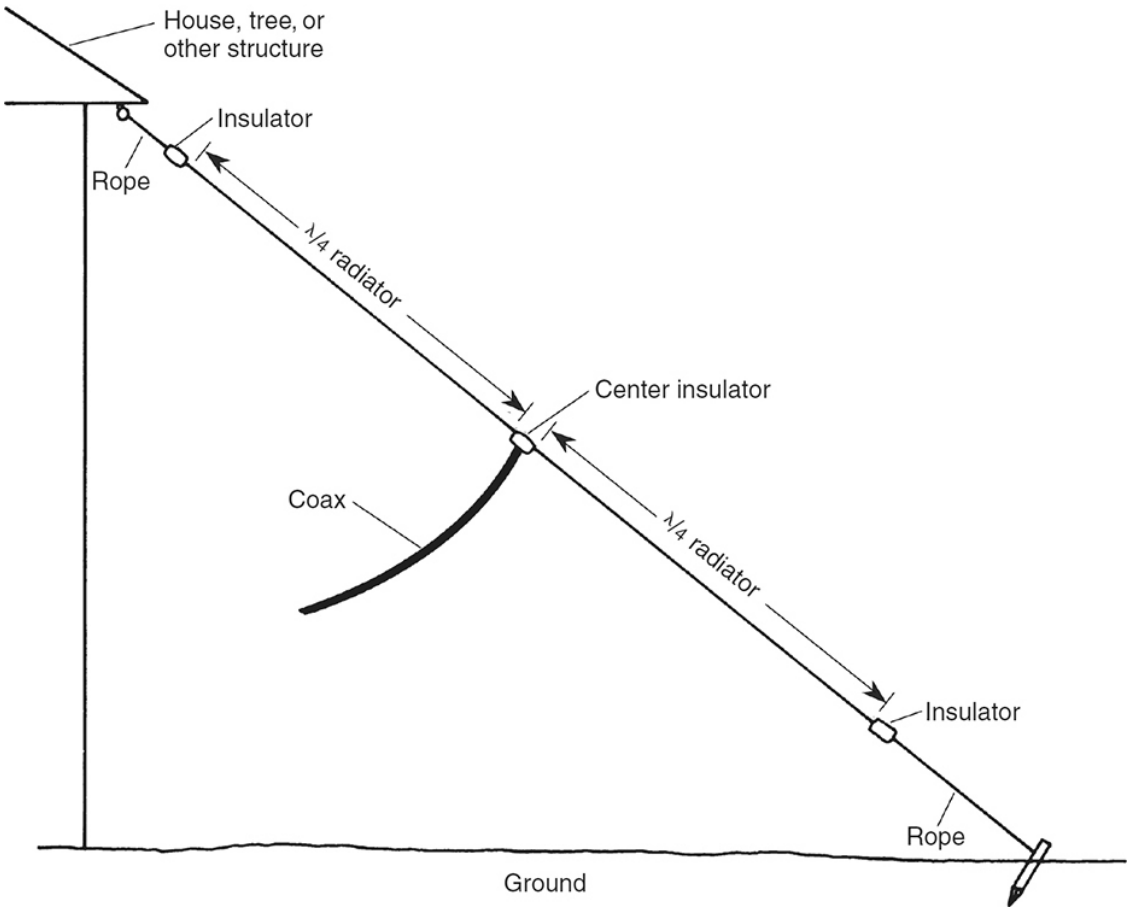


FIGURE 17.5 Sloping dipole.

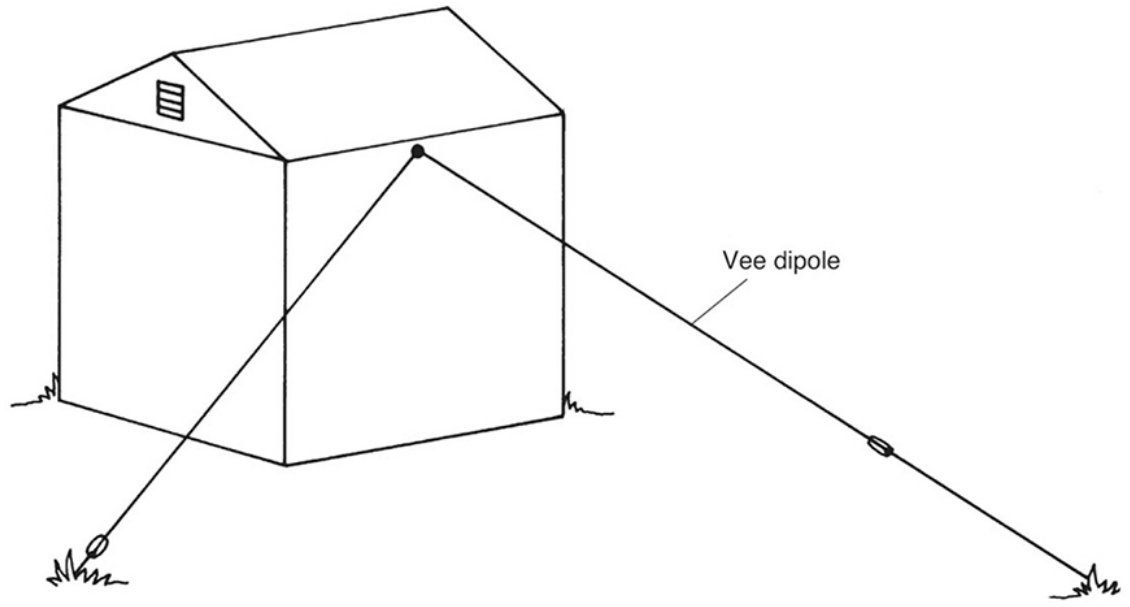


FIGURE 17.6 Vee dipole.

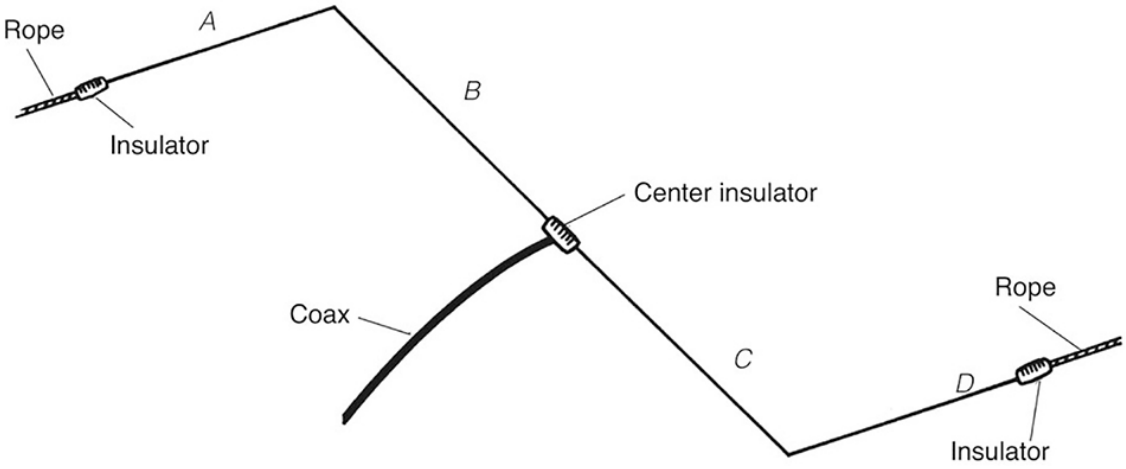


FIGURE 17.7 Bent dipole.

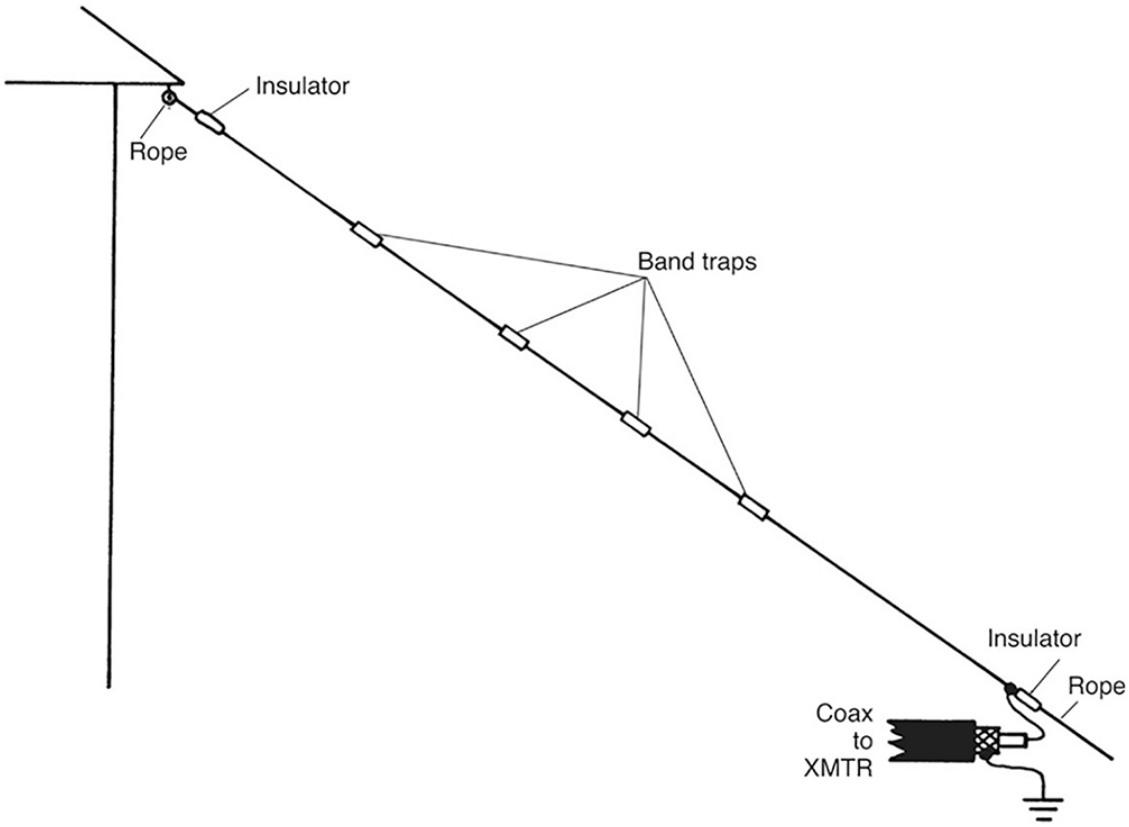


FIGURE 17.8 Sloping trap vertical.

Another method is the vee dipole shown in Fig. 17.6. In a regular dipole installation, the ends of the antenna are along the same axis. (In other words, they form an angle of 180 degrees.) In Fig. 17.6,

the angle between the elements is less than 180 degrees but greater than 90 degrees. In some cases we might need to bend the elements partway along their length, rather than installing them in a vee shape. Figure 17.7 shows an angled or bent dipole with four segments.

The length needed for resonance will vary somewhat from the standard dipole length and the feedpoint impedance is very likely to be something other than 73 Ω. The pattern may be distorted by any house wiring, aluminum siding, or other nearby conducting materials with at least one dimension roughly comparable to $\lambda/4$ or more. Although these antennas may not work quite as well or as predictably as a high dipole clear of nearby objects, the performance can be surprisingly good.

Another limited-space wire antenna is the half-sloper shown in Fig. 17.8. Although single-band versions are often seen, the example shown is a multiband version. Resonant traps result in relatively low SWR across multiple frequency ranges, allowing the use of coaxial cable for the feedline and potentially eliminating the need for an ATU. In contrast to center-fed dipoles, this antenna is a sloping monopole that is strongly dependent upon a good RF ground at the feedpoint for reasonable transmit efficiency.

Figure 17.9 shows another antenna that's useful for limited-space situations on the lower frequencies. Although it is easily constructed from low-cost materials, the antenna is also available commercially. It consists of a 4- to 16-ft section of aluminum or copper tubing. Some of the commercial antennas use telescoping tubing that can be carried in luggage. As with the longwire, this window sill antenna is tuned to resonance with an L-section coupler. A good ground or radial system will improve performance.

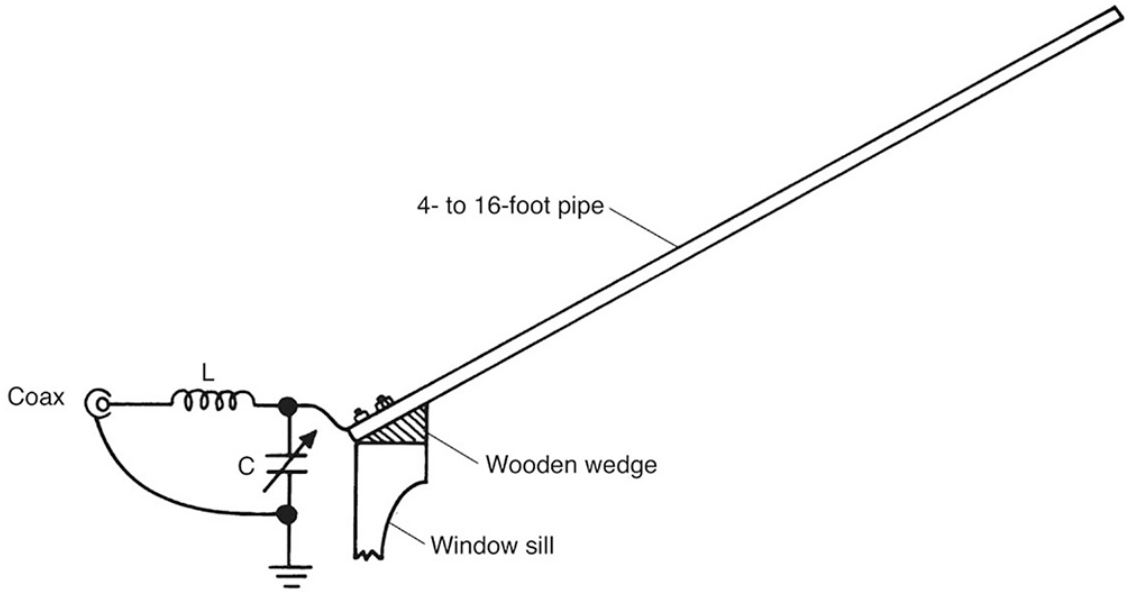


FIGURE 17.9 Window sill antenna.

Figure 17.10 shows the use of a mobile antenna for a fixed or portable location. Simply mount a mobile whip, such as the *Hustler*, on the window sill or other convenient mount. Since this is a vertical monopole, an adequate ground system is essential. A monopole should have a counterpoise or elevated ground consisting of at least two or three radials per band.

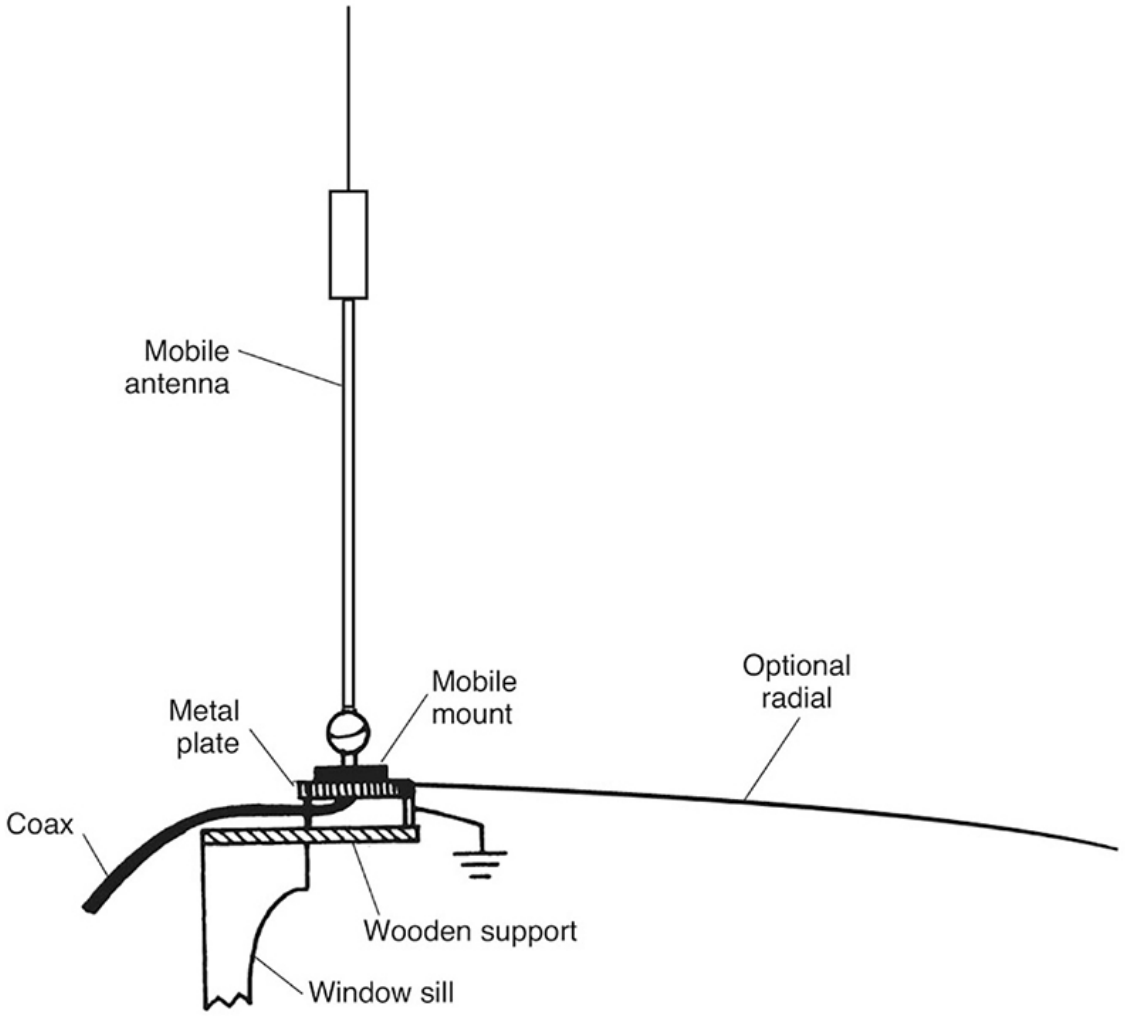


FIGURE 17.10 Use of a mobile antenna as a window sill antenna.

A directional rotatable dipole is shown in Fig. 17.11. This antenna is made from a pair of commercial mobile antennas connected back-to-back on a horizontal length of

$$1\text{-in} \times 2\text{-in}$$

lumber and fed in the center with a coaxial cable or open-wire line. The two antennas can be clamped to the wood center support with hose clamps or strong cable ties. The tips of the mobile antenna determine where each half-element is resonant, and you must be sure to adjust both sides of the antenna to the same length. On HF these are extremely short dipoles for the frequencies involved, so their tuning will be sharp, the useable bandwidth narrow, and the

efficiency low. One advantage to using them in a horizontally polarized configuration is that no RF ground system is required. Another advantage, particularly if you live in a high-rise condo or apartment building with your own balcony and can lay the completed assembly out on a nonmetallic table or other support, is they then pick up as much as 6 dB ground reflection gain at a decent height. Try it—the results may surprise you!

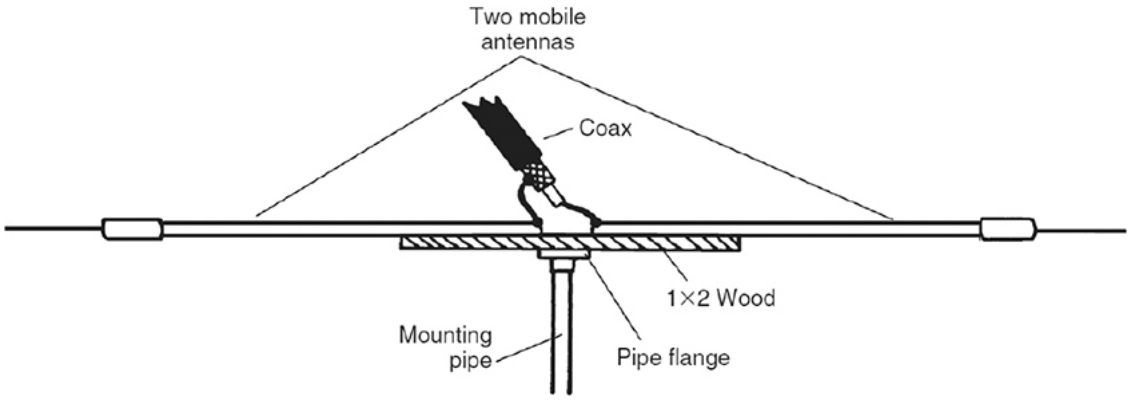


FIGURE 17.11 Use of two mobile antennas as a rotatable dipole.

Two examples of helically wound antennas are shown in Figs. 17.12 and 17.13. An insulating mast is wound with a half-wavelength of antenna wire. Mobile HF whips from Mosley Antennas utilize this technique. The overall length of the antenna is considerably less than a half-wavelength except at the highest frequencies. In order to dissipate the high voltages that tend to build up at the ends of these antennas when operated at higher power levels, a capacitance hat is often used. These hats can be disks (pie tins work well) or conducting rods about 16 to 24 in long. The version shown in Fig. 17.12 is a vertical monopole antenna, and like other such verticals, it must be installed over either good ground or a counterpoise ground. The version shown in Fig. 17.13 is a horizontally polarized dipole. Helical whips are somewhat less efficient than other mobile loading coil schemes because the helical coil uses much more wire than a lumped-component coil to achieve

the same loading. The extra wire equates to added resistive power loss, which increases bandwidth but lowers efficiency.

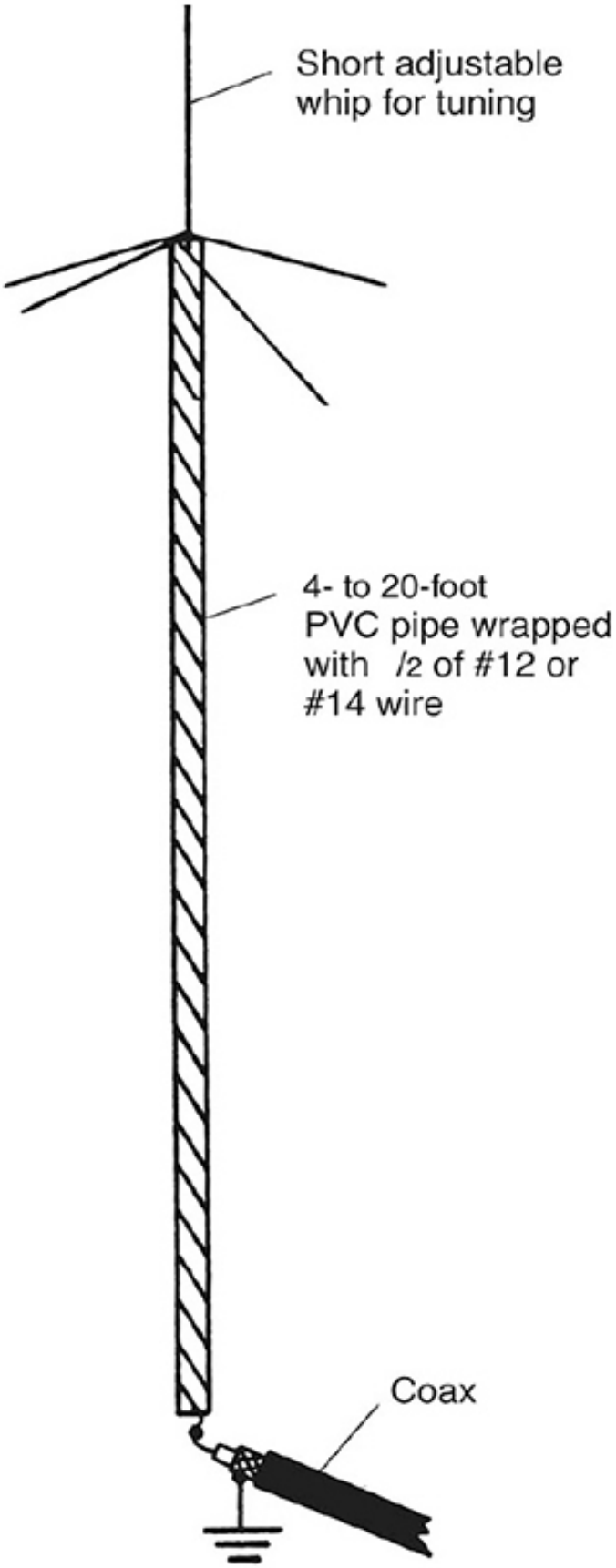

Short adjustable
whip for tuning
4- to 20-foot
PVC pipe wrapped
with /2 of #12 or
#14 wire
Coax

FIGURE 17.12 Helically wound vertical antenna.

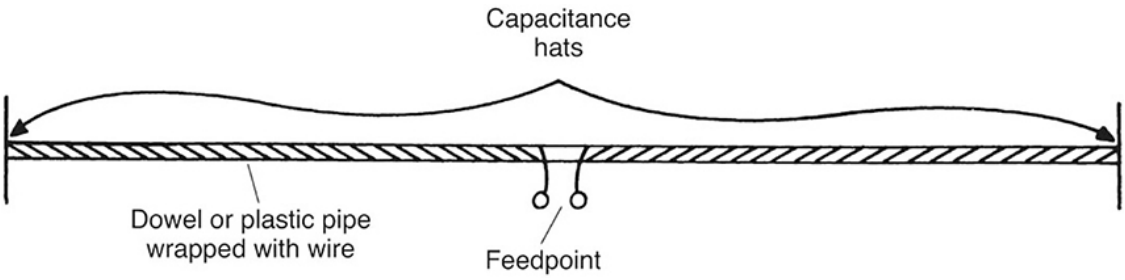


FIGURE 17.13 Helically wound dipole antenna.

CHAPTER 18
Mobile and Marine Antennas

From the earliest days of wireless, radio communication equipment has been put in moving vehicles. Over time, mobile radio has moved higher in frequency for practical reasons. The higher the frequency, the shorter the wavelength and the smaller the antenna. On the 11-m citizens band, a quarter-wavelength whip antenna is 102 in long, and for the 10-m amateur band only 96 in long. At VHF and above, full
$\lambda/4$
mobile antennas are practical. Partially as a result of this, much mobile activity takes place in the VHF and UHF region. Mobile operation in the amateur 144-, 220-, and 440-MHz bands is popular because of several factors, not the least of which is the small size and cost of vehicle-mounted verticals (or whips) for these bands.

18.1 Mobile Antennas for VHF and UHF

At 2 m and above, the roof of a car has enough nearly flat surface area to act very much like an ideal ground plane. The antenna is ideally located near the middle of the roof, and the shield, or grounded, conductor of the coaxial feedline should have a low-resistance connection to the roof at the base of the antenna. If you are comfortable with snaking RG-58 or RG-8X cable between the roof and the liner of your vehicle and know how to install an old-fashioned chassis-mount SO-239 connector with a pilot hole and chassis punch with gasket or caulk to seal against rain, you can install a decent 2-m mobile antenna.

If the car roof is not suitable due to limited clearance or other issues, the next best location is usually the center of the trunk lid. Collinear whips for VHF and UHF are relatively inexpensive and can

bring at least one element of the collinear up above the roofline for 360-degree coverage. Locating the mobile antenna at the rear of the vehicle can also reduce noise pickup from electronics in the car.

If needed, a VHF or UHF whip can be mounted on a fender or along the side or rear of the hood, in the gap between the hood and other vehicle body sheet metal sections. Whips can be mounted to the oversize outside rearview mirror brackets found on larger vehicles, though the ground there is poorer than for antennas mounted over flatter surfaces.

For vehicles with nonmetallic bodies, there are special difficulties. Low whips, such as bumper-mounted ones, will encounter less loss of signal from massive pieces of metal near the whip. The frame and drivetrain members may also provide something of a ground plane.

If all else fails, bumper mounts, side fender mounts, and trailer hitch mounts can be used, even though the typical VHF or UHF whip (even the collinear) will be well below the roofline of the vehicle. At HF, these options are quite common.

18.2 Mobile Antennas for HF and MF

Radiating a strong signal from a mobile installation is a frequency-dependent problem. Many vehicles provide a satisfactory ground plane and sufficient vertical clearance for a
$\lambda/4$
or taller whip on 2 m and above. But for HF, a strong mobile signal can be challenging. Even in the presence of a perfect ground plane, most mobile vertical monopoles below 30 MHz are such a small fraction of
$\lambda/4$
that radiation efficiency and radiation resistance are extremely low by comparison to fixed station antennas.

Electrically short antennas have several practical issues. The input impedance is highly reactive, requiring a matching network that limits operation to a narrow bandwidth. Short antennas

generally have a low radiation efficiency. The antenna's radiation resistance is small, meaning that a high current must be input to the antenna to achieve high radiated power. High input current means large ohmic losses, and much of the power input to the antenna is lost to heat.

To overcome these problems, HF or MF mobile installations require careful design. A top hat and center loading coil can be used to increase the effective antenna length of a short monopole. This increases the radiation resistance and reduces the reactance.

As the operating frequency is lowered, the amount of metal needed to form an adequate ground plane increases. Even at 6 m, the roof of most vehicles is not large enough to approach. As the operating frequency is lowered, ground losses in the mobile antenna system will increase. As with fixed or base station vertical monopoles, few antenna design and construction details are as important as those of the ground system directly beneath the vertical. In most cases, the vehicle body is an oddly shaped mass of metal, and the most efficient mobile installations attempt to optimize antenna performance despite close proximity to that conductor.

Reduce RF losses in the vehicle's metallic mass to the smallest extent possible. Because the radiation resistance of the mobile monopole is so low, the quality of the ground beneath is more important than it is for a fixed vertical. Locate the antenna so as to keep as much of the metal as possible below the primary radiating portion of the antenna. A superior mobile installation will have as many major metal components of the vehicle as possible electrically bonded together, as well as to the ground (shield) side of the transmission line at the base of the mobile antenna. Similarly, losses in connections, connectors, transmission lines, antenna tuning units, and even the monopole itself can eat up most of the RF energy from the transmitter. Pay attention to every link in the system.

Don't forget the receiving side of the equation. Locate the mobile antenna as far as possible from vehicular electronics. Route power and RF cables connecting your equipment to the antenna, the

vehicle battery, and any other devices as far from other vehicular cabling as possible. Place common-mode chokes on all cables entering and exiting the receiver or transceiver. Many other techniques are available for minimizing RFI.

Effective high-frequency (HF) mobile operation requires different antennas than those commonly installed for VHF or UHF. Quarter-wavelength antennas are feasible on only the 10-, 11-, and 12-m HF bands. At the 21-MHz (15-m) band, the antenna size must be approximately 11 ft long, much too long for practical mobile operation. Height limits imposed by utility lines, overpasses, parking garages, and tollbooths limit antenna length to about 9 ft. For many amateurs the 8-ft whip is most popular because it is resonant on 10 m. At all lower frequencies, the 8-ft whip is too short and it becomes capacitive, requiring an equal inductive reactance to cancel the capacitive reactance of the antenna at the chosen operating frequency.

As the frequency drops, not only is the radiating monopole too short for the frequency, but the ground plane beneath it is becoming less and less effective. When the effects of imperfect ground planes and short vertical radiating elements are combined, an HF mobile installation may have overall efficiencies between 1 percent and 10 percent of those of a conventional quarter-wavelength radiator with an extensive ground plane.

18.3 Types of Short Mobile Antennas

Given the impossibility of having a
$\lambda/4$
whip on the HF and MF bands while mobile, the most common compromise is to add a combination of discrete and/or distributed components to the mobile whip in hopes of increasing the feedpoint resistance and eliminating or at least minimizing the feedpoint reactance.

Figure 18.1 shows three basic configurations of coil-loaded HF antennas for frequencies lower than the natural resonant frequency

of the antenna. The purpose of the coil is to cancel most, if not all, of the capacitive reactance of the electrically short antenna at the operating frequency. The antenna is series-fed with coaxial cable from the base. Point A is connected to the coaxial cable center conductor and point B is connected to the shield and the car body, which serves as an imperfect ground plane.

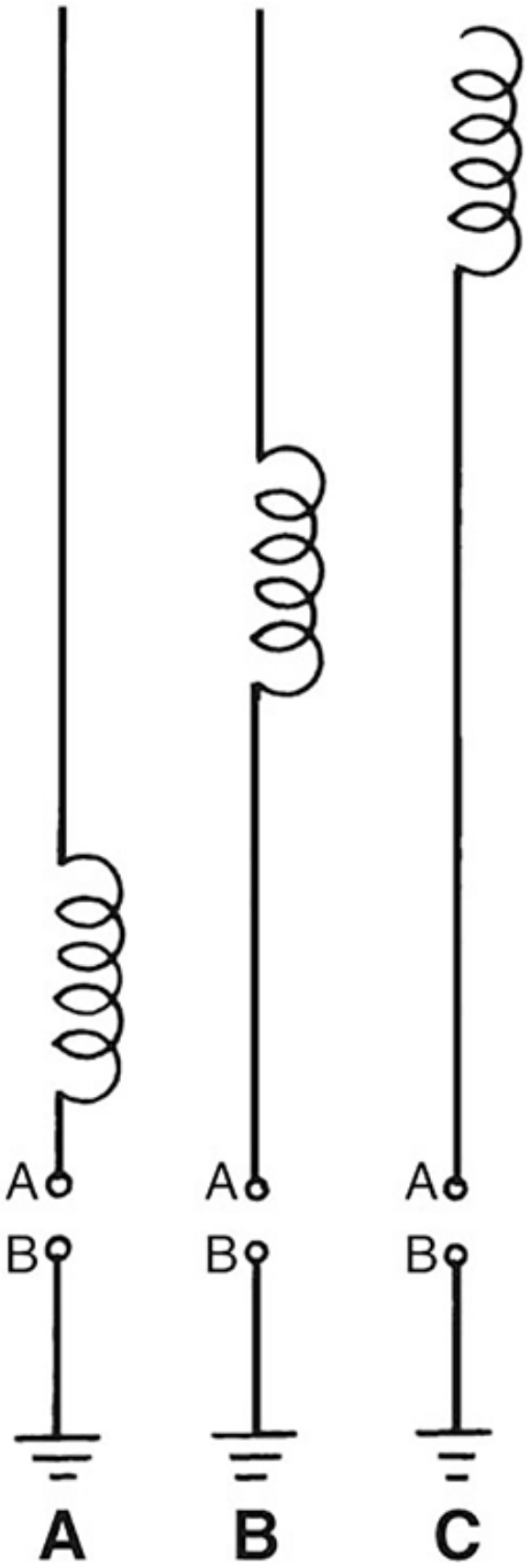

A
B
A
B
A
B
A
B
C

FIGURE 18.1 Loaded mobile antennas.

The system shown in Fig. 18.1A is base-loaded. Although convenient, simple to build, and the easiest to support mechanically, base-loaded verticals are among the least efficient radiators around and should be avoided if at all possible. Their inefficiency is a direct result of the high-current portion of the antenna element being replaced by a lumped-component loading coil, an approach we have eschewed in earlier chapters.

Figure 18.1B depicts a center-loaded whip, with improved current distribution. This configuration is the most common among commercially available HF mobile antennas, although the coil may be high enough and/or large enough to require special support.

Finally, we see the top-loaded coil system in Fig. 18.1C. When accompanied by a capacitive top hat, this is the most efficient configuration, as has been proved time and time again at numerous mobile contests at hamfests around the world. It also is the most top-heavy of the three configurations, requiring the largest coil, the strongest mast, and sometimes even guy wires or guy rods high on the mast from anchors near the top of the vehicle.

A modified version of the coil-loaded HF mobile antenna is shown in Fig. 18.2. The loading inductance is divided between two coils, L_1 and L_2. Coil L_1 is adjustable with respect to the antenna, while L_2 is fixed. Coil L_2 is tapped to match the impedance of the antenna to the characteristic impedance of the coaxial-cable transmission line. When tuning this antenna, two instruments are useful: a field strength meter and a VSWR meter or network analyzer. Placed at

least a few wavelengths from the mobile installation, the field strength meter indicates the signal radiated from the antenna, while the VSWR meter or network analyzer helps determine the state of the impedance match.

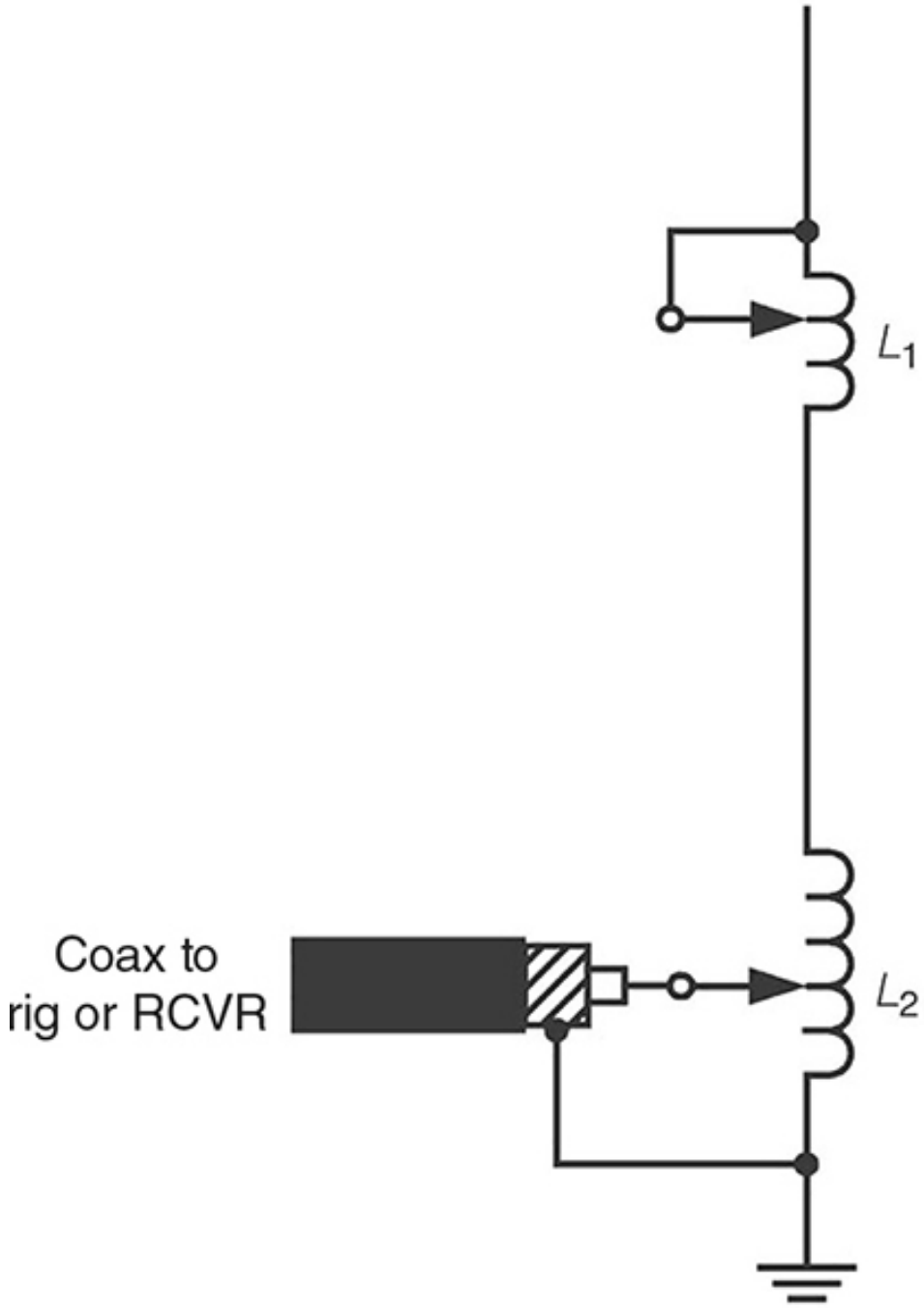


FIGURE 18.2 Practical mobile antenna for HF bands below 10 m; *L*1 resonates, *L*2 matches impedance.

Variations on this theme are found in Figs. 18.3 and 18.4. In the first variation, a helical whip, similar to the family of mobile whips available from Kenwood, is shown in Fig. 18.3. The inductor is distributed along the length of a fiberglass antenna rod. Rather than being a cylindrical metal tube or rod, the conductor is a wire helically wound on a fiberglass shaft and encased in shrink-wrap sleeving to minimize abrasion damage. An adjustable-length tip section tunes the antenna to resonance if the mobile transmitter or transceiver does not have a built-in ATU. The length of this tip should be set with the help of a VSWR meter. Fixed station versions of this type of antenna are popular on the 11-m citizens band, where they may be as short as 30 in or as long as 48 in.

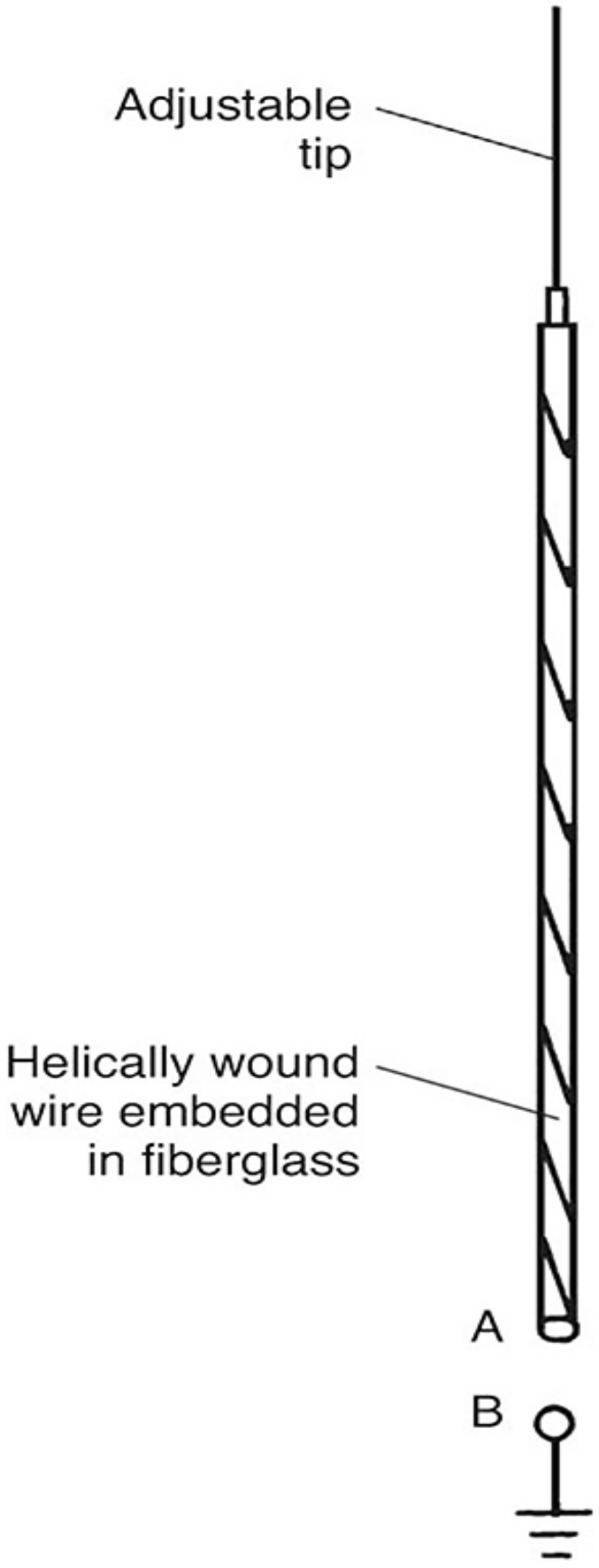

Adjustable
tip
Helically wound
wire embedded
in fiberglass
A
B

FIGURE 18.3 Helically wound mobile antenna.

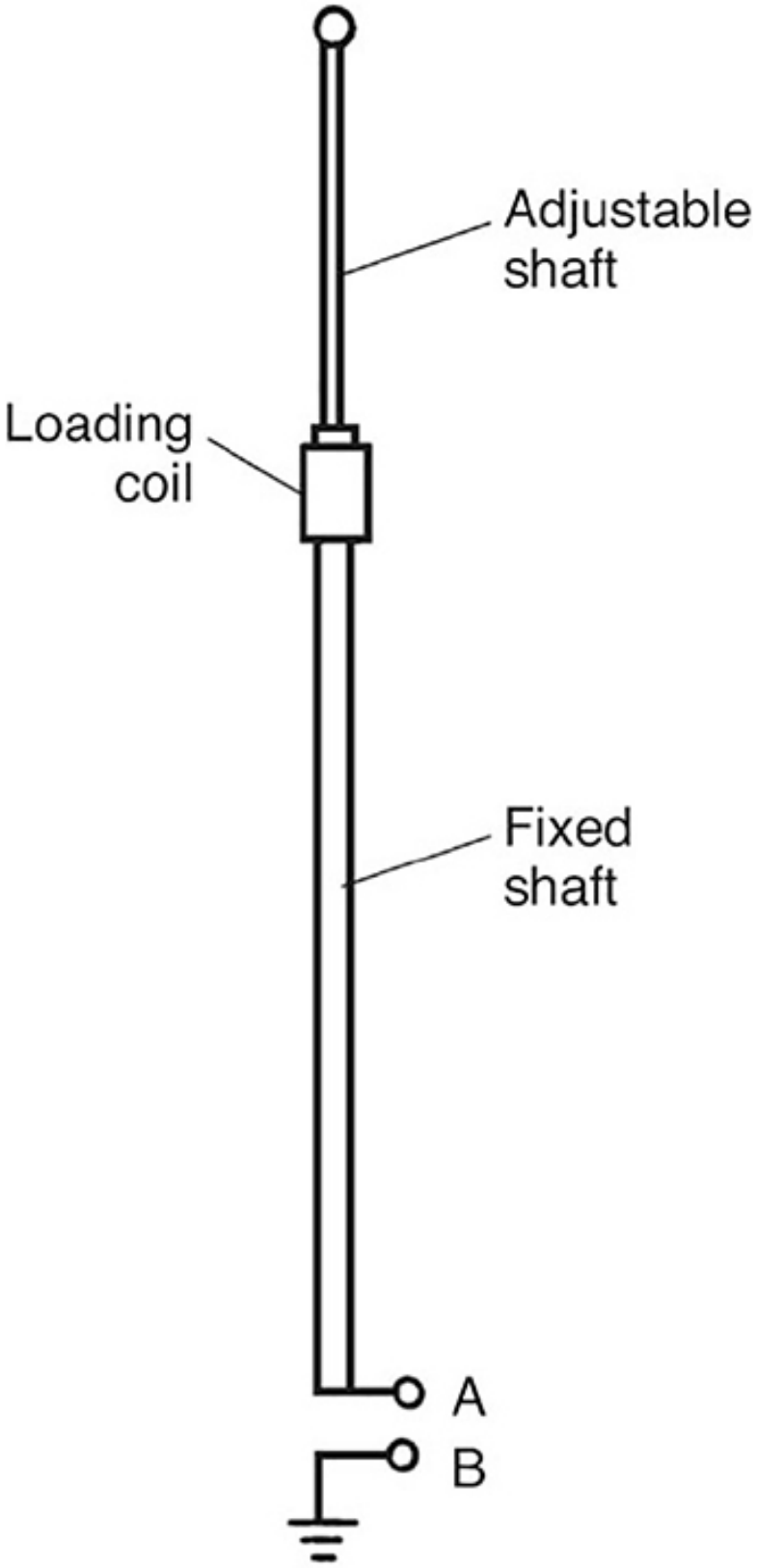

Adjustable
shaft
Loading
coil
Fixed
shaft
A
B

FIGURE 18.4 Top-loaded mobile antenna common in amateur radio.

Another variant is shown in Fig. 18.4. In this case, the lower end of the radiator consists of a metal tube topped with a loading coil. An adjustable shaft at the top end is used to tune the antenna to resonance. Commercially manufactured forms of this approach are easily found; the New-Tronics Hustler product line includes a family of lower rods, couplers, mounts, and adapters, and features resonator coils for both low and high power. The fixed lower rod is used on all bands, while the coil and adjustable stinger form a separate "resonator" for each HF band. Multiband operation can be accomplished by using a resonator mounting bracket such as the three-way unit shown in Fig. 18.5. (But don't use it while moving; the stock *Hustler* bracket is aluminum, which can quickly fatigue and crack when subjected to buffeting while under way.) With no resonators at all atop it, the MO-1 or MO-2 fold-over shaft can be used by itself on 6 m, especially if the mobile HF rig has an internal antenna tuner (example: the Kenwood TS-480S/AT).

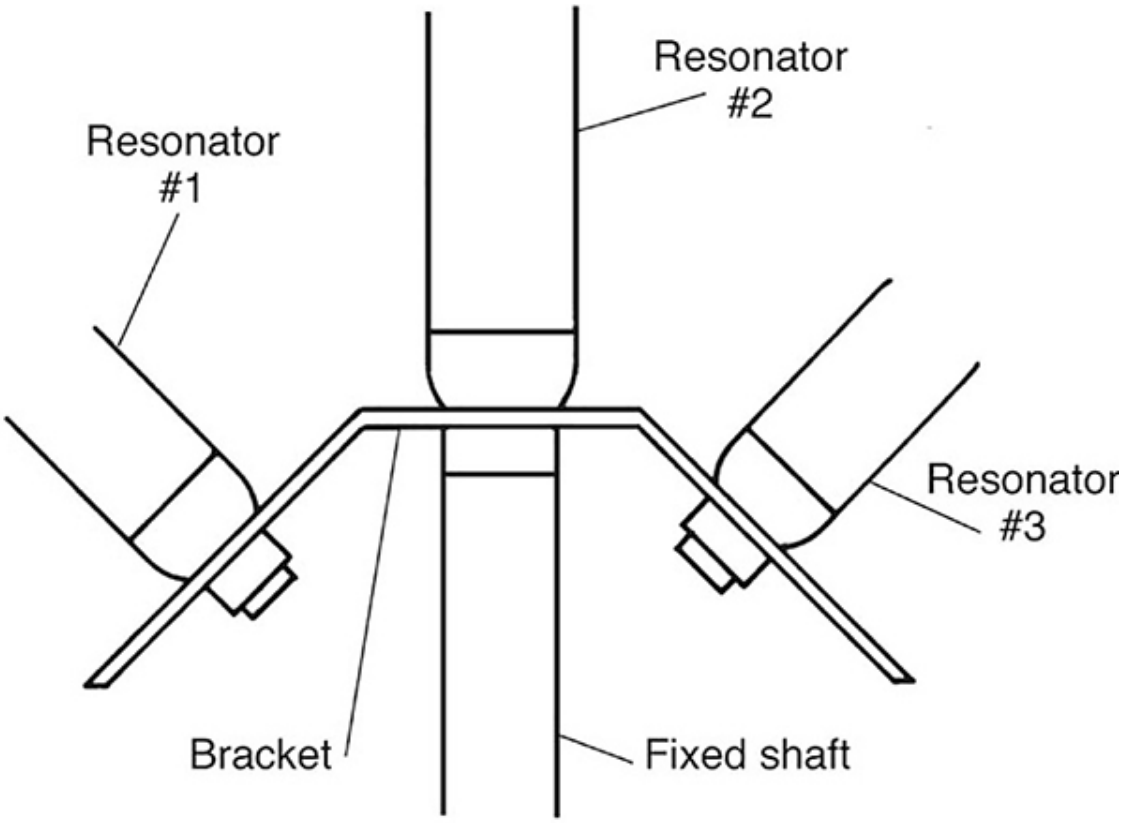


FIGURE 18.5 Multiband top-loaded antenna.

A common problem with short coil-loaded mobile antennas is that they tend to have a high *Q*, making them sharply tuned with a narrow useable bandwidth. As little as 25-kHz change of operating frequency will detune the antenna significantly on 80 m. Although an antenna tuner at the transmitter will reduce the VSWR to a point that allows the transmitter to operate, that type of tuner does not fully address the problem. The actual problem is that the antenna is very short relative to a quarter-wavelength, and the feedpoint impedance of the antenna varies drastically with small changes in operating frequency. The only cure for this problem is to readjust the resonator's adjustable shaft as the band segment is changed. This solution requires the operator to get out of the vehicle in order to do a good job of retuning. In recent years more and more mobile operators have switched to using a motor-driven variable inductor for the loading coil. Today several manufacturers like High Sierra and Tarheel offer both base-loaded and center-loaded coils

and antenna assemblies that are motor-driven or relay-selectable to permit frequency changing from inside the vehicle. Some models even feature a remote-control tilt to allow driving into the garage.

Arguably the best enhancement that one can make to any short HF whip is a capacitive top hat. Usually consisting of a half-dozen or more radial spokes emanating from the top of the stinger to a circumferential support wire or tube, a top hat increases the electrical length of the entire assembly, bringing more of the natural high-current portion of the monopole out of the feedline and into the vertical element. Benefits include overall antenna efficiency improvements and increased VSWR bandwidth. Top hats for 21 MHz and below should be as large as you can make them; it's unlikely you'll exceed the dimensions at which they become an electrical detriment. Mobile HF antennas from some manufacturers have top hats with a 3- or 4-ft diameter; in many designs, guying of the upper half of the antenna is required.

18.4 Tuning HF Mobile Antennas

Tuning a typical mobile HF antenna is not particularly difficult, since there's very little you can adjust. Here's a sequence that works for center-loaded verticals with fixed-inductance loading coils:

- Initially set the whip or stinger to the manufacturer's recommended length. If you have no guidelines, set the total length to the maximum you can within the practical constraints (overpasses, utility lines, etc.) mentioned earlier.
- If your loading coils are not adjustable, attempt to find a length for the stinger (the portion of the whip above the loading coil) that produces a low VSWR back at the transceiver in the lower portion of your desired range of operating frequencies on a given band. While transmitting briefly on spot frequencies (no further apart than 25 kHz on 80 and 40) throughout the appropriate band, determine the frequency of minimum VSWR, if

possible. Use a VSWR meter at either end of the coaxial cable between the mobile rig and the antenna. Alternatively, use the VSWR meter built into many mobile rigs. But do not have an antenna tuner in the line at this stage of the process. If there is not an obvious minimum VSWR somewhere in the desired range of operating frequencies, try a different stinger length. If a minimum VSWR can be found and it's on the high side of your chosen center frequency, lengthen the stinger slightly and remeasure. If the minimum VSWR is on the low side of your desired center frequency, shorten the stinger and remeasure.

- Once you have obtained the lowest VSWR you can by adjusting the length of the stinger, reintroduce any matching networks, starting with any at the base of the vertical. As a last resort, use the antenna tuner in your mobile rig to minimize the VSWR. Presumably this will coincide with maximum RF out of the vertical if you have been monitoring the field strength meter mentioned earlier. The goal is to match the antenna to the feedline and the transmitter to maximize RF transfer and minimize any tendency by the transmitter to reduce power in the face of VSWRs higher than 1.0:1.
- At some point you may find you can't make the stinger any shorter because it has bottomed out in the clamp at the top of the loading coil. If so, you may need to cut some of it off. Be careful! It's hard to add length if you cut it prematurely. If the minimum VSWR is not unreasonable (say, 1.5:1) with the original stinger stuffed all the way into the clamp, be satisfied. Do any further VSWR reduction with a matching network in the line or back at the transceiver. In short: Learn when to leave well enough alone.

In other mobile whip installations, there may be an adjustable (or tap-selectable) loading coil in addition to the adjustable stinger. Figure 18.2 takes this idea one step further with an adjustable or tapped center-loading coil in series with a tapped impedance-matching coil at the base. Alternatively, the upper coil is a fixed inductance, and the only adjustment to the upper part of the vertical is to the length of the adjustable resonator shaft (stinger). Ideally, a dip meter (see Chap. 30) is available to set the upper inductance of the stinger so as to roughly resonate the vertical at the desired operating frequency. Then adjust the matching network at the base of the antenna or back at the transceiver. With this configuration, it may be possible to initially adjust the upper coil or stinger for maximum reading on a field strength meter, then follow up with adjustment of the base matching network for minimum VSWR on the transmission line back to the transceiver. Tuning the upper coil or stinger for maximum field strength can lead to ambiguous or even erroneous results if the output power of the transceiver changes substantially in response to the changing VSWR it sees as the stinger wire or upper coil is adjusted.

18.5 Marine Radio Antennas

From its inception, radio communication considerably lessened the dangers inherent in sea travel. The maritime industry adopted wireless earlier than other fields. The infamous Titanic sinking might have cost less in terms of human life if the wireless operator aboard a nearby ship had been on duty. Today, shore stations and ships maintain 24-hour surveillance and vessels can be equipped with autoalarm devices to wake up those who are not on duty. Even small pleasure boats are now equipped with radio communications, and these are monitored around the clock by the U.S. Coast Guard.

When operating type-accepted radio equipment on marine frequencies, the small shipboard operator can have a selection of either HF single-sideband communications or VHF-FM communications. The general rule is that a station must have the

VHF-FM and can be licensed for the HF-SSB mode only if the VHF-FM is also aboard. The VHF-FM radio is used in coastal waters, inland waters, and harbors. These radios are equipped with a high/low power switch that permits the use of low power (1 W typically) in the harbor but higher power when under way. The HF SSB radios are more powerful (100 W typically), and are used for offshore long-distance communications beyond the line-of-radio-sight capability of the VHF-FM band. A radio amateur can usually utilize the full panoply of modes, subject to local regulations. For radio regulatory purposes, the world has been divided into three regions, and the rules differ from region to region. When using amateur radio equipment outside home coastal waters, find out about the pertinent regulations before embarking.

The problems of antennas on boats are the same as for shore installations but are aggravated by certain factors. Space is limited on smaller boats, so most HF antennas are, like mobile antennas, quite short relative to the wavelength in use. RF grounds are harder to come by on a fiberglass or wooden boat, so additional grounds must usually be provided.

An example is shown in Fig. 18.6. The radio is connected to a whip antenna through a transmission line and a tuner on HF, while also being grounded to an externally provided ground plate. Over the years, shipboard radio grounds have taken many different forms. The ground might be copper or aluminum foil cemented to the boat hull. Or it might be a bronze plate or hollow bronze tube along the centerline of the boat or along the sides just below the waterline. The ground is connected to the engine. Attention must be paid to the electrical system of the boat when creating external grounds to prevent electrolytic corrosion from inadvertent current flows.

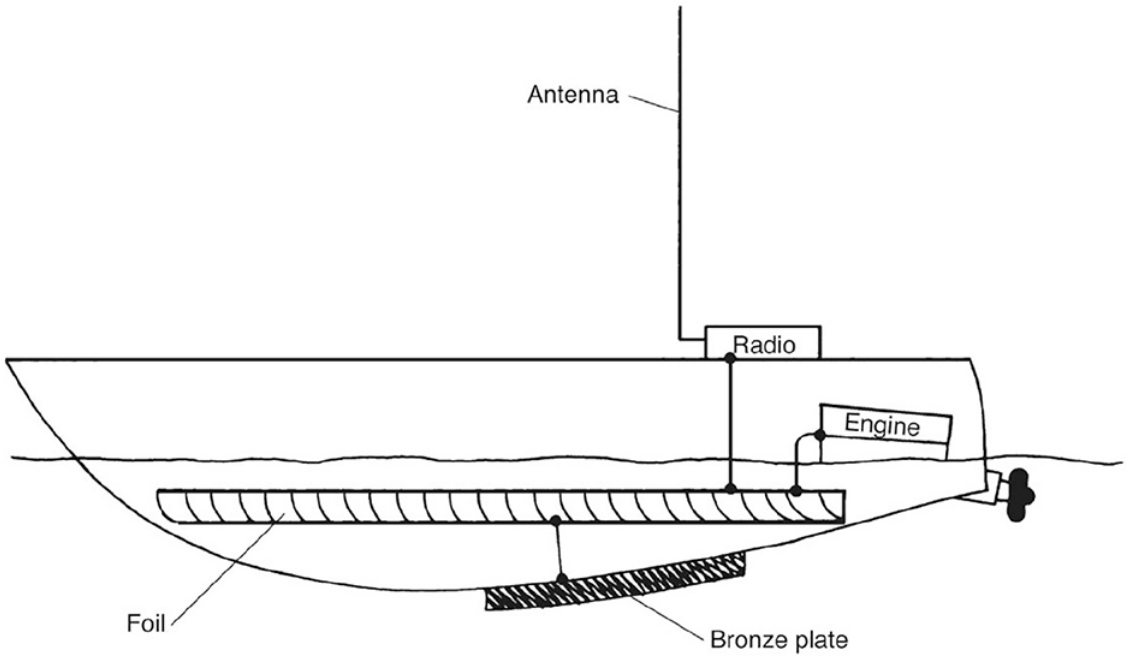


FIGURE 18.6 Motorboat grounding system.

On sailboats, the whip antenna (especially VHF-FM) might be mast-mounted as shown in Fig. 18.7, or it might be an aluminum mast itself. Alternatively, a stay might serve double duty as an HF antenna. In any event, the same grounding schemes apply, including the additional option of a metal keel or metal foil over a nonmetallic keel.

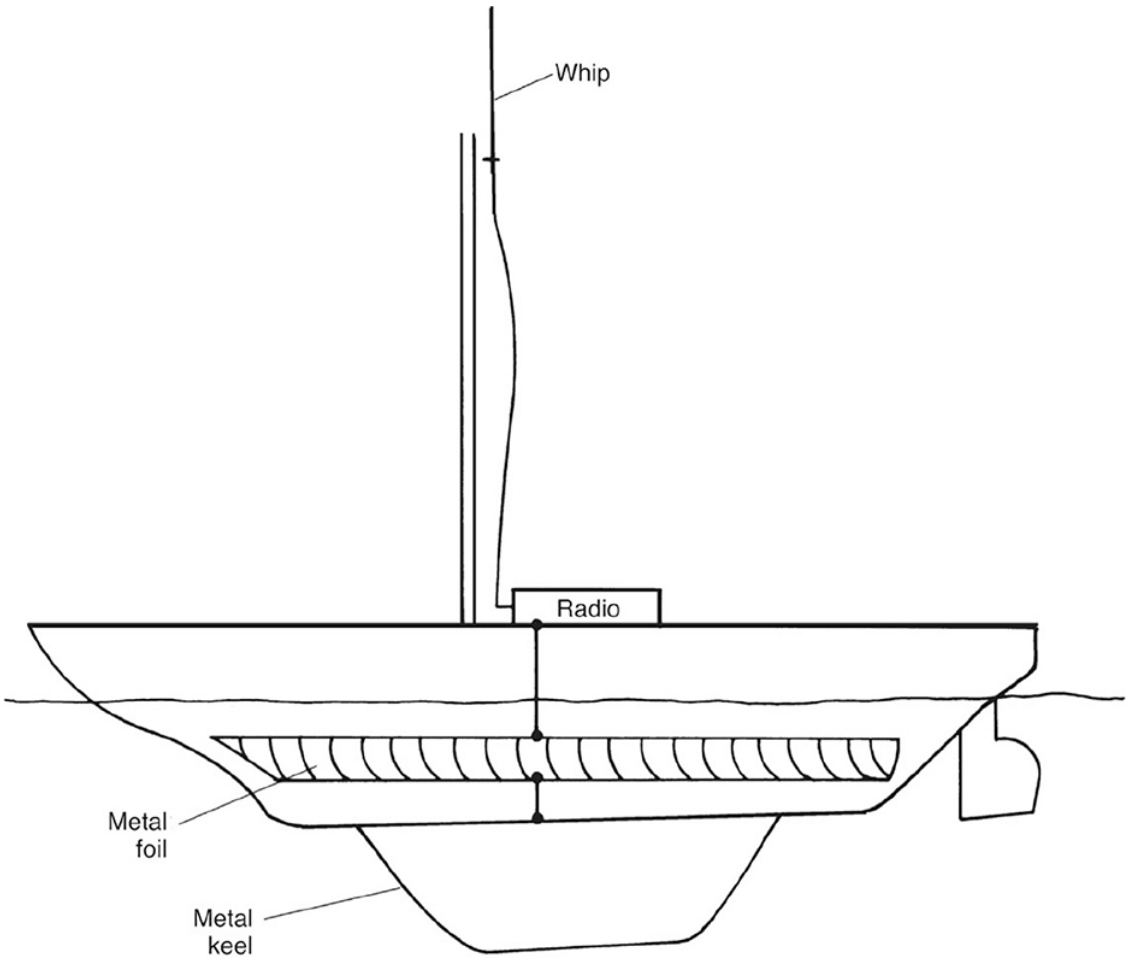


FIGURE 18.7 Sailboat grounding.

The whips used for boat radios tend to be longer than land mobile antennas for a given frequency because they are not subject to the same height limitations. The VHFFM whip (Fig. 18.8) can be several quarter-wavelengths and take advantage of collinear gain characteristics. Whips for the HF bands can be 10 to 30 ft in length and often look like trolling rods on powerboats. You will also see citizens band radios and whips on board.

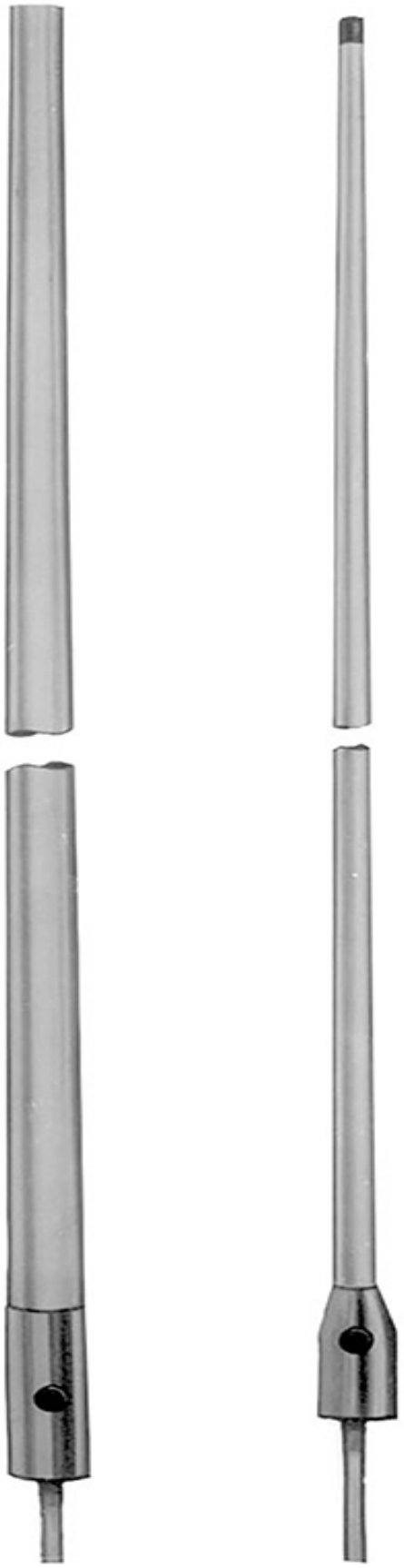
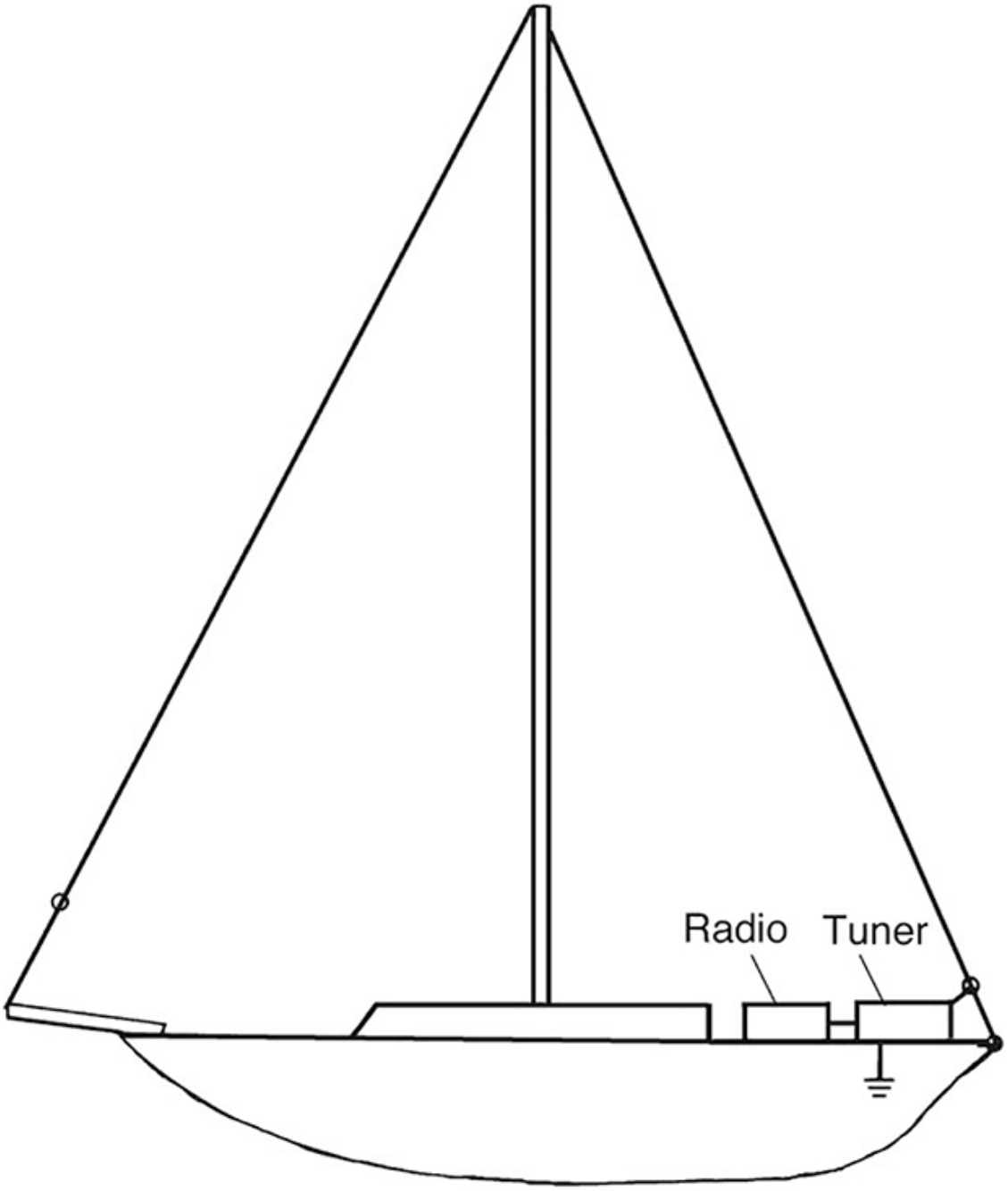


FIGURE 18.8 VHF-FM boat antenna.

Longwire antennas also find use in marine service. Figures 18.9 and 18.10 shows two possible types of installation. The antenna in Fig. 18.9 shows a wire stretched between the stern and bow by way of the mast. The antenna is end-fed from an antenna tuner or "line flattener." The longwire shown in Fig. 18.10 is similar in concept but runs from the bottom to the top of the mast. Again, a tuner is needed to match the antenna to the radio transmission line. The antennas used in this manner are actually not longwires in the rigorous sense but, rather, random-length antennas.



F IGURE 18.9 Sailboat HF antenna.

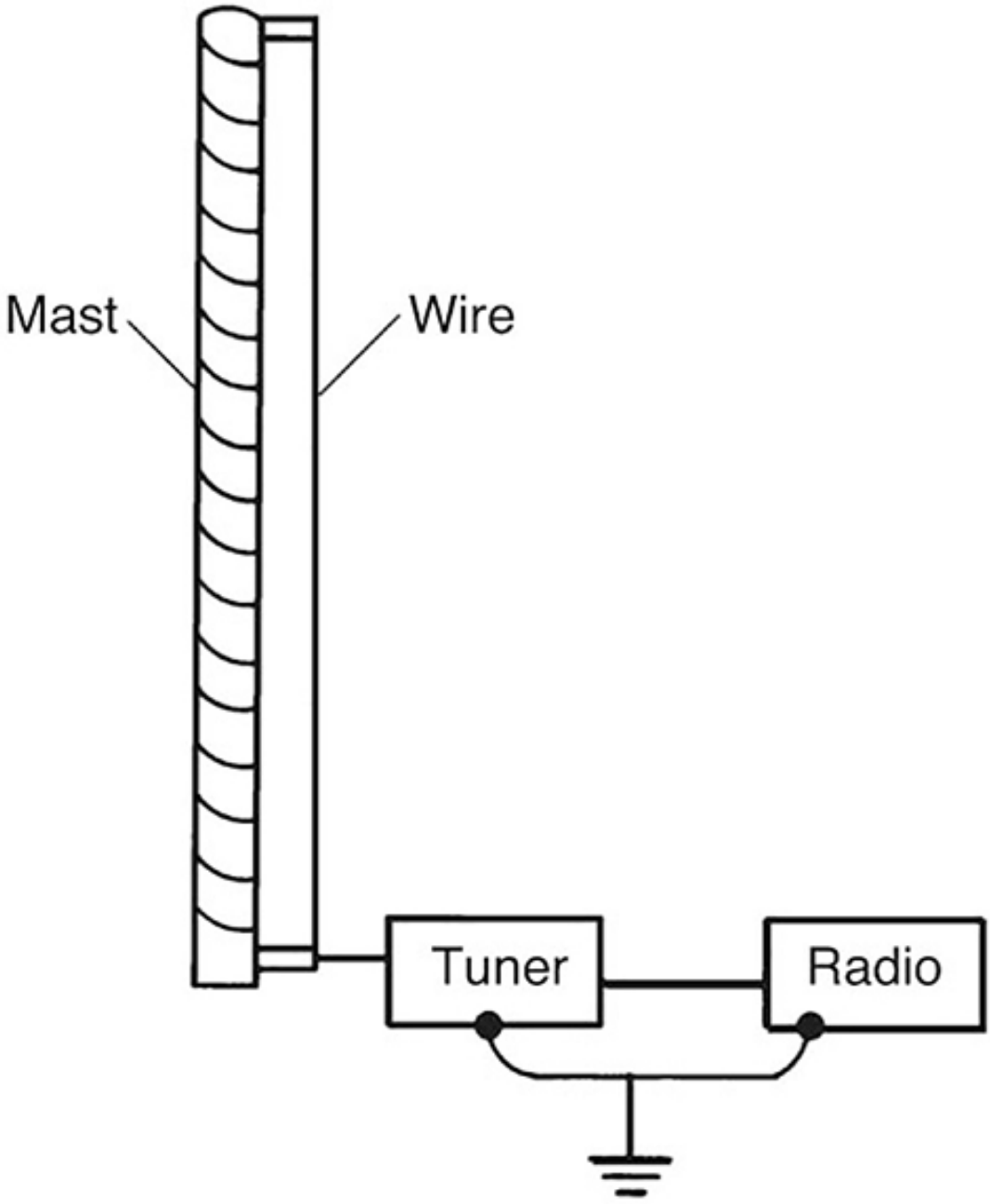
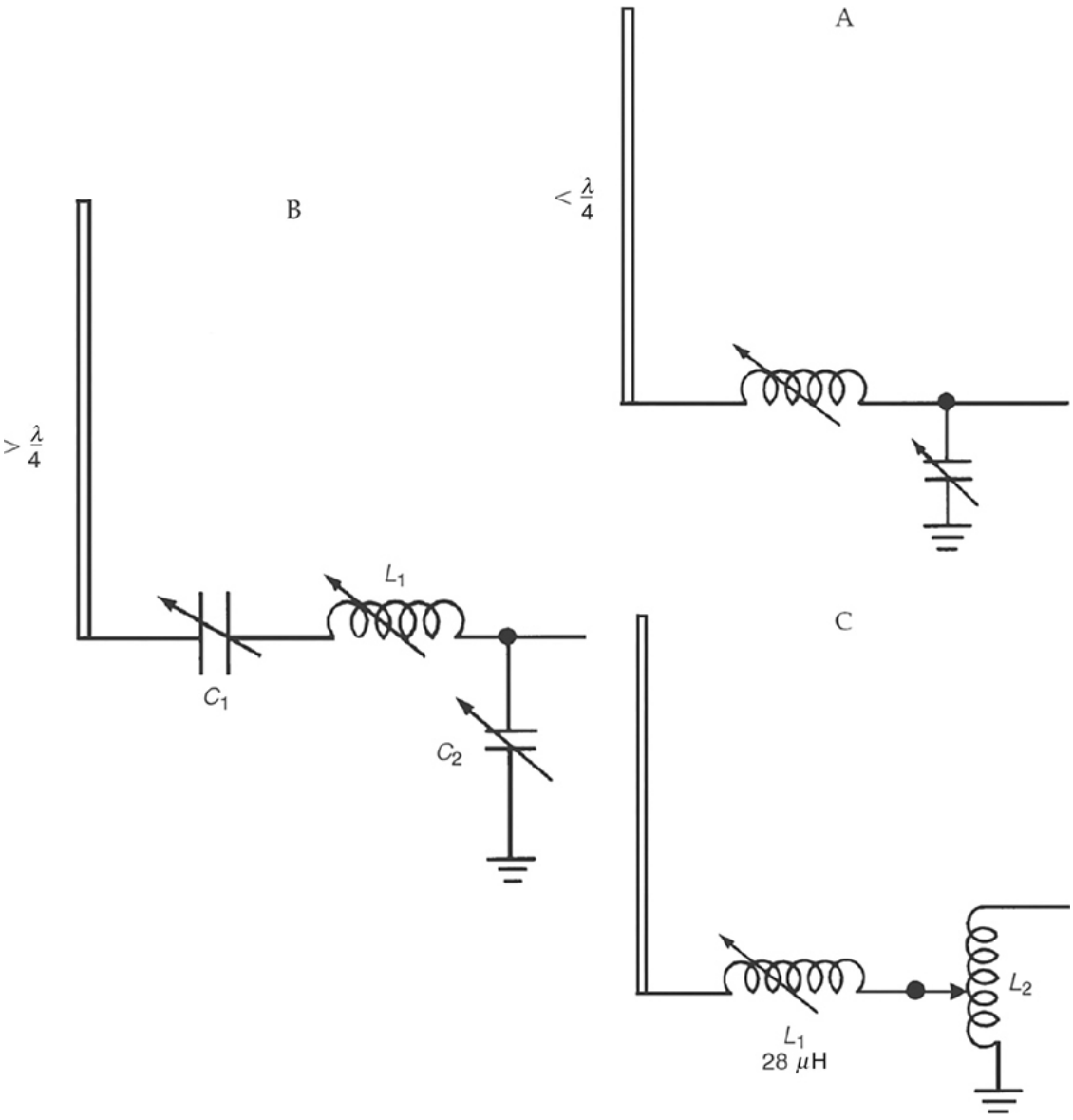


FIGURE 18.10 Mast antenna.

The tuner can be any of several designs, shown in Fig. 18.11. The reversed L-section coupler shown in Fig. 18.11A is used when the antenna radiator element length is less than a quarter-wavelength. When the antenna is greater than a quarter-wavelength, the circuit of Fig. 18.11B is the tuner of choice. This

circuit is a modified L-section coupler that uses two variable capacitors and an inductor.



F IGURE 18.11 Antenna tuners for broad HF antennas.

Finally, we see in Fig. 18.11C the coupler used on many radios for random-length antennas. Two variable inductors are used:

L_1

resonates the antenna, and

L_2

matches the feedpoint impedance to the characteristic impedance of the transmission line. In some designs, the inductors use switchable taps. The taps are selected when the operator selects a channel. With frequency synthesized transceivers, the antenna must be tuned every time the frequency is changed or an automatic preselected tuner is used.

The line flattener (Fig. 18.12) is a standard transmatch antenna tuner that reduces the coaxial-cable transmission line VSWR. This type of tuner is especially useful for transmitters that react badly to high VSWR. Many transceiver designs incorporate shutdown (foldback) circuits that progressively reduce power as the VSWR increases. The line flattener tunes out the VSWR at the transmitter. It does nothing to tune the antenna but facilitates maximum power transfer from the transmitter to the transmission line.

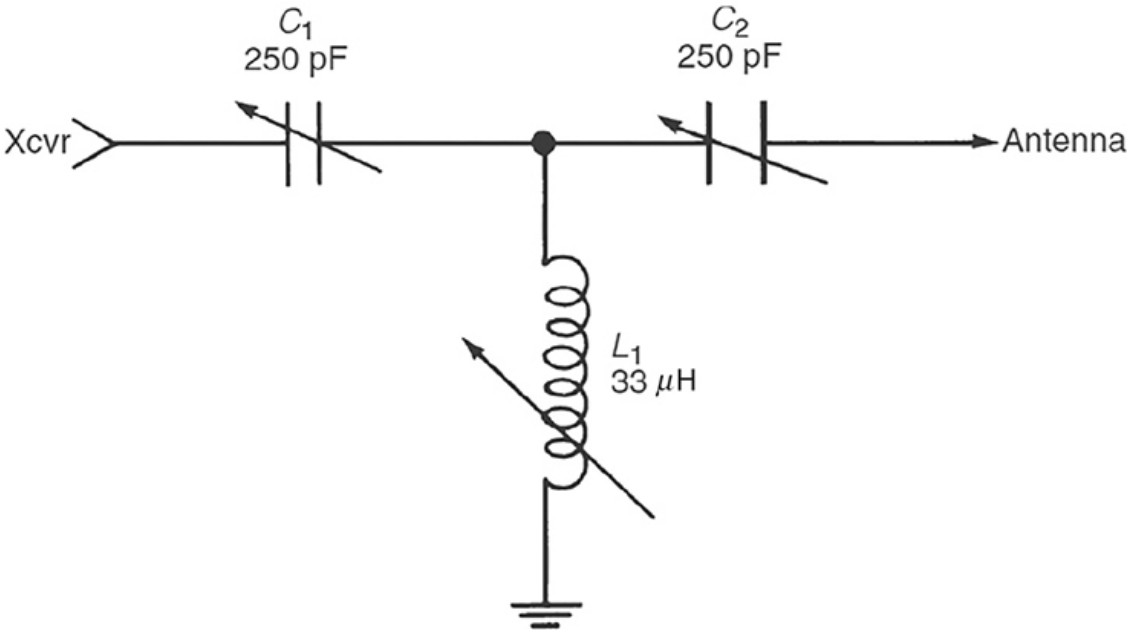


FIGURE 18.12 Line flattener tuner.

CHAPTER 19
Emergency and Portable Antennas

When disaster strikes and traditional infrastructure is damaged or unavailable, ham radio operators step in to assist with emergency communications. This takes place during times of civil unrest and in war-torn parts of the world and after hurricanes, earthquakes, and other natural disasters. Ham radio operators support search and rescue operations and help with public events like parades, races, and competitions in the backcountry. Other types of valuable service include weather spotting, support during medical emergencies, and coordinating international aid and relief.

During times of distress, ham radio operators must cope with damaged equipment and deal creatively with hard situations. This might mean trying to establish critical long-range communications with a low-power transmitter and portable antenna instead of high wattage transmitters with larger and more efficient fixed antennas.

In the most extreme situations, the ability to communicate can means life or death for someone. Given the inefficiency of the loaded whips typically used as mobile antennas in the low-HF region, the low power levels generally found in mobile rigs, and the crowded conditions on the 80-, 75-, and 40-m bands, it becomes a matter of more than academic interest how they might increase the signal strength from their portable (or mobile) emergency station. Anything they can do, easily and cheaply, to improve their signal is like having money in the bank. Fortunately, there are several tricks of the trade that will help.

In cases of emergencies on most highways, we are likely to be in range of some repeater, so we would use a VHF band (probably 2 m) to contact police or other emergency services—perhaps through manual assistance from another ham on the repeater, perhaps

through a repeater autopatch. With the wide availability of repeaters it can be valuable for an amateur backpacking or four-wheeling into remote areas to be familiar with nearby repeater locations and frequencies.

As we saw in Chap. 18, the HF mobile configuration is inefficient by its nature. Mobile antennas are typically short and have a low radiation efficiency, making them difficult to match, and often are coupled with a poor RF ground, leading to power dissipation in the lossy ground around the antenna. In most emergency situations, the preferred coverage distance is close for HF propagation modes. The use of a conventional mobile antenna mounted on the vehicle results in a vertically polarized signal with maximum amplitude at relatively low wave angles, while ionospheric propagation modes for the distances involved overwhelmingly favor an antenna that puts most of its energy into the higher wave angles, which would favor a low, horizontally polarized radiator.

If the mobile station is temporarily in a fixed location, there are ways to improve the performance of mobile HF antennas. We will discuss the limitations of typical mobile setups below, with the intention of identifying simple workarounds for the temporary situations that most emergency and portable operations imply. We will assume that the vehicle and the mobile gear it is carrying are stationary for the duration of the emergency or portable activity.

19.1 Improving Radiation Efficiency

The ideal solution when possible is to replace a mobile antenna with a longer, more efficient, stowable antenna that can be brought out and erected when needed. Radiation efficiency increases dramatically as the length of the antenna approaches $\lambda/4$. One candidate is a surplus military HF whip antenna. Intended for jeeps and communications trucks, these antenna and tuner combinations are collapsible and are as efficient as any on the market.

Alternatively, a quarter-wave length of wire (or, with a suitable antenna tuner, anything longer than the mobile whip) can be attached where the mobile whip was located previously. A coil of wire attached to a ball or other weight can be thrown to get one end of the antenna wire up in the air, 20 or 25 ft, draped over a tree branch. If necessary, support the far end of the wire a little distance from the rear of the vehicle so that the lack of height is made up for by the added horizontal distance the wire is covering. This creates a sloping vertical, which is a reasonably good antenna.

19.2 Reducing Ground Losses

For frequencies up to 6 m the typical vehicle body is not large enough to provide an adequate ground plane. A better solution is to provide a counterpoise or elevated ground plane. As we saw in the chapter on verticals (Chap. 11), a vertical monopole, such as a mobile whip, is really only half an antenna. The other half is distributed throughout the ground return for the antenna, and a failure to provide an adequate ground system is equivalent to further compromising the antenna itself.

Figure 19.1 is the schematic representation of a proposed way to enhance the ground return efficiency for a mobile whip used when a vehicle is temporarily stationary. At its minimum it consists of the addition of three or four radials, each

$\lambda/4$

in length on the band(s) of interest and spread as equally as possible around 360 degrees. A ground-plane antenna should be mounted $\lambda/8$ or higher with some kind of additional support, but the degradation in performance is gradual as the height becomes lower, so you do the best you can with what you have.

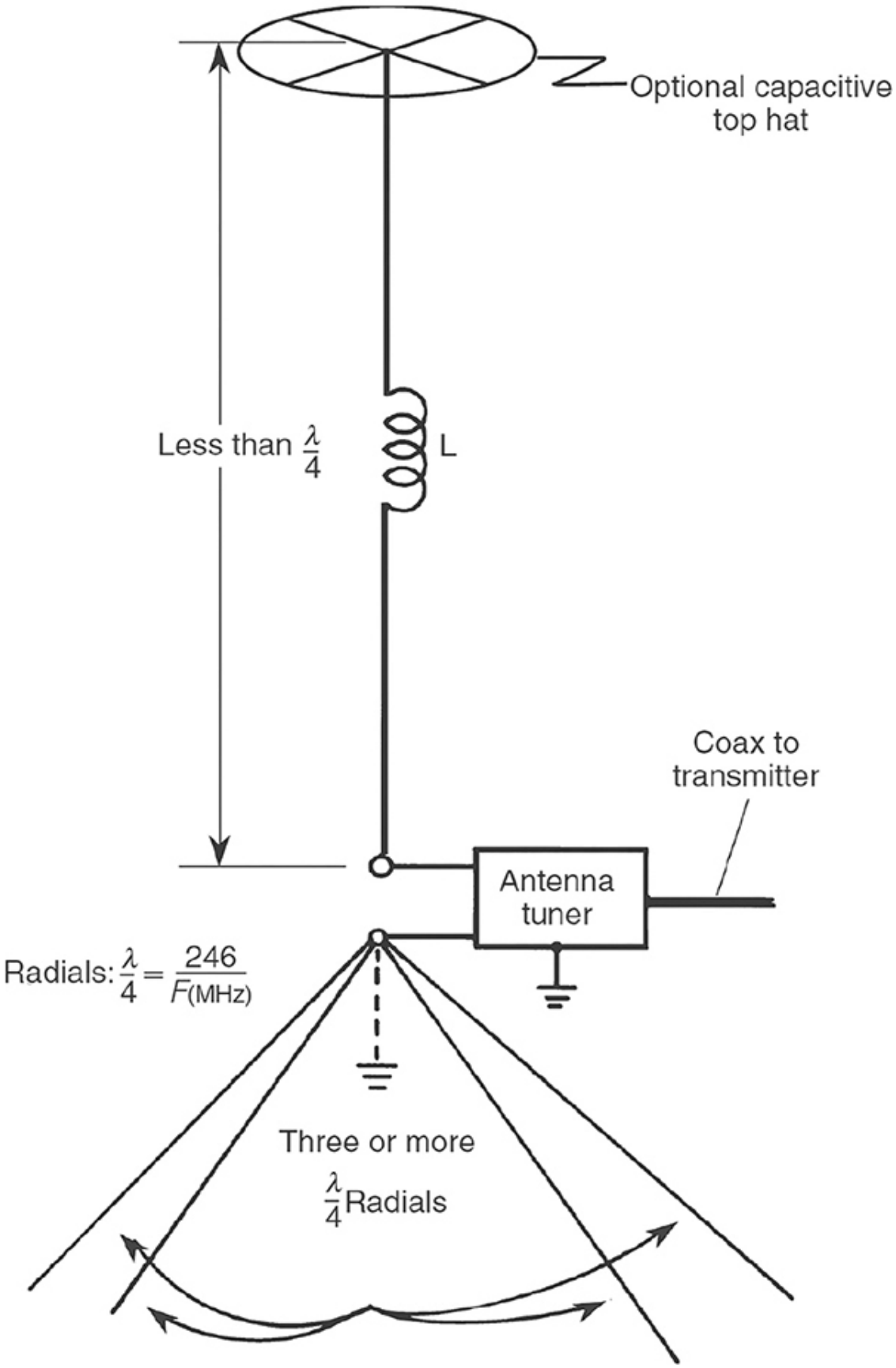

Optional capacitive top hat
Less than $\frac{\lambda}{4}$
L
Coax to transmitter
Antenna tuner
Radials: $\frac{\lambda}{4} = \frac{246}{F(\text{MHz})}$
Three or more $\frac{\lambda}{4}$ Radials

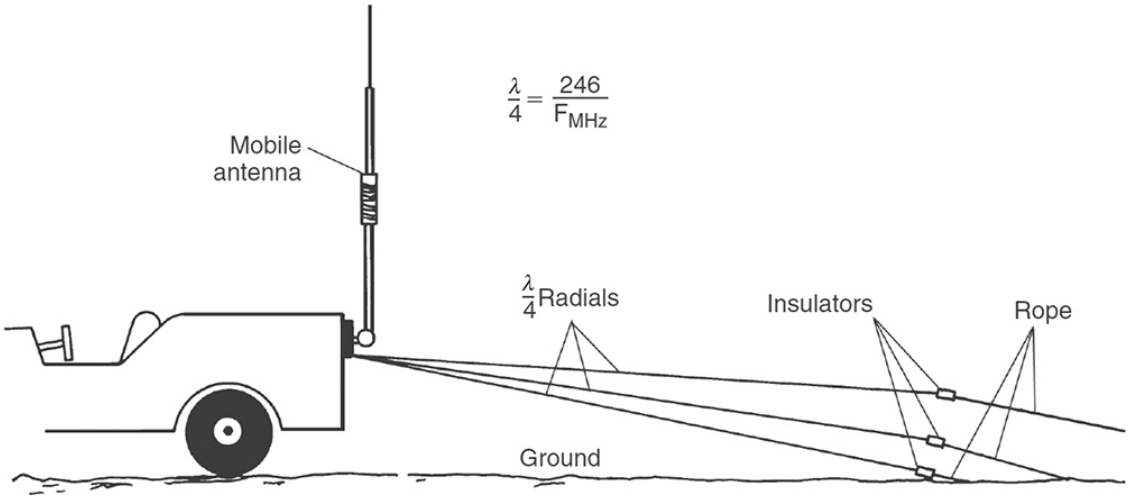
FIGURE 19.1 Radials can enhance a basic mobile antenna system at a fixed location.

Radial wires should be strong enough to withstand being repeatedly installed, taken down, coiled up, and stored. Current-handling capability in radials is seldom an issue. Antenna suppliers sell spools of bare #18 wire for radials; a 500-ft spool takes up little space in a car trunk. For many portable and emergency setups, insulated wire may be best because there may be a need to drape one or more radials across metallic objects. Metal-to-metal contact with other objects can cause extra noise in the receiver. For 40 m, four

$\lambda/4$

radials total 132 ft. Allowing some extra wire for connections to the base of the whip and for attaching to rope or twine at the other end means 150 ft of #14 wire.

Figure 19.2 shows a system that can further improve the performance of a mobile rig in stationary situations. The mobile antenna uses the normal base mount attached to the rear quarter-panel of the car adjacent to the trunk lid. An all-metal grounding-type binding post is installed through an extra hole drilled in the base insulator (see Fig. 19.3). Radials for portable operation are attached at this point. The binding post should be sized to accommodate three or four #18 radials. Additional radials can be attached to the base indirectly by connecting them to the top of one or more of the radials already on the binding post.



F IGURE 19.2 Multiple radials are even better.

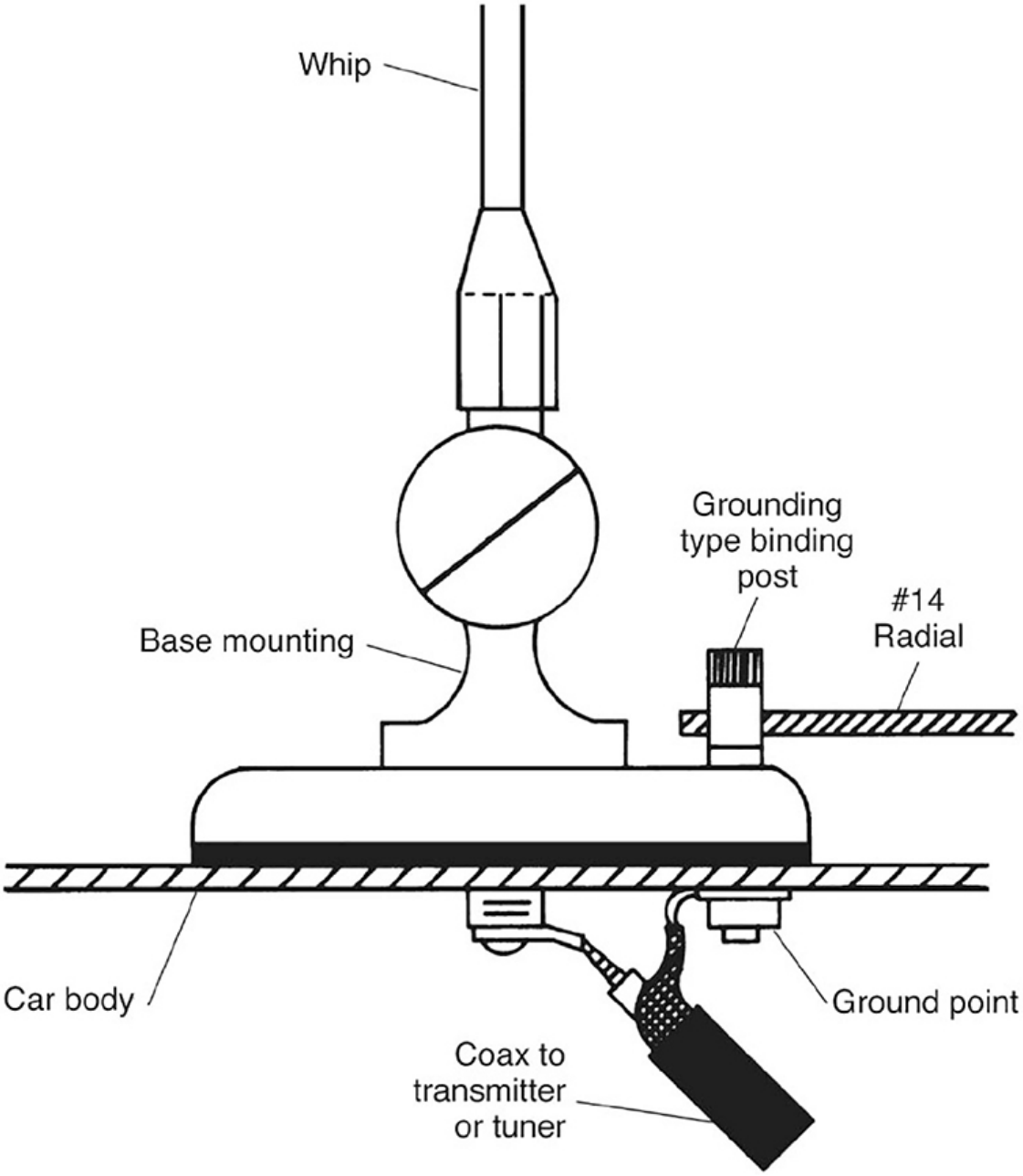


FIGURE 19.3 Connection of radials to a mobile antenna.

19.3 Selecting the Right Polarization

To improve the HF signal from an emergency or portable station setup in critical situations, the antenna polarization can be optimized for the coverage distance and surrounding terrain characteristics. Vertical polarization may occasionally be used for

low-band operation in amateur radio competitive events, but most scenarios call for a horizontally polarized antenna.

When operating a mobile system in stationary mode, putting up a dipole is an often overlooked approach. In addition to increased radiation at high wave angles, the dipole eliminates the need for ground radials and can cut the required amount of wire in half.

The biggest impediment to erecting a dipole sufficiently high in the air is the number and spacing of suitable available supports. Trees, flagpoles, roofs of buildings, and upstairs windows can all be candidates. Utility poles should normally be avoided to avoid the risk of shorting power lines. Depending on the distances to be covered, a height of 30 or 35 ft for an 80-m dipole can be adequate.

For low-power operation, the dipole can be fed with lightweight RG-58 or RG-8X coaxial cable. Open-wire line (OWL) can also be used, but most will find it easier to keep a stored roll of coaxial cable. In a pinch, a zip cord or electrical wire makes an easily obtainable feedline. The impedance can be around 72 Ω, which is reasonably good, but losses can become excessive on the higher bands. In a worst-case scenario, lengths of rotor cable, telephone cable, or Cat 5 or Cat 6 Ethernet cable can be pressed into service.

Figure 19.4 shows the common dipole and the normal equation for determining the approximate length. Typically, it is cut long and then trimmed until the VSWR drops to its lowest point, a nicety that might not seem altogether important in an emergency. One or both ends of the dipole must be supported on trees, masts, or some other elevated structure.

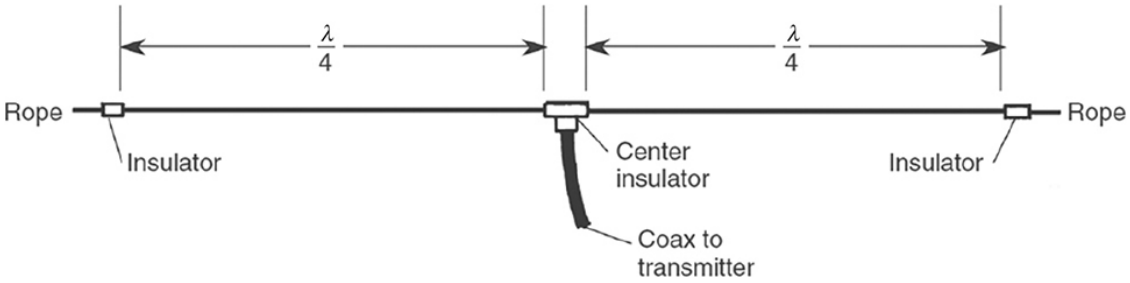


FIGURE 19.4 Basic dipole.

If only a single high support is available, a sloping dipole or an inverted-vee can be substituted for the dipole. For the antenna of Fig. 19.5, assume as a starting point that the length of each leg is 6 percent longer than would be obtained by using the dipole formula of Fig. 19.4.

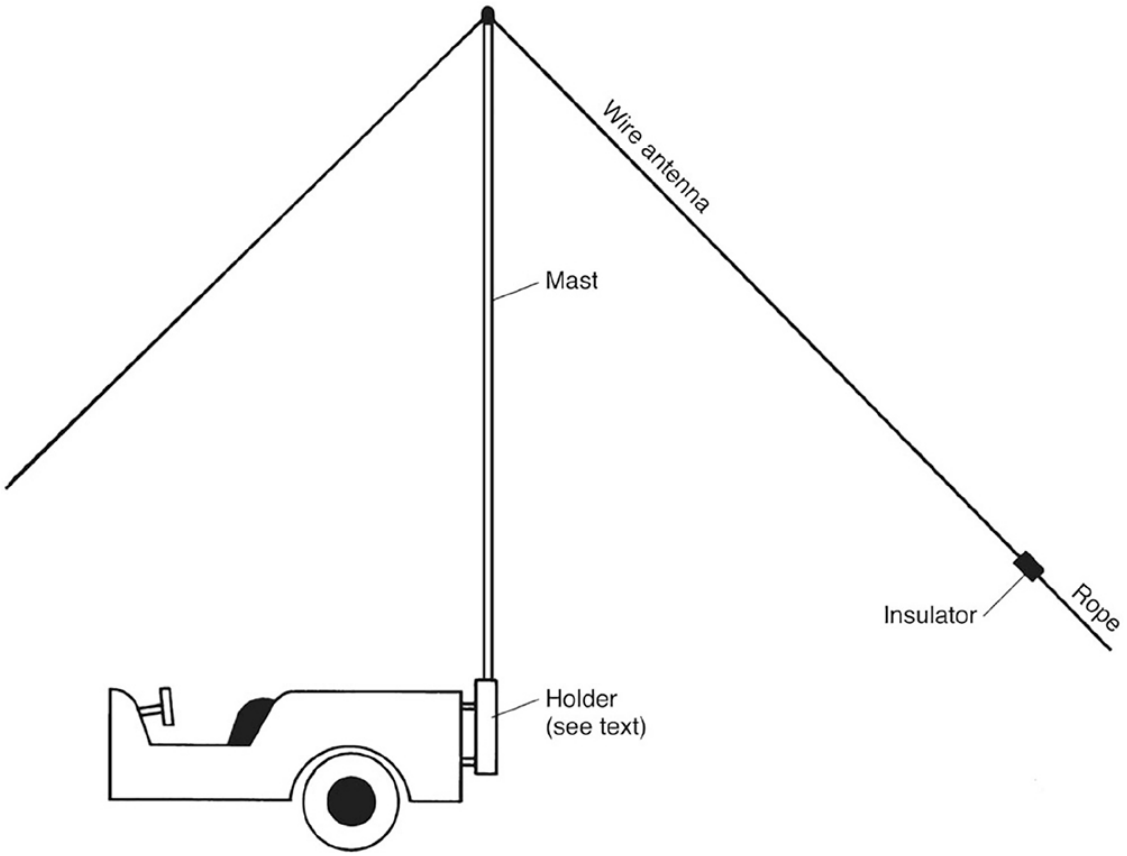


FIGURE 19.5 Portable inverted-vee antenna.

A final approach is to combine the vertical with the horizontal and run a random-length wire up as high as you can and then horizontally as far as you can. A total of 140 to 160 ft works well for the 160 m band and a total length of 70 to 100 ft should work well for 80 m. This is an inverted-L, which is an efficient cousin to a conventional

$\lambda/4$

monopole. Its radiation and matching characteristics are covered in detail in Chap. 20. This antenna is simple to erect and it provides a

mix of high-angle and low-angle radiation. For proper operation, the inverted-L must have radials, a counterpoise, or a tuned RF ground at the feedpoint.

These portable antennas will almost always outperform the basic mobile whip. For temporary, emergency-related use, we can get away with installation and support techniques that would be unthinkable in more permanent installations, although when possible antenna tuners or VSWR meters should be used for best performance.

19.4 Temporary Supports

Anyone serious about being prepared for emergency or portable radio operation will develop a list of needed items. The extent of the list will depend on the magnitude of the planned operation and the conveyance mechanisms available. Components for a functioning radio station for backpackers will be fewer than for one for stowing some gear in an off-road vehicle. Size and weight are constraints in both cases, but the scale is different. The largest item for portable or emergency use is often the antenna support, if required. In many parts of the world, there are trees all around but in other areas, sand or fields are all that is visible for miles around, and some form of temporary mast(s) must be employed.

Where trees are available, they can be used to get support lines up high. If trees are not plentiful, various mast alternatives, ranging from extremely heavy (military surplus or steel TV antenna sections) to reasonably light (plastic pipe and fiberglass rods) are available. The choice is determined primarily by whether you expect to have help erecting the mast and by how much room you have for stowing and transporting it. Some masts consist of telescoping sections that get progressively smaller and more limber toward the top. Most masts collapse to 6 or 8 ft in length for transport and stowage but can be pushed up to heights of 18, 25, 30, 40, or 50 ft, depending upon the type selected. Larger models are considerably heavier than shorter models and they require two or

more people to install them. Erecting a 40- or 50-ft telescoping mast is not a one-person job.

Masts made of PVC plastic plumbing pipe are somewhat lighter. If mast sections can be carried on top of the vehicle, then lengths up to 10 ft are available and can be joined together at the site with couplings. If the sections are going to be transported in the vehicle, shorter lengths may be needed. PVC pipe is relatively thin walled and can become too flexible as it is extended in length. Sizes below 1.5-in diameter will not easily stand alone without guying. Nor is PVC pipe particularly self-supporting. While a single 10-ft section might be, two or more sections together will not support themselves plus the weight of any reasonable antenna. Guying with rope or cord will be necessary.

Heavier temporary or portable antennas such as an 80-m dipole require heavier but more rigid steel TV antenna masts. Available in 5- and 10-ft lengths, these masts are flared on one end and crimped on the other for joining together to form longer lengths. Mating guy rings can be used for easier installation. Military and steel TV masts are used during *Field Day*, the annual American Radio Relay League (ARRL) emergency preparedness weekend exercise. The inverted-vee antenna is a natural for this kind of event, though a *Hustler* mobile antenna at the top of a 20-ft mast with four radials works surprisingly well given the low output power (30 W) being used.

It can be helpful to look around and see how other groups of people are solving similar problems. Off-road and recreational vehicles sport a variety of satellite dishes and other amenities that require supports and mounting hardware. RV dealers and suppliers have products that can be adapted easily to specific antenna support needs.

For temporary installations lasting only a few hours or a few days, we need not worry about long-term integrity or the practicality of the installation. Mounting the mast to the back of a four-wheeler or pickup truck with a pair of U-bolts works nicely for short-term deployments.

For lightweight masts located away from the vehicle the support could be an X-shaped base made of

$2\text{-} \times 4\text{-in}$

lumber (Fig. 19.6) or another type of base such as a large Christmas tree stand. A TV antenna rooftop tripod mount (Fig. 19.7) is easily adapted for use on the ground. None of these alternatives can be depended upon to provide a safe self-supporting installation and all of them must be guyed, even if used for only a short period. Because of the temporary nature of the installation, aluminum or wooden tent pegs can be used to anchor the guy wires, depending on the soil characteristics in the area. These are inappropriate for long-term installations but work fine for the short run.

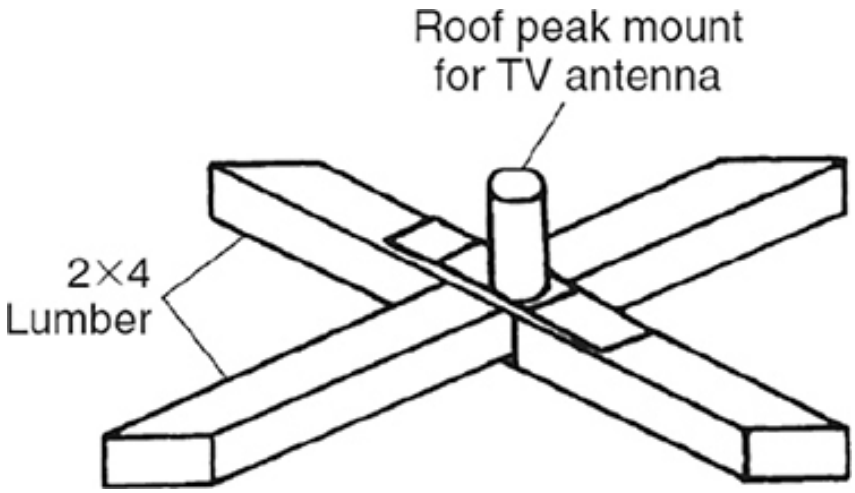


FIGURE 19.6 Wooden antenna stand.

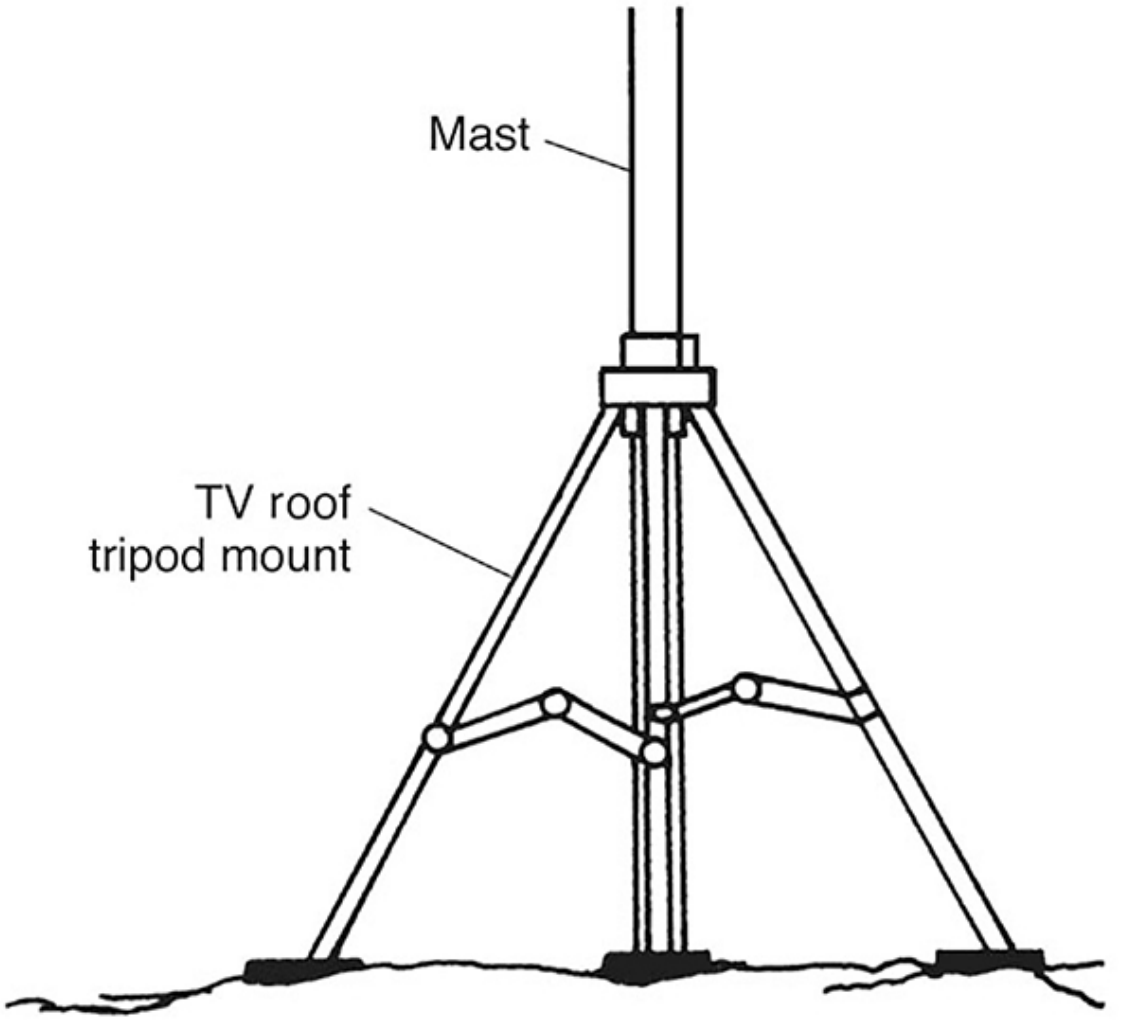


FIGURE 19.7 TV antenna tripod also works in portable situation for HF antenna masts.

19.5 Summary

The ideas presented in this chapter are meant to help the reader maximize her signal under adverse conditions in the field. While methods of getting antennas up in the air under difficult circumstances are substantially different from those techniques we would normally employ at our home or business, the electromagnetic principles that maximize signal transmission and reception do not change. In particular, remember these key points:

- Horizontal (or inverted-vee) antennas should be as close to an electrical half-wavelength as possible; don't rule out the use of traps to allow easier multiband operation in a limited space.
- Make any vertical monopoles as close to an electrical quarter-wavelength as you can; if you can get the base of the vertical off the ground, create a ground-plane antenna with two or three $\lambda/4$ radials on each band to be used. If the vertical must remain ground-mounted, try to put down eight or more radials total, with lengths as long as you can make them, up to a maximum of $\lambda/4$ on the lowest frequency of intended use.
- If you have only a single elevated support available, consider either a sloping dipole or an inverted-vee on the lower HF frequencies. Spread the legs of the inverted-vee as far apart as possible.
- Select your antenna to match the type of communications you need. If you are primarily interested in contacting others in your own general area (i.e., within a few hundred miles) on HF, use a horizontal wire. If you're on a DXpedition to a remote South Sea island, you'll probably want to use a vertical on the beach.
- Make sure you have a meter for reading both forward and reflected power or VSWR. HF transceivers usually have that capability built in. At VHF and UHF, it may be necessary to bring along a stand-alone wattmeter or antenna analyzer.
- Similarly, make sure you have an antenna tuner for the HF bands—either built into the rig or as a separate accessory. For VHF and UHF ground planes, as well as mobile whips, however, all you may need is an Allen wrench to adjust the top section of the vertical whip.
- Be sure to bring rope, twine, insulators, and short lengths of coax and other cables, with the proper

connectors on each end.

CHAPTER 20
Antennas for 160 Meters

One of the growth areas of amateur radio interest worldwide in recent years has been the 160-m band. *Topband*, as it is called by its aficionados, presents technical and operational challenges and opportunities. These include the physical dimensions required for efficient antennas, the high noise levels often found at these low frequencies, and the complexities of atmospheric propagation in that band.

Because high radiation efficiency requires antennas that are a reasonable fraction of the wavelength in size, efficiency in the medium-frequency band (300 kHz to 3 MHz) is the main challenge. The most efficient transmitting antennas for the MF and HF bands are a half-wavelength (horizontal dipoles) or a quarter-wavelength (vertical monopoles) in their long dimension. A half-wave dipole for 160 m is about 270 ft long, usually exceeding the capabilities of a typical suburban lot. Towers for the AM broadcast band (530–1700 kHz, just below the 160-m band) are usually hundreds of feet tall.

20.1 Vertical Monopoles

For an amateur without limits on space, budget, or zoning requirements, the ideal transmit antenna for 160 m is similar to the antennas used for AM radio. This means one or more $\lambda/4$ ground-mounted verticals with anywhere from 12 to 60 radials originating at the base of each tower.

Since AM radio broadcast stations are optimized for signal strength in a local coverage region, whereas amateurs are interested in long distance communication, one might think that a different antenna would be better. Would a

$\lambda/2$
dipole for 160 m be a better choice? To radiate well at the low elevation angles that are most useful for long-haul communications, a horizontally polarized antenna should be at least a half-wavelength above ground. On 160 m, this suggests a minimum height of nearly 300 ft for the antenna supports. Although a low dipole will normally provide acceptable nighttime coverage on 160 m throughout a radius of a few hundred miles around the antenna, sunlight greatly increases absorption of near-vertical signals in the lower layers of the ionosphere, making daytime short-haul communication problematic. Further, during periods of very low sunspot activity, the critical frequency,
f_C
, can drop below 1.8 MHz, with the result that nighttime skip entirely disappears on the topband. An efficient vertically polarized antenna can cover the same region with a strong ground wave that is present 24/7 and is not dependent upon conditions in the ionosphere.

Surprisingly, vertical antennas require nearly as much backyard space (or footprint) as a horizontal half-wave dipole. While a 40-m vertical (33 ft high) has modest guying requirements, a full-size quarter-wave vertical for 160 stands 125 to 140 ft tall and requires a guy wire footprint having a diameter of 200 ft or greater. Local building codes may not require inspections or permits for the 33-ft antenna but may impose rigid requirements on the higher structure. An antenna with overall height taller than the maximum allowed for a principal structure (often 35 or 40 ft in the United States) may not be allowed at all by local zoning ordinances. On suburban or urban lots, a typical 40-m vertical might well be able to fall over without crossing the property line or landing on power lines. Not so the 160-m vertical.

There are ways to reduce the size of basic antenna types for the 160 m band at the cost of degraded performance. The user typically gives up a few decibels of distant signal strength for a given power delivered to the antenna and has a narrower VSWR

bandwidth before retuning of the antenna is necessary. For a thin wire antenna to radiate as effectively as possible, we want as much area as possible under the antenna's current-versus-length curve while minimizing losses in the feed, antenna structure, and ground system. Our preferred antennas are those that give maximum exposure to the natural high-current portions of the antenna element(s), exhibit a high radiation resistance relative to other losses in the system, and are a reasonably close match to available feedlines. If the antenna is a monopole, the ground system must be well designed.

20.2 λ/4 Vertical Monopole

The simplest way to give maximize exposure to the high-current portion of the antenna is to make sure that horizontal dipoles are $\lambda/2$ and verticals are $\lambda/4$ in length or slightly longer. Since the title of this book includes the word "practical," we will concentrate on the $\lambda/4$ ground-mounted vertical instead of $\lambda/2$ dipoles, which for the bands considered here must be erected at heights of 300 ft or more.

If tall trees are available, the simplest physical configuration for a vertical is to hang a 130-ft wire from one using any of the techniques in Chap. 31. The wire should be insulated and if high-power operation is contemplated a catenary or halyard system may be required to prevent arcing from the antenna to tree limbs when transmitting. The height of the antenna should be adjusted so that the bottom end of the wire is within 1 or 2 ft of the ground beneath it for direct access to the feedline, a base-mounted antenna tuning unit (ATU), or some other form of base matching network. Safety

precautions must be observed to prevent people and animals from coming in contact the antenna.

For proper operation as a grounded monopole, a system of wire radials must be installed on or near the surface of the ground directly beneath the vertical. Wires in contact with the earth or within a small fraction of a wavelength need not be cut to a resonant length. A rule of thumb is to extend ground wires at least as far out along the earth's surface as the vertical is tall. A good radial field for a $\lambda/4$ vertical operating in the 160 m band has wires approximately 120 to 150 ft long coming together electrically and physically at the base of the vertical.

How many radials are sufficient? AM broadcast stations use at least 120 such wires equally spaced around the full 360 degrees. But in some tests 60 radials was the point at which no further improvement in ground system performance could be measured. In amateur practice, anything over 20 or so radials is adequate as long as the length of the vertical radiating element is sufficient to keep the radiation resistance reasonably high.

As discussed in Chap. 11, a full-sized $\lambda/4$ ground-mounted monopole has an input impedance of 36 Ω at its base. For transceivers capable of RF outputs up to about 200 W, the resulting SWR at or near the resonant frequency of the wire is not a serious mismatch and it may be possible to feed the antenna directly with 50-Ω coaxial cable. The bottom end of the vertical wire goes to the center conductor of the cable and the common junction of all the radials that make up the RF ground system beneath the wire attaches to the shield or braid side of the cable.

At higher power levels it may be necessary to add an ATU at this point in order to convert the native impedance of the vertical to that of the actual transmission line used: 50 Ω, 75 Ω, 300 Ω, as appropriate. There are several reasons for this. The power amplifier being used may not tolerate SWRs much greater than 1.0:1 and

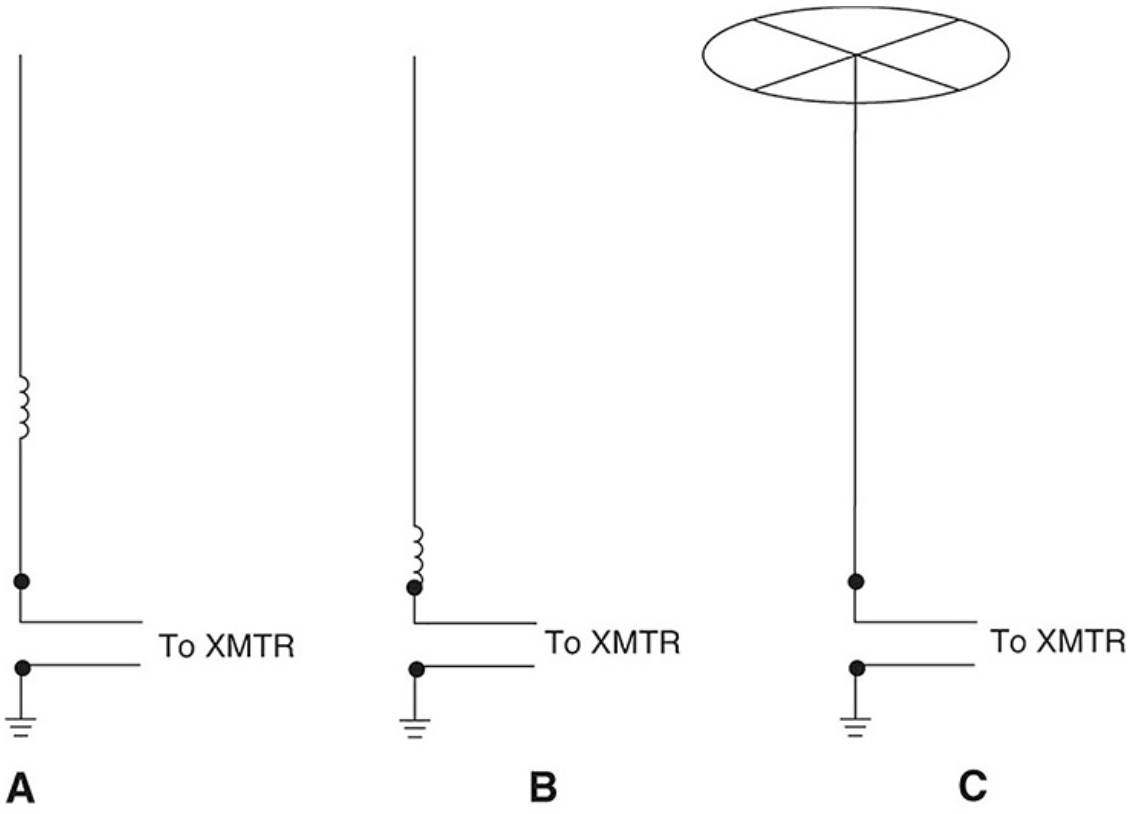
either shuts down or reduces its output power substantially in response to high SWRs. The transmission line may be very near its voltage breakdown rating and any SWR on the line will serve to reduce the maximum power it can handle. If the user intends to operate over the entire 160-m band, he or she may need a way to reduce the SWR on the transmission line when operating far from the resonant frequency of the antenna.

Adding any form of ATU can increase total loss in the antenna and transmission line system, usually as a consequence of I^2R losses from high circulating currents in inductors. In a properly designed ATU, the added loss added by the ATU itself may be less than the reduction in transmission line losses from high SWR and it may be the only way to get full performance and legal maximum power from the transceiver or amplifier feeding the line.

20.3 Shortened Verticals

The antennas in this section are all variations on the basic $\lambda/4$ vertical. Each is a shortened monopole augmented by distributed or lumped circuit elements to keep the radiation efficiency as high as possible and minimize the VSWR on the transmission line back to the transmitter.

Several configurations are shown in Fig. 20.1. The basis for all these antennas is a vertical that is physically short compared to a standard quarter-wavelength monopole with an ideal ground plane beneath it. These antennas will not work as well as a properly installed full-size antenna, but they will serve to get you on the air with a passable signal from cramped quarters that would otherwise preclude operation on these frequencies.


To XMTR
To XMTR
To XMTR
A
B
C

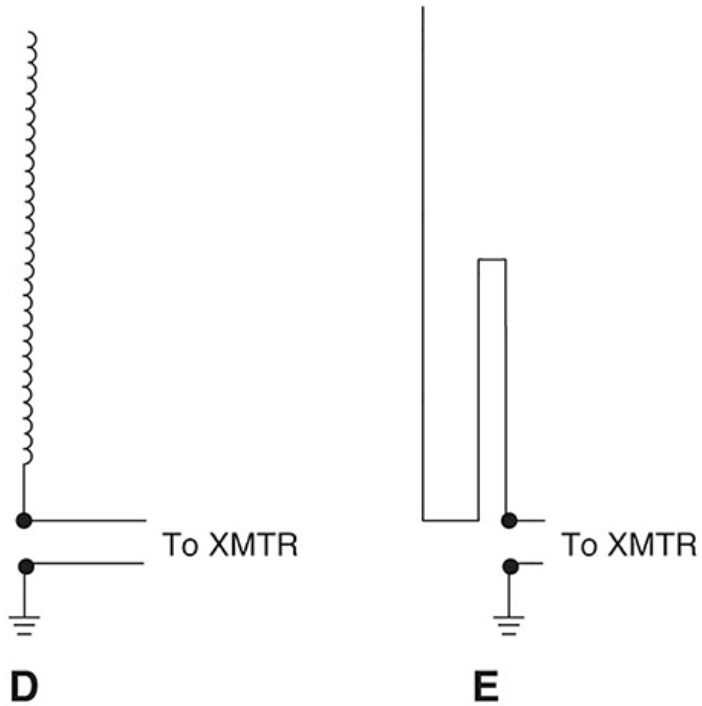

To XMTR
To XMTR
D
E

FIGURE 20.1 Examples of loaded vertical antennas. (A) Center. (B) Base. (C) Top. (D) Continuous. (E) Linear.

A vertical monopole too short for its operating frequency (i.e., less than $\lambda/4$) will exhibit capacitive reactance. To resonate such an antenna, it is necessary to cancel the capacitive reactance with an equal amount of inductive reactance, such that $|X_L| = |X_C|$. By placing an inductance in series with the radiating element, we can effectively lengthen it electrically, as measured at the feedpoint of the combined antenna element and series inductor.

An antenna that is reactance-compensated is said to be loaded. The examples considered here are inductively loaded. Three generic forms of loading are popular: discrete loading, continuous loading, and linear loading.

Discrete loading means that there is a discrete, or lumped, component (inductance or capacitance) connected to the antenna radiator. The simplest form of inductive loading inserts a loading coil near the center (Fig. 20.1A) or at the base (Fig. 20.1B) of the radiator element. Because there is little current near the open circuit end of a vertical, inductive loading at the top is ineffective. Instead, a capacitive hat (Fig. 20.1C)—often in conjunction with a center loading coil—is used to increase the area under the current-versus-length curve and hence raising the radiation efficiency of the antenna.

You may recognize these configurations as being the same as those found on mobile antennas. Low-band mobile antennas can be used in both mobile and fixed installations. Although it is tempting to use mobile antennas for fixed locations because they are readily available, they are less efficient than longer antennas. The reason is that the mobile antenna for the lower HF bands has historically been based on the standard 96- to 102-in whip antenna used by amateur operators on 10 m or by citizens band operators on 11 m,

with the overall height of the antenna limited by safety and mounting considerations. In fixed locations, longer and more efficient radiator elements are usually possible. A 16- to 30-ft-high aluminum radiator element can be constructed from readily obtainable materials.

Loaded antennas tend to be rather high *Q* devices and their bandwidth is narrow. An antenna tuned for the center of a band may present a high VSWR at the ends of the band, especially on 160 m, where the U.S. amateur band extends ±5 percent from the center frequency. One way around this problem is to make the inductor variable so that slightly different inductance values can be selected for different band segments. On 1.8 Mhz this can be done readily only for the base-loaded version (Fig. 20.1B). For the other configurations, it is easier to use an adjustable matching network or ATU at the base of the antenna.

A continuously loaded antenna has the inductance distributed along the entire length of the radiator (Fig. 20.1D). Typical of these antennas is the helically wound vertical, in which about a half-wavelength of insulated wire is wound over an insulating form (such as a length of PVC pipe or a wooden dowel) with the turns of the coil spread out over the entire length of the insulated support.

I^2R

losses in the winding are somewhat higher than in a straight vertical member because the conducting element is not only longer but often has a smaller cross-sectional area.

Linear loading (Fig. 20.1E) is an arrangement whereby a section of the antenna is folded back on itself like a stub. Antennas of this sort have been built from the same type of aluminum tubing as regular verticals. For 160-m operation, a length of 60 ft for the radiator represents 41 degrees, while a normal

$\lambda/4$

vertical corresponds to 90 degrees. The difference between 90 degrees and 41 degrees of electrical length is made up by the hairpin structure at the base. A potential disadvantage of the linear loading approach is the requirement for an additional insulator

somewhere along the monopole's physical support structure if the antenna is supported from the bottom (as a tower or rigid mast might be), rather than suspended from above (as a wire would be).

The relative cost and mechanical complexity of continuously loaded and linearly loaded verticals for 160 m generally render them inferior to other approaches, especially the top hat of Fig. 20.1C.

20.4 Loading the Tower

Many amateur radio installations include rotatable Yagi antennas on top of an appropriate support structure, such as a ground-mounted guyed triangular lattice tower. Most towers are in the 40- to 90-ft range. This is often considered too short to be an efficient vertical monopole radiator on 160 m. But with proper attention to installation details and a modest set of radials, most such tower and antenna combinations can become excellent transmitting and receiving antennas for 160 m.

To making a "short" tower an efficient radiator on 160 (or 80) m, the apparent electrical length must be increased to approximately $\lambda/4$. This brings the high-current part of the antenna out into the open so as to increase the effective $I \times h$ (or $I \times l$) product. Increasing the apparent electrical length also raises the radiation resistance, thus beneficially altering the resistive divider formed by the radiation resistance and ground losses and improving radiation efficiency. A low-loss ground return path for the RF fields surrounding the vertical must be provided. In the absence of saltwater at the base of the tower, 20 to 60 copper radial wires, each approximately as long as the tower is tall, should be connected through an appropriate anticorrosion metal to the grounded side of the transmission line or base-mounted ATU.

As implied at the beginning of this topic, an excellent way to turn a short grounded tower into an efficient 160-m antenna is to utilize one or more HF Yagi antennas at the top to provide top loading, but there are at least two other ways to do the same thing. If the tower is guyed, allow the top set of guy wires to extend out and down from the top of the tower by a calculated amount before inserting a set of insulators. Three or four dedicated top hat wires can be equally spaced around the horizon at or near the top of the tower, below any rotatable antennas, and extend them out as close to horizontal as possible.

Here are some drawbacks and compromises associated with these approaches:

- Using a portion of each guy wire in the top set of guys is inefficient compared to the horizontal boom and elements of a Yagi because the top guy wires usually follow a fairly steep slope back down alongside the tower itself.
- Use of the top set of guys as a top hat means you should select an appropriate length for the uninsulated sections of guy wire attached to the top of the tower prior to installing the tower. This can be done with an antenna modeling program such as those discussed in Chap. 29.
- Horizontal or sloped lengths of wire at the top of the tower—whether constructed from the top set of guy wires or entirely separate wires—may detune rotatable antennas on the mast if there is only a few feet of separation between the lowest rotating antenna and the wire attachment level.
- In most installations it will not be possible to find supporting termination points for the far ends of any added top hat wires that allow any less of a slope than that encountered when using the guy wires themselves. But additional wires can be added, removed, shortened,

or lengthened at any time, if necessary. Since they are not guy wires, the only mechanical requirement on them is that they be strong enough to support their own weight plus the weight of any rope or cord used to tie them off and hold them away from the tower. When used with a tower surrounded by tall trees, it may be possible to string the top hat wires almost horizontally by using halyards from high in the trees.

- The use of Yagis for top loading is normally limited to HF beams or very long boom VHF/UHF beams, since smaller beams will most likely not provide adequate capacitive loading. Since Yagi dimensions are fixed (with the exception of SteppIR or other adjustable beams), the equivalent electrical length of your tower plus beams is not fully controllable.
- Many Yagi beams do not have their elements directly connected to their booms, so only the boom fully contributes to the top loading effect.
- Insulators and matching networks for some Yagi beams may not tolerate high voltages across their terminals at frequencies far from their intended bands of operation.

As a first approximation to the length of top hat wires or uninsulated guy wire segments at the top of the tower, subtract the tower height from 140 ft. If you have a choice, err on the side of longer total length, rather than shorter. If you are relying on one or more Yagis at the top of the tower, get the turning radius of the lowest beam from the manufacturer's specifications or calculate an equivalent top hat radius from

$$r = \sqrt{b^2 + e^2} \quad (20.1)$$

where *b* is the distance (usually half the total boom length) from the mast to where the elements farthest from the mast are attached and *e* is half the average length of elements nearest the

ends of the boom. Only elements directly grounded to the boom should be considered as fully contributing to the top hat effect.

If it is to ever be used as a radiating monopole, a guyed tower must have at least one insulator for each guy wire attached to the tower. With the exception of any initial lengths established in the top set of guy wires for top hat purposes, the guy wire length from the tower to the first insulator should be less than a few feet. Even though the tower itself may be grounded at its base, failure to insulate each and every guy wire near the point where it attaches to the tower changes the electrical characteristics of the structure, complicates calculations and assumptions, and can reduce the efficacy of the tower as an antenna. General purpose towers can be insulated at the time of installation in case the tower is used as a low frequency antenna.

As discussed in Chap. 11, if a grounded vertical monopole is increased in length beyond $\lambda/4$, its elevation pattern is flattened so that radiation and reception at low vertical angles is enhanced at the expense of higher angles. For most intended uses of an efficient transmitting vertical on 160 or 80 m, this is desirable. If the electrical length exceeds $\lambda/4$, the feedpoint impedance at the base of the antenna will increase and the reactive component will appear as a series inductance, which is easily canceled with a transmitting air variable capacitor in series with the feedpoint.

When the base of a vertical is insulated from ground so that it can be fed directly from the transmission line, the technique is called a series feed. This is most often employed in conjunction with wire verticals. For mechanical simplicity and low cost, towers and rigid 160-m verticals in most amateur installations are grounded at the base. For these, a shunt feed is used.

The impedance at any given point on a grounded vertical monopole increases with height up to $\lambda/2$

. Numerous ways of connecting a feedline to such a structure have been employed. Figure 20.2 shows one of the simplest: an unbalanced output ATU feeds a wire connected to the side of the tower at an appropriate point. This is half of the delta match often employed to provide a balanced feed to the center of dipoles and Yagi driven elements. The height for the tap on the tower depends on the design and tuning range of the ATU. The objective is to find a height where both the resistive and the reactive parts of the impedance seen at the lower end of the wire by the ATU are within its tuning range.

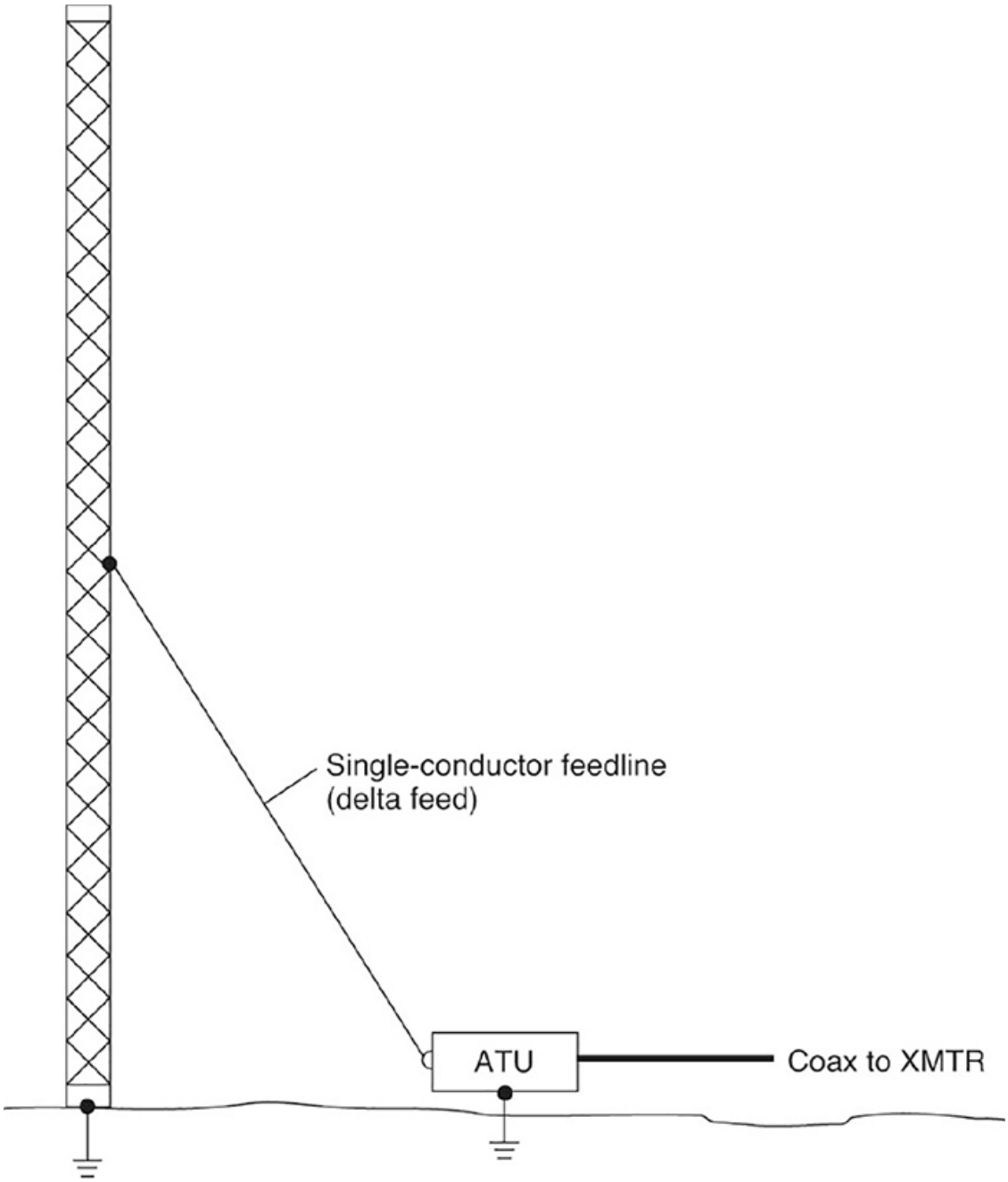


FIGURE 20.2 Delta-matching a grounded vertical tower antenna.

Similar to the sloping wire or half-delta feed of Fig. 20.2 is the shunt feed of Fig. 20.3. This is known as a gamma match. The gamma match is an unbalanced form of the T-match used to feed HF and VHF Yagi-driven elements. The difference between this and

the delta match is that the gamma wire or rod is held at a constant distance from the side of the tower, allowing the ATU and gamma rod to be placed close to the tower. The gamma match is adjusted to provide minimum SWR on the transmission line at the center of the operating frequency range by alternately changing the tap point on the vertical and the spacing of the gamma rod from the radiating element while tuning the series capacitor.

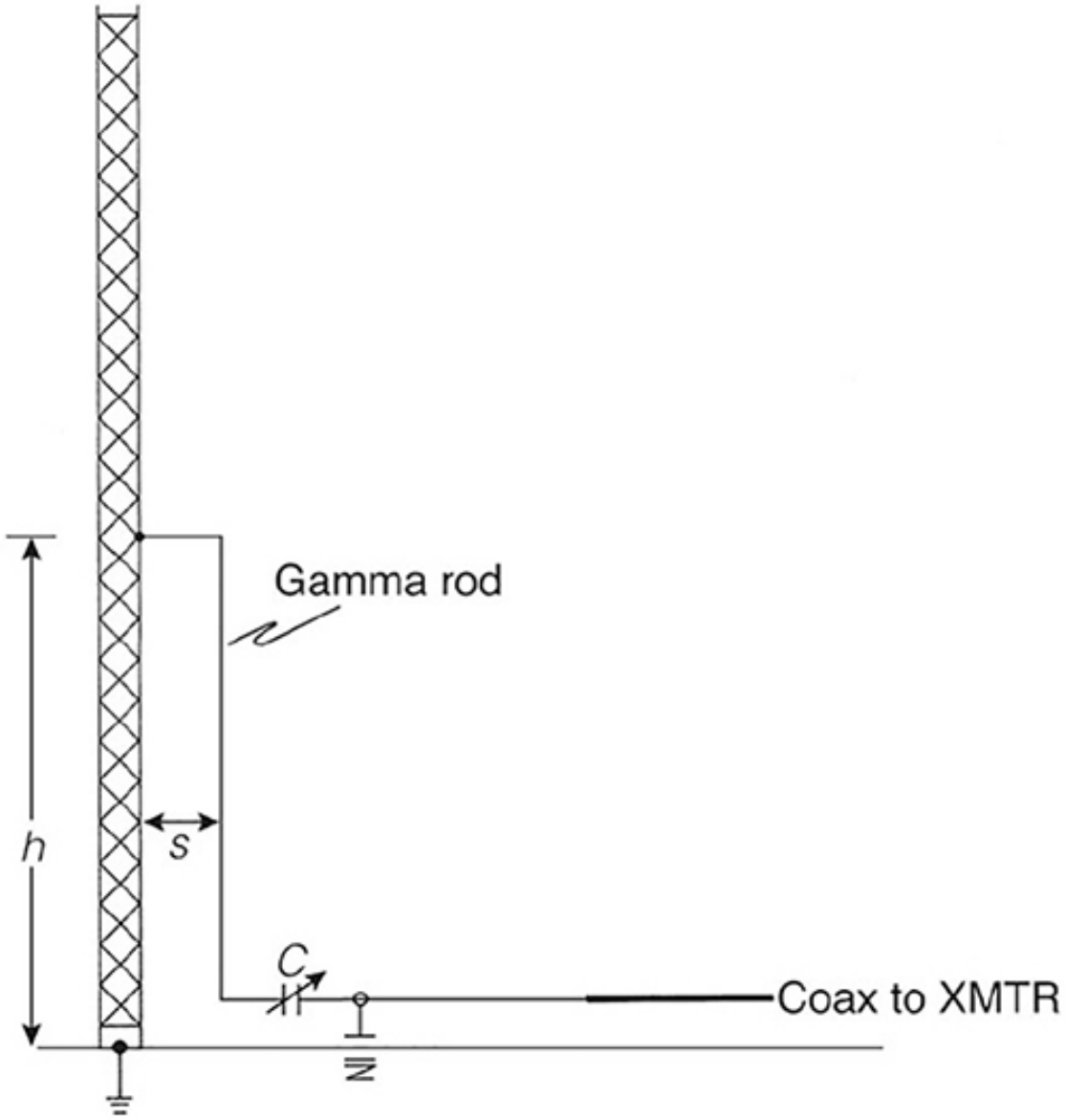


FIGURE 20.3 Gamma-matching a grounded vertical tower antenna.

The shunt feed is similar to an autotransformer, converting the higher impedance of the tap on the tower to the lower impedance of the feedline. If dimensioned properly, the need for a separate ATU can be eliminated or the ATU can be replaced with a single air variable capacitor having adequate spacing between its plates for the transmitter power levels involved. The capacitor is placed in series between the bottom end of the shunt rod and the hot side of the coaxial cable or unbalanced transmission line feeding the antenna.

Antenna modeling software (Chap. 29) provides a way to check tower heights and top loading schemes. The value is in determining the proper length and spacing for the gamma rod or wire. Antenna modeling software can eliminate literally dozens of trips up and down the tower.

As an example, consider a 90-ft guyed tower built with Rohn 45 triangular sections and topped off with a 10-ft heavy-duty mast (Fig. 20.4). At the 92-ft point is a four-element 20-m monobander with a 32-ft boom and full-size elements approximately 36 ft long. Above that, at the 98-ft level, is a two-element 40-m "shorty 40" monobander with a 20-ft boom and shortened (linear-loaded) elements about 42 ft long. All elements of both beams are directly grounded to their respective booms. The tower is grounded at its base and approximately 16 radials of varying length extend outward in all compass directions. Guy wires have their uppermost insulators located 3 ft or less from their respective attachment points to the tower. Since a 98-ft tower is a substantial percentage (70 percent) of a quarter-wave on 160 m, the capacitive top loading provided by the two beams should easily bring the total electrical height to more than 90 degrees or an equivalent height of more than 135 ft at 1.8 MHz. Rather than making dozens of trips up the tower to adjust the gamma rod tap point, EZNEC antenna modeling software can be used to find that a spacing of between 6 and 8 ft in combination with a tap point of about 57 ft optimizes bandwidth at the desired center frequency with a feedpoint impedance very close to 50 Ω resistive and a small amount of inductive reactance to allow

fine-tuning of the VSWR on the feedline with a single air-variable transmitting capacitor at the base.

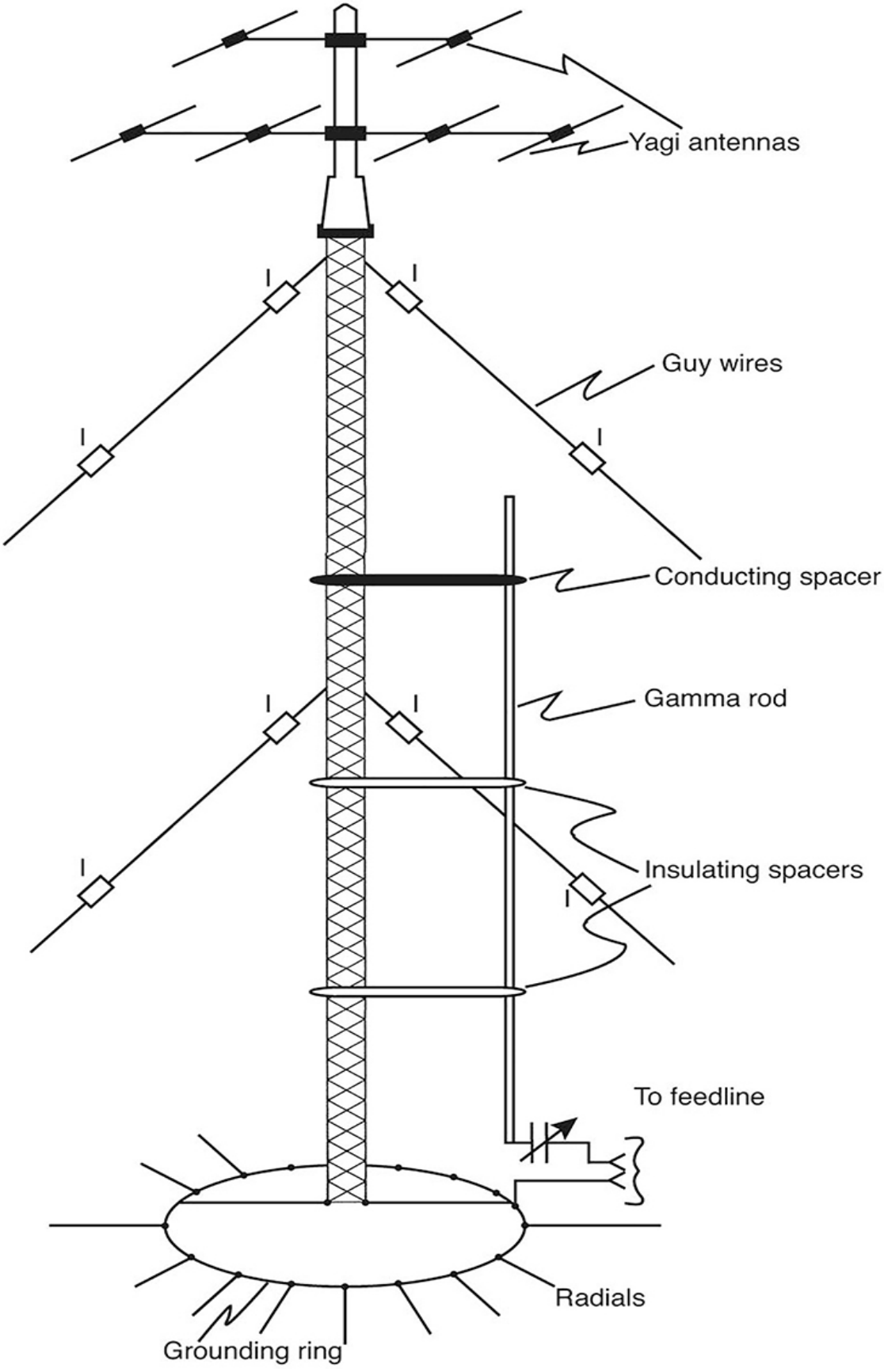

Yagi antennas
Guy wires
Conducting spacer
Gamma rod
Insulating spacers
To feedline
Radials
Grounding ring

FIGURE 20.4 Top-loaded 90-ft tower.

For this example design, an 8-ft length of sturdy aluminum angle stock from a local hardware store supports the gamma rod (made from lengths of copper plumbing tubing and couplings) 7 ft from the side of the tower and provides the horizontal electrical connection between the gamma rod and the tap point on the tower. Two 10-ft lengths of scrap

$$1\text{-} \times 2\text{-in}$$

lumber act as insulated spacers holding the gamma rod 6 ft away from the nearest edge of the tower at the 20 and 40 ft heights. The matching network consists of a single 200-pF transmitting air variable capacitor connected in series between the bottom end of the gamma rod and the center (signal) conductor of the 52-Ω coaxial feedline. With this design, the antenna tunes easily and has a much broader SWR bandwidth (under 2.0:1 from 1.800–1.877 MHz, with a minimum of 1.43:1 at 1.827 MHz). In on-the-air use in competitive activities such as DXing and contests, this top-loaded shunt-fed tower with its 16 radials of various lengths is a superior performer that is usually only beaten by arrays of multiple verticals and ocean-front station locations.

To summarize for vertical antenna towers,

- When cost, strength, and mechanical complexity push strongly toward an uninsulated ground-mounted vertical, consider the use of a gamma match.
- Break up conducting guy wires holding up a tower or rigid vertical with insulators.
- When possible, use some form of top loading in preference (or in addition) to bottom, middle, or linear loading. The objective is to maximize signal strength by bringing the natural high-current portions of the antenna up out of the feed-line as much as possible.

- The vertical monopole needs a low-resistance RF ground system. The shorter your vertical is, relative to a full quarter-wave, the more important your ground system becomes.
- Consider whether elements are grounded to the boom when selecting beams for the higher HF bands if there is any chance the tower they are going to be mounted on will be used as a vertical monopole for the low bands.
- Inexpensive or free antenna modeling software can help in the design process.

20.5 Insulated-Base Loaded Verticals

In some situations, insulating the tower base or feedpoint from ground may be necessary or desirable. This can be needed when an array of vertical monopoles used to provide greater forward gain and/or front-to-side (F/S) or front-to-back (F/B) ratios than can be obtained from a single (omnidirectional) radiating element. As discussed in Chap. 13, getting full performance from a multielement array requires control of both the magnitude and the phase of the feedpoint current in each element of the array. This control is easier to maintain when the feedpoint is not grounded—that is, when series feed is used.

A loaded insulated-base vertical is essentially one side of a shortened dipole or one side of the driven element of a beam that uses loading coils or some other technique to shorten the elements. The performance of such a vertical will be no better than the conductivity of the ground system beneath it. Unless the vertical is mounted in seawater, radials are necessary.

Radial wires lying on the ground or in very close proximity to ground are not resonant wires; there is no one specific length that is optimum. In general, 20 or more radials between

$\lambda/4$

and $\lambda/8$ in free-space length will provide an excellent ground for a vertical monopole, but there is nothing magic about either the

length or the number. As many as possible should be used, and if they have to be shorter than normal to fit a backyard, so be it.

As with the grounded vertical, loading of an insulated vertical can take many forms. The simplest and most efficient are capacitive top hats with or without a small amount of base loading inductance.

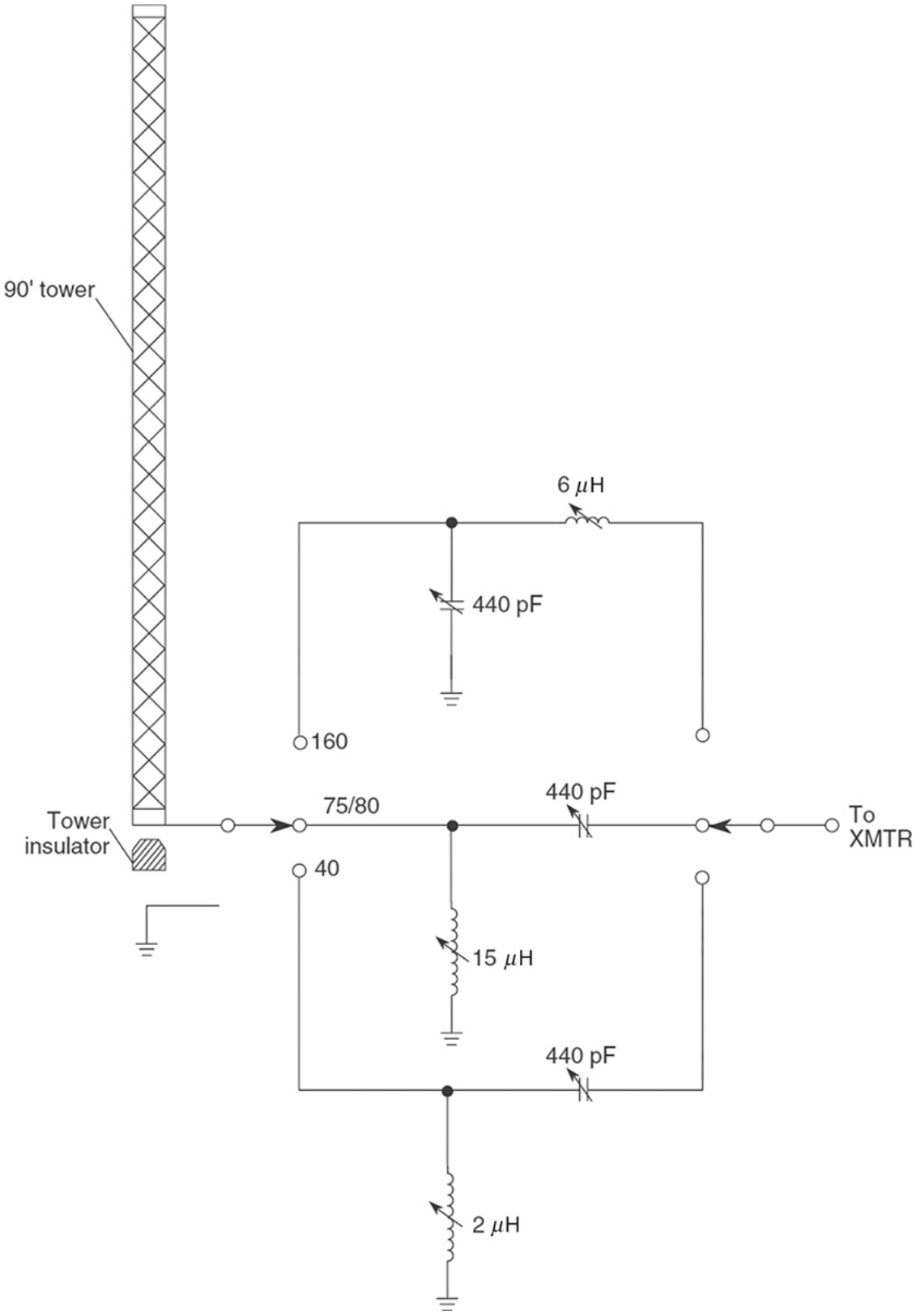
Up to an effective electrical height of 5⁄8*λ* (or 225 degrees), increased height results in a modest increase in the low-angle radiation of a vertical monopole, obtained at the expense of high-angle radiation. This additional gain reaches its maximum at 5⁄8*λ*, beyond which the low-angle gain gradually decreases. A monopole having an equivalent electrical length of

$\lambda/4$

or 90 degrees on 160 m can do well on 80 m with attention to matching its high impedance, but will suffer reduced performance on 40 m unless a phasing section is added such that the physical structure acts like a collinear array.

In an array of multiple vertical elements, each element should have its own radial field and the fields should be as similar as possible. If radials from different elements overlap, it may be useful to connect them together, but the possible improvement in ground conductivity may be limited. Radials that overlap should not be allowed to intermittently touch or develop oxidized connections. Either connect them together firmly or insulate them from each other.

Figure 20.5 shows a situation in which a 90-ft tower is insulated from ground and operated on 160, 80, and 40 m. The feed is an L-section or reverse L-section ATU. On 40 and 75/80 m, the feedpoint impedance of the tower is inductive, so a series capacitance and shunt inductance are used for the ATU (i.e., reverse L-section coupler). On 160 m, the feedpoint impedance is capacitive, so the inductor and capacitor are reversed.



F IGURE 20.5 Methods for matching a series-fed tower.

For insulated towers, the insulator section is not limited to being located at ground level. On 160 m, elevated-base ground-plane antennas can be used. One advantage is that fewer radials are needed. If the base of the ground plane is $\lambda/8$ or higher above earth ground, between four and eight elevated radials will provide efficiencies comparable to 20 or more radials lying on the ground for a vertical monopole whose insulated base is at ground level. The disadvantage is that the overall height of the tower increases by as much as $\lambda/8$, or 65 ft (see Chap. 33 for a clever alternative).

Finally, when considering any form of insulated-base vertical, insulated-base towers are more expensive and lightning protection is simpler with a grounded-base tower.

20.6 Phased Verticals

For topband (160 m) users, arrays of phased verticals can add transmit gain in switch-selectable compass directions. To maximize radiation efficiency and bandwidth, the array elements are $\lambda/4$ tall or higher, but arrays of shortened verticals have merit as well with efforts to minimize losses in ground systems and matching units.

An array used on 160 m (and 80 m) is the four-square (also written 4-square). These arrays are base-insulated and series-fed. Some amateurs install them as much as $\lambda/8$ above ground and use elevated radials. Whether elevated or ground-mounted, a 160-m four-square of $\lambda/4$ elements is a marvel to behold!

Typically, feedpoint currents are phased as 0 degrees in the rear element, −90 degrees in the two side elements, and −180 degrees in the front element (i.e., the element nearest the desired direction for maximum gain), but many variations in both element phasing and interelement spacing to optimize one or another parameter

exist. The definitive compendium is "ON4UN's Low-Band DXing" by John Devoldere, available from ARRL.

Four-squares are all-driven arrays, but parasitic arrays are equally useful and often quite a bit less expensive or time-intensive to install. Arrays of full-size

$\lambda/4$

elements at 160 m can be impractical for most users, but there are more accessible alternatives. Chief among those—whether employed in an all-driven or a parasitic array—is the inverted-L, described later in this chapter.

The use of directional arrays for transmitting antennas also adds value on receive. Tests comparing full-size four-squares with dedicated receive antennas, such as 1λ to 2λ Beverages, suggest that these transmitting arrays are usually as good as or better than the specialized receiving antennas.

20.7 Horizontal Antennas

Although subject to the limitations discussed at the beginning of this chapter, a horizontal dipole can be an inexpensive solution to the problem of putting out a decent nighttime signal on 160 m, especially if tall trees are available. There are at least four variations on a dipole: fully horizontal, sloping, inverted-vee, or bent.

Horizontal dipoles are good for distances out to 1500 mi or so, beyond which a single top-loaded vertical may be better. With the appropriate ATUs capable of handling high feedpoint impedances when operating a dipole on its second harmonic, the antennas are superb performers on 80 m, where they exhibit perhaps 2 dB more gain in the main broadside lobe than a conventional

$\lambda/2$

dipole.

For limited spaces, bent dipoles can be better than inverted vees, for reasons described in detail in Chap. 8. The former antenna requires two high supports, however, while the latter needs but

one. With only one high support but no particular space limitations, the sloping dipole is often a better choice than the inverted-vee.

20.7.1 Shortened Horizontal Antennas

Strategies that work for loading short vertical antennas can also work for horizontal antennas. In the horizontal case the goal is to make the antenna appear electrically to be a half-wavelength balanced antenna rather than a
$\lambda/4$
unbalanced antenna. One practical form of shortened horizontal antenna uses inductive loading in the middle of each leg, either with loading coils or with traps that allow use of the same antenna on 80 m.

20.8 Inverted-L Antennas

One of the simplest wire antennas, but one that can produce excellent results, is the
$\lambda/4$
inverted-L (Fig. 20.6). Typically considered when there is no chance of having a high enough support to hang a full
$\lambda/4$
wire vertically over its entire length, this antenna is simplicity itself.

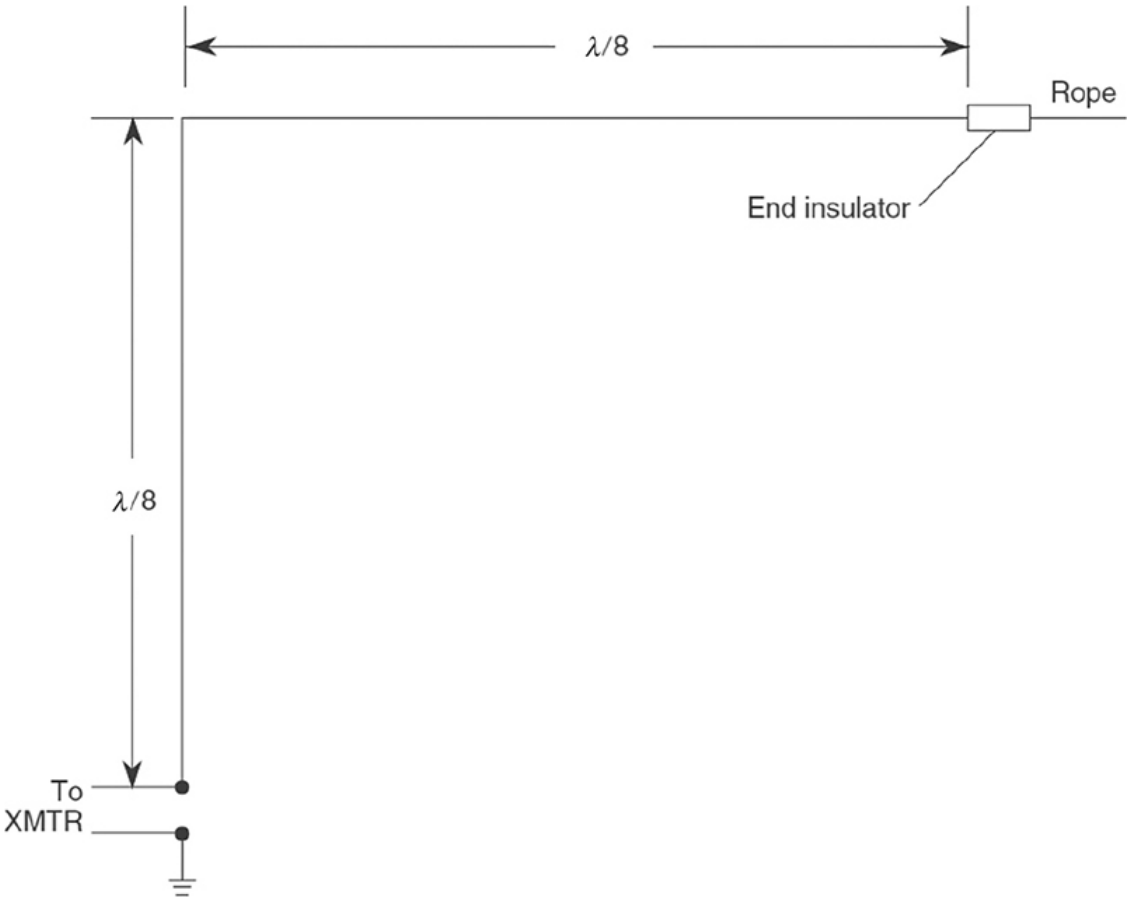


FIGURE 20.6 The inverted-L.

Starting as close to the antenna coupler (ATU) or transmitter as is reasonable, run as much of a 170-ft piece of wire vertically as high as you can. Then run the remainder of the wire horizontally until the full 170 ft have been deployed. This might result in a 65-ft vertical section and a 105-ft horizontal section, but wide variations from these lengths are often found. Even vertical sections as short as 30 ft have merit.

Making the total wire length somewhat longer than $\lambda/4$ on 160 m ensures that the natural high-current portion of the standing wave is on the vertical portion of the wire rather than trapped inside the feedline, and partly to boost the radiation resistance. This also makes the input impedance slightly inductive so that the feedpoint reactance can be tuned out easily with a

series air variable transmitting capacitor. Fed as described, the antenna should exhibit a reasonable match to 52- or 75-Ω coaxial cable.

The inverted-L is a cousin to a quarter-wave vertical monopole. Some prefer to look at it as a bent vertical, while others see it as a short vertical with asymmetrical top loading. If constructed as shown in Fig. 20.6, a small portion of the transmitter power will be radiated from the horizontal half of the wire, providing some high-angle signal useful for working stations within a few hundred miles. Because the horizontal portion of the wire extends in a single direction from the vertical portion, it causes a slight distortion in the azimuthal radiation pattern of the antenna. If you don't have a tower you can shunt-feed, but you need a good antenna for 160 m that will work for both distant and close stations, the inverted-L is hard to beat.

Because it is a member of the family of vertical monopoles, the inverted-L requires an adequate ground system. Chapter 33 and the beginning of this chapter cover RF grounding in more detail. For an inverted-L, run 12 to 50 radials of 50 to 150 ft, depending on the available space. A rule of thumb is to put down as many equally spaced radials as necessary to have their far tips be no more than about 0.05λ apart. On 160 m, that's about 8-ft spacing between tips. As explained in Chap. 11, the longer you can make your radials, the more you should use, and the higher your overall antenna and ground system efficiency will be. If property lines, driveways, or buildings keep you from making all the radials the same length, radials in proximity to the ground are not resonant, so their exact length is immaterial.

The center of the radial circle should be at the feedpoint of the inverted-L. Tie the inner end of all the radials to the shield side of the coaxial transmission line using a chassis-mount SO-239 UHF receptacle or whatever best matches the connector on your feedline. Cover the joint with electrical tape and clear acrylic spray and cover it with a waterproof container. If you are using a matching capacitor, cover that as well. One popular method of constructing the inverted-L employs an existing tower for the

vertical section and a length of wire for the horizontal section. If you already have a 60-ft tower to accommodate a beam antenna used on higher frequencies, it is relatively easy to turn it into an inverted-L for 160 m. If the base of the tower and any guy wires are insulated, connect the shield of the coaxial feedline and the tips of the radials together at the bottom end of the insulator and connect the center conductor of the feedline or one end of the matching capacitor to one leg of the tower above the insulator. Never connect copper wire directly to a galvanized tower. See Chap. 32 for additional information.

If the tower is not insulated from ground, run a wire up the side of it using aluminum angle-stock with insulators, rigid PVC pipe, or wood to hold the vertical portion of the wire at least a few feet away from the tower. Depending on its natural electrical length, the tower may affect the tuning of the inverted-L, but that's not necessarily bad. The effect of the tower may vary as feedlines to HF Yagis or other antennas at its top are switched in and out at their far end in the radio room.

Success with the inverted-L depends more on the quality of its system of radials than to the exact height of the vertical section of radiator. The inverted-L can approach the performance of a full-size $\lambda/4$ monopole, but only if these two factors are optimized:

- The radial field must be extensive, so as to minimize losses in the earth beneath it. This is more important than for a full-size $\lambda/4$ or taller vertical radiator because the radiation resistance of an inverted-L is typically lower (20 to 25 Ω) than that of the taller (37 Ω) vertical, so it takes a lower ground resistance to dissipate no more transmitter power in ground losses than in the case of the full-length $\lambda/4$ radiator.
- The total length of wire in the inverted-L should be at least 150 ft in order to get the high-current portion of

the standing current out of the feedline and onto the open vertical portion of the antenna.

20.8.1 Dual-Band Inverted-L

A 160 m wire antenna can also be used for 80 m with appropriate feeding. The problem with trying to use the 160 inverted-L on its second harmonic is that the impedance at the feedpoint near the ground is very high. This can be matched with a good parallel-tuned ATU or a step-up transformer, but the system is likely to have poor bandwidth.

Another problem with using the antenna on its second harmonic is that more of the transmitter power (roughly half) is going to be radiated from the horizontal section of wire. The total wire is a half-wavelength on 80, and the current maximum is at the midpoint of the wire, at the point where the vertical and horizontal sections come together.

A work-around that minimizes both problems is to insert an 80-m trap at the top of the vertical section or a slight distance out on the horizontal section if the total vertical run is less than 60 ft or so. The trap will permit some shortening of the horizontal top section, allowing the total wire length to be less than 150 ft.

20.8.2 Phased Inverted-L Antennas

Because it is inexpensive and normally series-fed, the inverted-L is an excellent candidate to be the basic element or building block of an all-driven 160-m phased array. Considerations in earlier chapters for ensuring proper and stable phasing pertain to the inverted-L.

Inverted-Ls make good building blocks for parasitic arrays, too. An inline array of three inverted-Ls is similar to one side of a vertically polarized three-element Yagi, the other side of each pseudo-Yagi element being replaced by ground. Using interelement spacings of 0.15λ results in a total baseline (or boom equivalent) length of about 170 ft. Forward gain and F/B ratio are optimized by adjusting the vertical portions of the reflector and director, possibly with the aid of an antenna modeling program. Unlike the horizontal

Yagi this array is derived from, a good radial system is required under each element for maximum efficiency and stability.

20.9 Flat-Top or T Antenna

For distant links, we want to minimize radiation in the vertical direction. One way to do that, space permitting, is to convert the "inverted-L" into a "T" by extending another horizontal wire, the same length as the existing horizontal segment, in the opposite direction (Fig. 20.7). Whatever standing wave of current exists in the first horizontal wire will be balanced by an equal current in the opposite direction in the new wire.

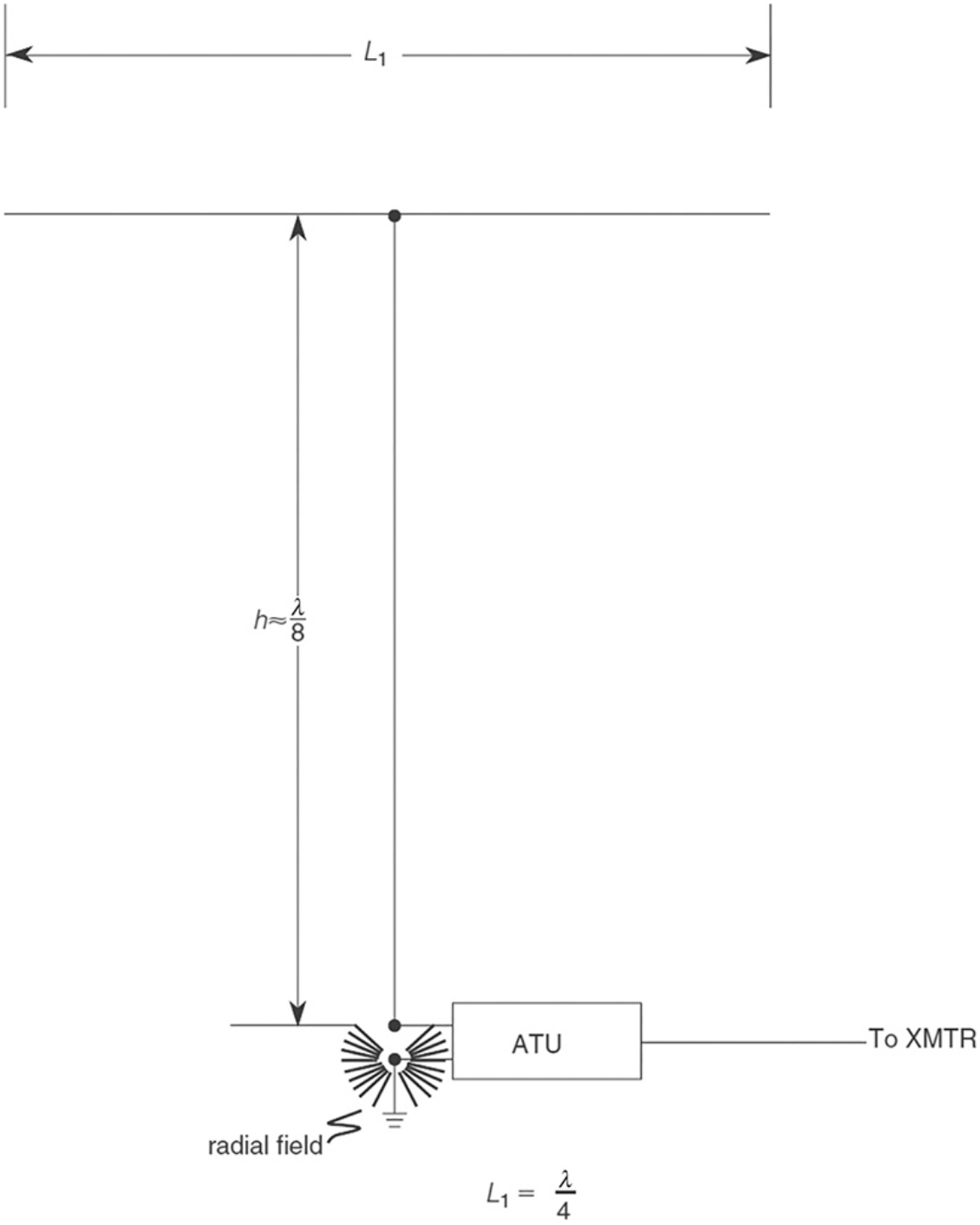


FIGURE 20.7 Flat-top or T antenna.

At distant points nearly straight up (meaning the ionosphere directly above), the total signal from the antenna will approach zero. Nearly all the radiated power emanates from the vertical portion of the wire, increasing its low-angle performance, and

coincidentally eliminating some of the azimuthal pattern skew introduced by the single horizontal wire. The flat-top reduces high-angle radiation on both 160 and 80 m. The radial system can be the same as that for the inverted-L.

The downside to the flat-top is that it requires twice as much span for the horizontal wires than the inverted-L does. If available space in the long dimension is an issue, the two sides of the flat-top portion of the T can be bent or run in zigzag fashion. To minimize high-angle radiation, such bending can be applied to both sections of wire symmetrically.

20.10 80-Meter Dipole on 160 m

Although not optimal, an 80-m dipole can be used on 160 m with some success. There are two ways to do this.

The first approach requires an ATU capable of tuning a wide range of reactance on 160 m. Match the antenna to the transmitter on 160 m by tuning the ATU for minimum SWR at the chosen 160-m operating frequency. Even though the dipole is too short to fully expose the high-current portions of the natural radiating standing wave, there's enough wire length to get signal out of the feedline and on its way.

Drawbacks are that the antenna feedpoint impedance is highly reactive and a poor match to any available transmission line and a suitable wide-range ATU can be expensive. The line will experience high standing waves with increased

$I^2 R$

losses that are characteristic of high SWR. Open-wire line may be the best feedline for this application because it has less added loss in the presence of high SWR than coaxial cable or twin-lead does, and it is less likely to break down because of the high voltages encountered. On the other hand, for a few hundred watts or less, any RG-8 or equivalent cable should be able to handle the voltages.

The second approach is to use an existing 80-m dipole and its feedline as a pseudo-flat-top or T-antenna on 160 m. This is

attractive if you have a reasonable feedline length (say, more than 60 ft) and if the center of your dipole is substantially higher than your radio room so the feedline goes not only horizontally (to get from indoors to outdoors) but up as well. This provides a vertical component to the feedline path that can serve as the vertical section of a rough inverted-L.

For this approach, disconnect the transmitter end of the feedline from its normal connection, tie the bare ends of the two feedline conductors together, and reconnect the joined ends to the hot (center) terminal of the SO-239 coaxial socket on your ATU or transceiver (Fig. 20.8). Tune the ATU for minimum SWR. If you are using twin-lead or open-wire line, twist the indoor ends of the two sides of the feedline together and stuff them back into the center (or "hot") receptacle of the SO-239. If you are using coaxial cable, build up a small adapter consisting of an SO-239 socket and either a wire pigtail (to push into the SO-239 socket) or a PL-259 plug attached via a short wire to the added SO-239.

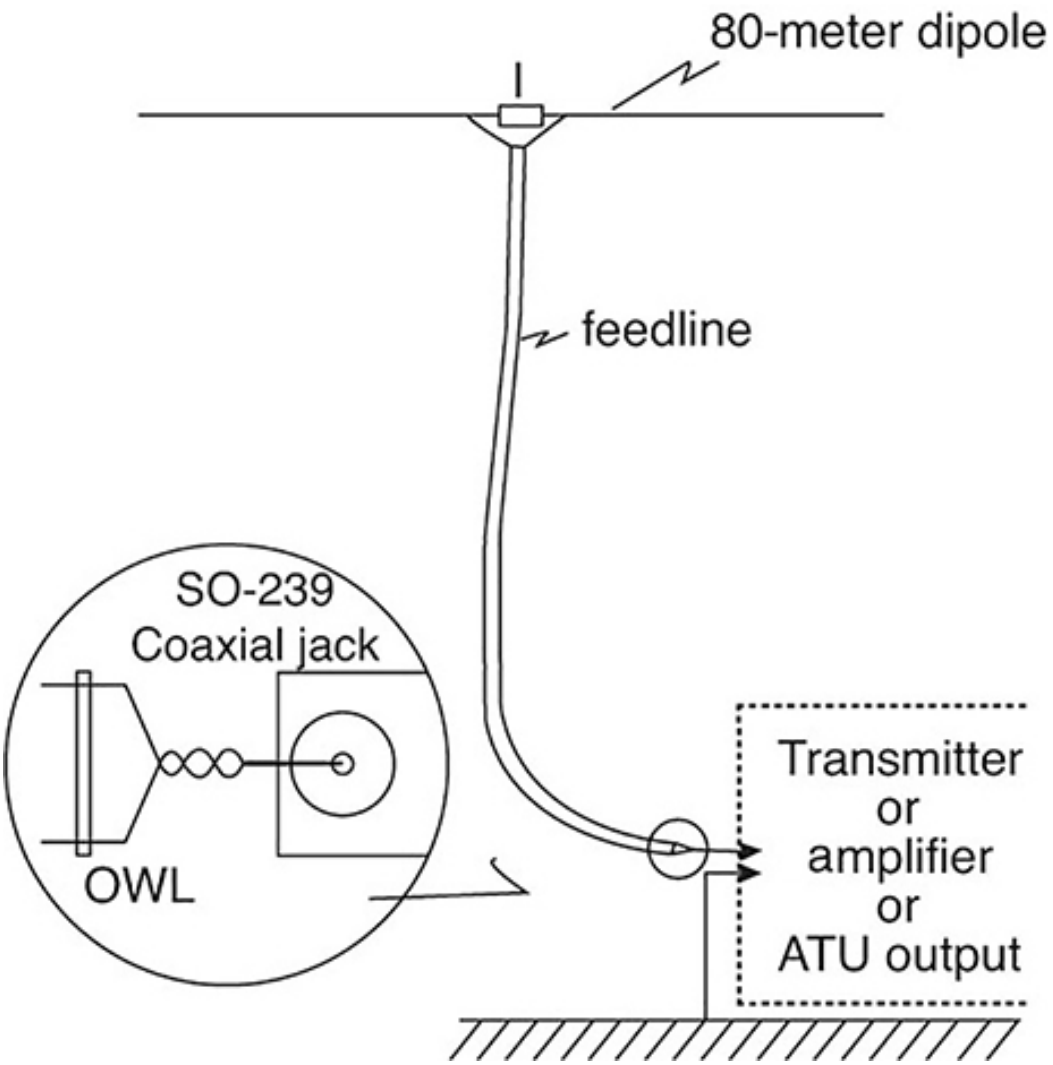


FIGURE 20.8 Pseudo-T from an 80-m dipole and feedline.

A good RF ground system is a necessity if this antenna is to be effective. If the radio equipment room is above the first floor of the building, this can be a challenge. Without a ground system, not only will the antenna fail to perform but the radio equipment will be "hot" with transmitted RF and capable of causing painful burns. At the very least, a single

$\lambda/4$

counterpoise wire must be attached to the chassis at the point where the new feedline is connected. Use insulated wire (except where it connects to the chassis) and try to get it outdoors and close to earth in as short a run as possible.

One negative of this approach is that you have to disconnect and reconnect the 80-m feedline whenever switching from 80 to 160 m, or vice versa. And this technique works only if neither side of the feedline is grounded between the transmitter SO-239 and the antenna. If all your coaxial feedlines go through a remote switch or lightning arrestor connected to ground, your signal will be shorted to ground before it gets to the antenna.

As described in the preceding T-antenna paragraphs, radiation from the two sides of the dipole mostly cancels at high elevation angles. Because the feedline in this example travels both horizontally and vertically on its way to the dipole center, it generates both vertically and horizontally polarized radiation. As a result, it's not possible to use this scheme to completely eliminate high-angle radiation.

The primary radiation from this antenna comes from the feedline, not the 80-m dipole. The purpose of the dipole is to provide a total length ($\text{feedline} + \text{one side of the dipole}$) that is close to $\lambda/4$ on 160 m, and to hold up the far end of the feedline.

20.11 Receiving on the Transmit Antenna

If your station location enjoys a low background noise level on 160 m, your transmit antenna may well be your best receiving antenna. When transmitting arrays with higher gain are unavailable or noise is too high, better antennas for detecting weak signals may be necessary. Chapter 16 describes a variety of receiving antennas including Beverages and small loops that can be much better. Alternatively, noise cancellation techniques can be used, and hardware is commercially available to do this for many applications.

CHAPTER 21
VHF and UHF Antennas

The VHF band extends from 30 MHz to 300 MHz, and the UHF band is defined as 300 MHz to 1 GHz by the Institute of Electrical and Electronics Engineers (IEEE) and 300 MHz to 3 GHz per the International Telecommunication Union (ITU).

Shorter wavelengths translate to smaller element dimensions. VHF and UHF antennas are often less expensive and simpler to construct than their HF counterparts, and their supports don't need to be as strong. Shorter wavelengths may allow the use of free-space design rules. A 2-m antenna atop a 70-ft tower is 10 wavelengths above ground, far enough away to model the behavior of the antenna as if it were in free space. Yagis and other antenna types that use aluminum tubing or rods for elements tend to deliver performance over a wider percentage bandwidth than at HF because the element diameter is much larger relative to wavelength.

At higher frequencies, arrays with more elements and hire directivity become feasible. Directive antennas with more gain, narrower beamwidths, and better rejection of signals off the sides and rear are relatively easy to implement. A 14-element Yagi for 144 MHz is a lot easier to put up and keep up than even a three-element Yagi for 14 MHz. The higher the forward gain of the antenna, the narrower its main beam and the more precise your ability to aim it must be. With higher directivity comes the need to point the antenna mechanically. In some cases, two rotors (called an el-az mount) are used.

Smaller element dimensions mean that greater precision and accuracy in assembly are required. Small hardware items, such as element-to-boom mounting brackets, become a critical part of each element's electrical characteristics and can affect the assembled

antenna compared with predictions based on models or calculated values.

Higher frequency radio waves experience greater blockage effects than lower frequencies. Most VHF and UHF propagation modes demand that your antennas be as high as possible, to attain line-of-sight paths extending as far from your site as possible before encountering the first obstacle. For VHF and UHF, a high location such as a hilltop is advantageous.

Caution *A 14-element 2-m Yagi is so lightweight that it can be carried up the tower in one hand by one person. But long-boom beams, even at VHF/UHF, have a relatively high "windsail area," and even relatively light winds can apply a lot of force or torque to them.* ***Install antennas with a helper and use hoists and other tools to actually lift and support the array while it is being attached to the mast or rotor***.

To speak of a "VHF" or "UHF" antenna is somewhat misleading because virtually all forms of antenna used on the HF or MW bands can also be used at VHF and UHF. The main factors that distinguish VHF/UHF designs from others are mechanical: Some things are simply much easier to accomplish with small antennas. For example, consider the delta match. At 80 m, the delta-match dimensions are approximately

$36 \times 43 \text{ ft}$

, and at 2 m they are

$9.5 \times 12 \text{ in}$

. Clearly, delta matching is a bit more practical at VHF.

Although VHF/UHF antennas are not substantially different from HF antennas, for various practical reasons there are several forms that are especially well suited to the VHF/UHF bands. In this section we will take a look at some of them.

21.1 Ground-Plane Vertical

At VHF and above, the ground-plane antenna is a vertical radiator situated immediately above an artificial RF ground consisting of quarter-wavelength radiators or a sheet of highly conductive material such as copper. Most ground-plane antennas are either 1/4-or 5/8-wavelength. For the latter case impedance matching is needed (see the example later in this chapter).

Figure 21.1 shows how to construct an extremely simple ground-plane antenna for 2 m and above. The base of the antenna is a single SO-239 chassis-type coaxial connector. Use the type that requires four small machine screws to hold it to the chassis, not the single-nut variety.

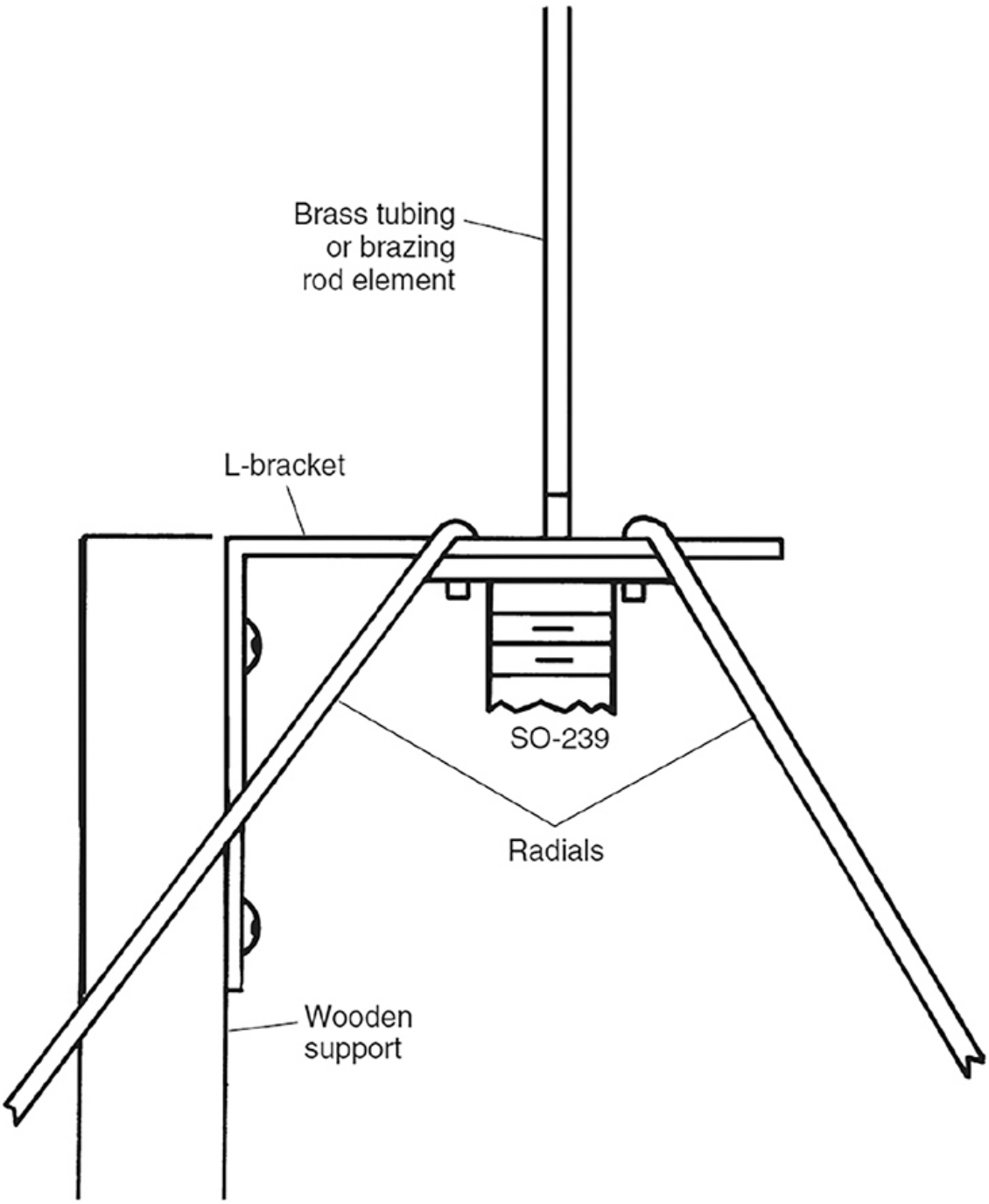


FIGURE 21.1 Construction of VHF ground-plane antenna.

The radiating element is a piece of 3/16-in or 4-mm brass tubing. These sizes fit over the center pin of an SO-239 with only a slight tap from a lightweight hammer. If the inside of the tubing and the connector pin are pretinned with solder, sweat-soldering the joint

will make a good electrical connection that is resistant to weathering. Cover the joint with clear spray for added protection.

Tubing or solid rods is used for the radials. At least three radials equally spaced around a circle are needed. (Only two are shown in Fig. 21.1.) For this method of construction, four might be a better number, with one radial attached to each of the four SO-239 mounting holes. Flatten one end of the radial with a hammer and drill a small hole in the center of the flattened area. Mount the radial to the SO-239 using hardware of the appropriate size (typically 4-40 screws and nuts for this particular SO-239).

The SO-239 can be attached to a metal L-bracket that is mounted to a length of

$2 \times 2\text{-in}$

lumber that serves as the mast. While it is easy to fabricate such a bracket from aluminum stock, it is also possible to buy suitable brackets in any well-equipped hardware store.

21.2 Coaxial Vertical

The coaxial vertical is a quarter-wavelength vertically polarized antenna that is popular on VHF/UHF. It is really a

$\lambda/2$

dipole, since the exposed quarter-wave center conductor is working against an equal length of exposed shield or braid in line with it, rather than in a ground-plane configuration. The method of construction forces the feedline to come away from the antenna inside the braid. This is an asymmetrical arrangement that has no particular merit and may benefit from common mode chokes at the bottom of the antenna.

Two varieties of this antenna are usually encountered, depending on how permanent it is supposed to be. In Fig. 21.2, we see the coaxial antenna made with coaxial cable. Although not suitable for long-term installations, it is useful for short-term, portable, or emergency applications. The length of either the exposed insulation or the exposed shield is

$$L(\text{in}) = \frac{2952}{f(\text{MHz})} \quad (21.1)$$

The antenna can be suspended from above with a short piece of cord or line. With this implementation the coaxial cable begins to deteriorate after a few rainstorms because the center of the dipole, where the braid and the center conductor go in opposite directions, is exposed to the elements. This effect can be slowed by sealing the end and the break between the sleeve and the radiator with silicone RTV or caulk.

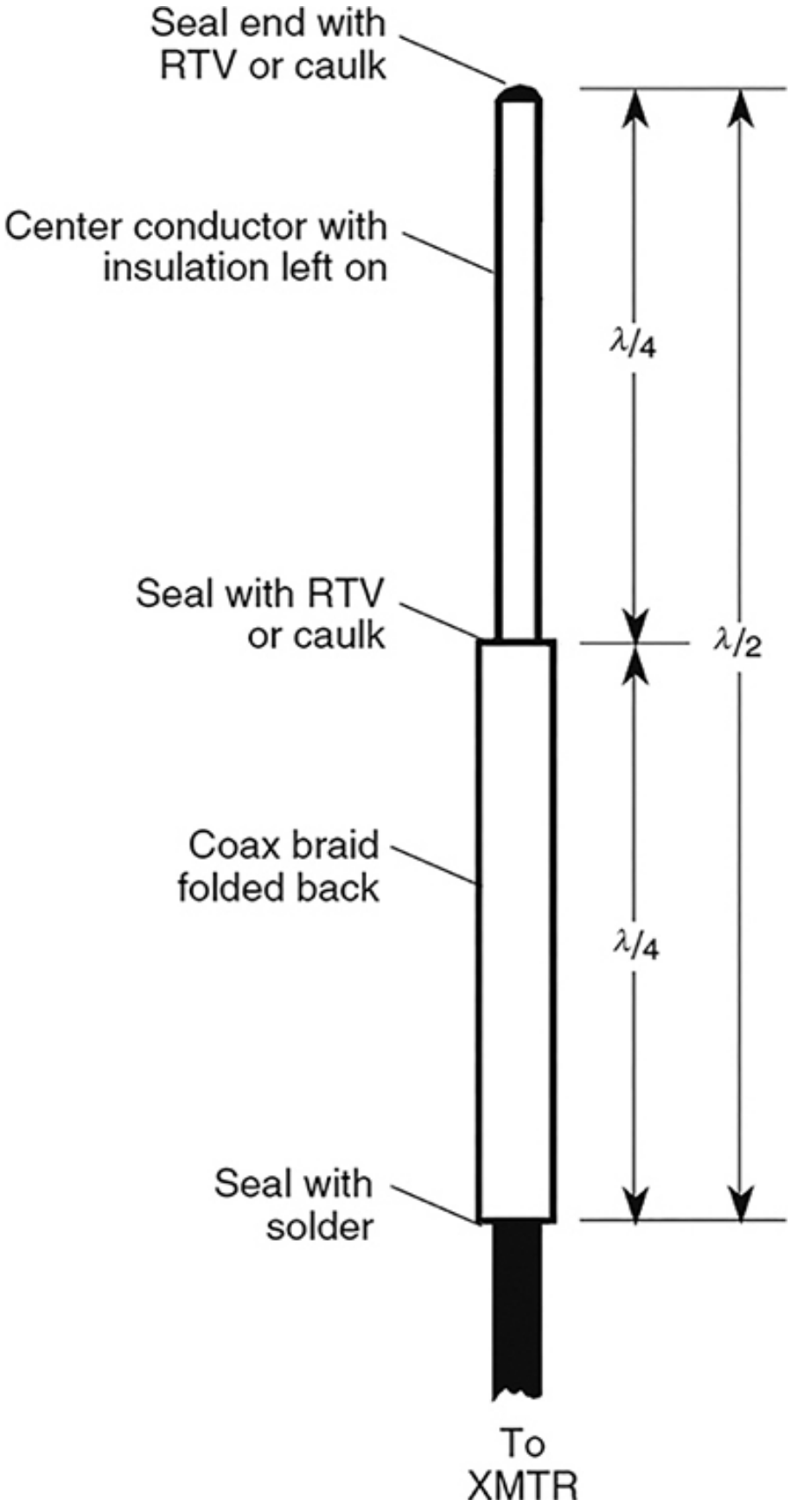

Seal end with
RTV or caulk
Center conductor with
insulation left on
λ/4
Seal with RTV
or caulk
λ/2
Coax braid
folded back
λ/4
Seal with
solder
To
XMTR

FIGURE 21.2 Coaxial vertical based on coaxial cable.

A weather-resistant implementation is shown in Fig. 21.3. The sleeve is a piece of copper or brass tubing (pipe) about 1 in diameter. An end cap is fitted over the end and sweat-soldered into place to prevent the weather from destroying the electrical contact between the two pieces. An SO-239 coaxial connector is mounted on the end cap. The coaxial cable is connected to the SO-239 inside the pipe, which requires making the connection before mounting the end cap.

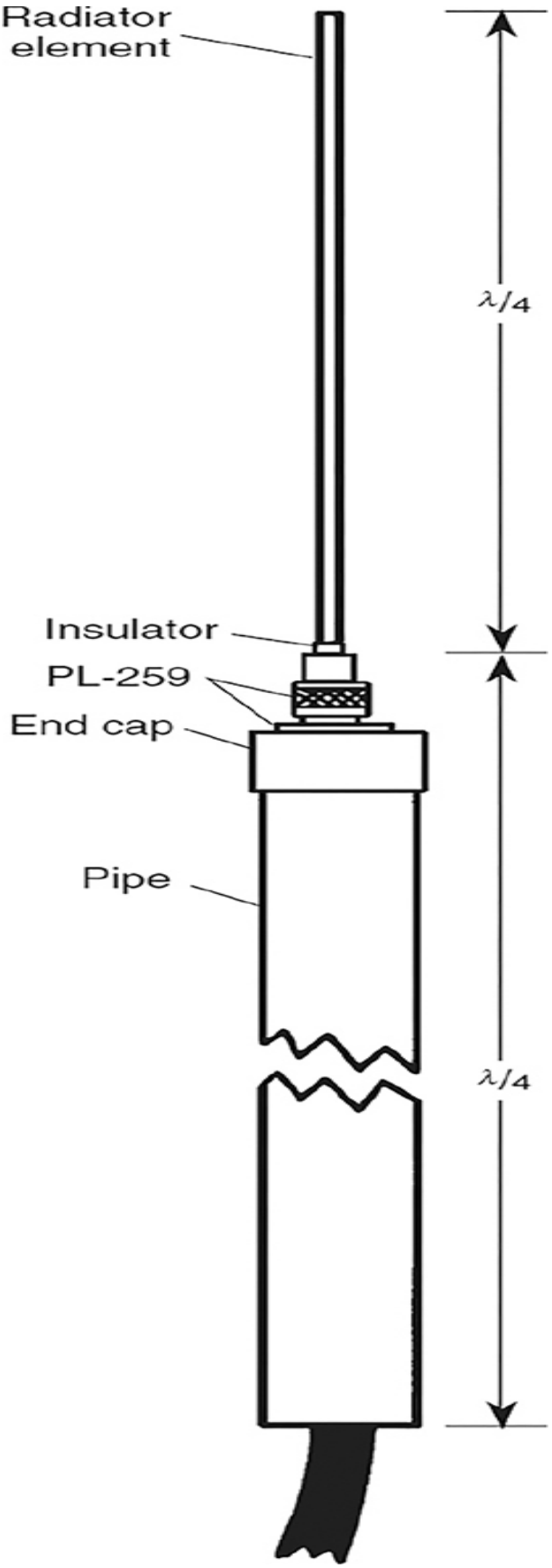

Radiator element
λ/4
Insulator
PL-259
End cap
Pipe
λ/4

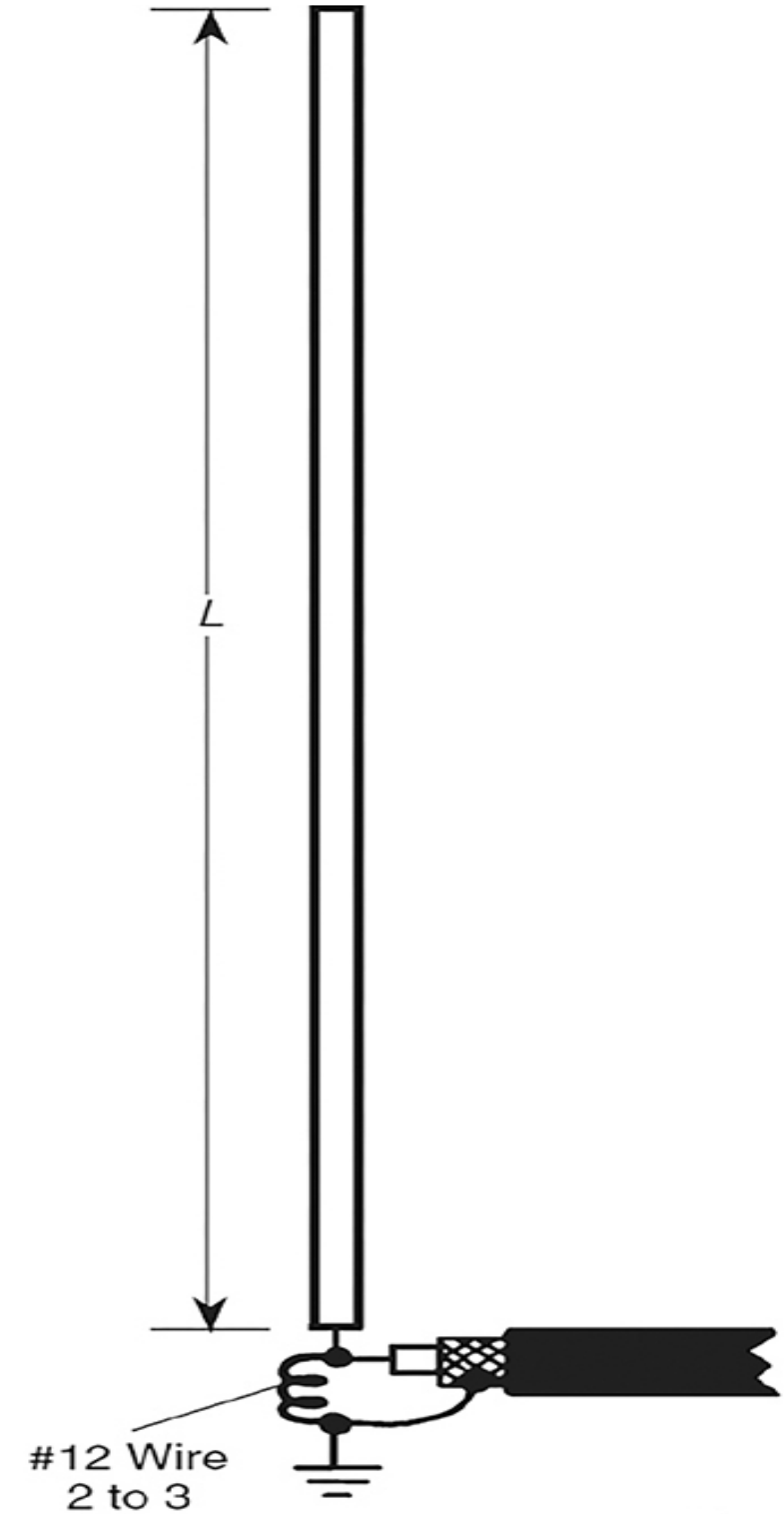
FIGURE 21.3 Tubing coaxial vertical.

The radiator element is a small piece of tubing or rod soldered to the center conductor of a PL-259 coaxial connector. An insulator is used to prevent the rod from shorting to the outer shell of the PL-259. An insulator salvaged from the smaller variety of banana plug can be shaved a small amount with a fine file and made to fit inside the PL-259. It allows enough center clearance for 1/8-in or 3/16-in brass tubing. Alternatively, the radiator element can be soldered to a banana plug. A standard banana plug nicely fits into the female center conductor of an SO-239.

21.3 5/8-Wavelength 2-Meter Antenna

The 5/8-wavelength antenna (Fig. 21.4) is popular on 2 m for mobile operation because it is easy to construct, and at low elevation angles it provides a small amount of gain relative to the standard $\lambda/4$ ground plane. The radiator element is 5/8-wavelength, so its length is

$$L(\text{in}) = \frac{7380}{f(\text{MHz})} \quad (21.2)$$



For $\frac{5\lambda}{8}$; $L_{FT} = \frac{615}{F_{MHz}}$

$$L_{IN} = \frac{7380}{F_{MHz}}$$

FIGURE 21.4 5/8-wavelength 2-m antenna.

The 5/8-wavelength antenna is not a good match to any of the common forms of coaxial cables. Either a transmission line matching section or an inductor match is normally used. Figure 21.4 shows an inductor match, using a matching coil of two to three turns of #12 wire wound over a 1/2-in-long 1/2-in outside diameter (OD) form. The radiator element can be tubing, brazing rod, or a length of heavy piano wire. For low-power systems, it can be a telescoping antenna that is sold as a replacement for portable radios or television sets. These antennas have the advantage of being adjustable to resonance without the need for cutting.

21.4 J-Pole Antennas

The J-pole antenna is another popular form of vertical on the VHF bands. It can be used at almost any frequency, although the example shown in Fig. 21.5 is for 2 m. The J-pole is an end-fed radiator with quarter-wave matching stub and is a monoband antenna. On harmonic bands, the J-pole can be decently matched, but the radiation pattern may not be ideal.

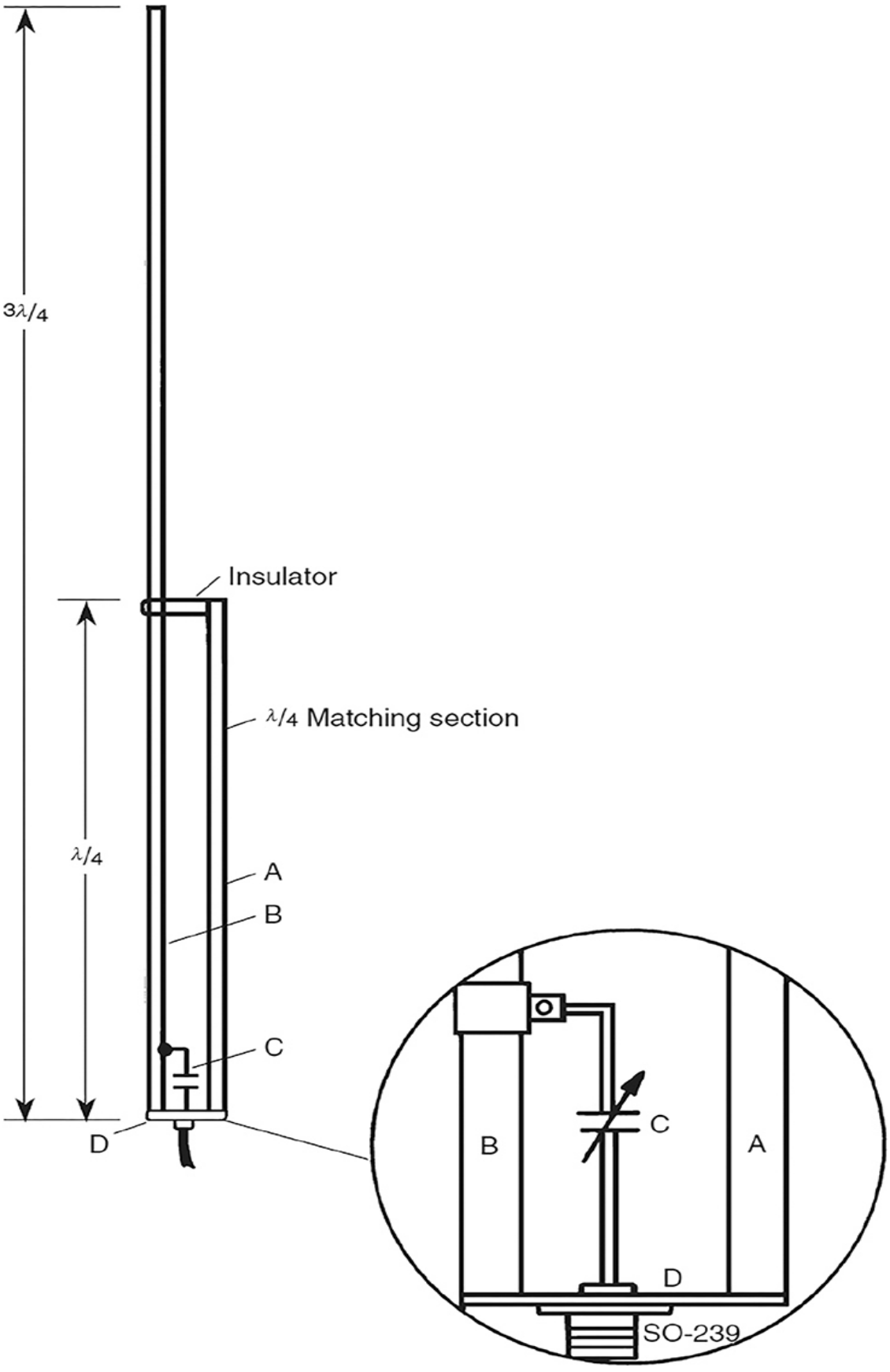

3λ/4
Insulator
λ/4 Matching section
λ/4
A
B
C
D
B
C
A
D
SO-239

FIGURE 21.5 J-pole antenna.

The antenna radiator is 3/4-wavelength long, so its length is

$$L(\text{in}) = \frac{8838}{f(\text{MHz})} \quad (21.3)$$

and the quarter-wavelength matching section length from:

$$L(\text{in}) = \frac{2952}{f(\text{MHz})} \quad (21.4)$$

Taken together, the matching section and the radiator form a parallel transmission line with a characteristic impedance that is four times the coaxial cable impedance. If 50-Ω coax is used and the elements are made from 0.5-in OD pipe a spacing of 1.5 in will yield an impedance of about 200 Ω. Impedance matching is accomplished by a gamma match consisting of a 25-pF variable capacitor connected by a clamp to the radiator and tapped about 6 in above the base.

21.5 Collinear Vertical

Gain in antennas is realized by taking radiated power from some directions and adding it to other directions. While many of the vertically polarized HF directive antennas we examined in previous chapters have taken energy from one or more azimuthal headings and used it to augment the power radiated in the "forward" direction, it is also possible to obtain gain over a limited range of elevation angles in all compass directions by utilizing energy that might have otherwise gone into elevation angles of little use to us. In this manner, it is possible to have a vertical antenna that exhibits gain for certain elevation angles around the entire 360-degree azimuth range.

Figure 21.6 is an example of such an antenna. Through the use of collinear elements for the two sides of a vertical dipole, it develops gain at low elevation angles in an omnidirectional pattern. Each array consists of a quarter-wavelength section A and a half-wavelength section C separated by a phase reversing stub B. The total length of the stub, out on one side and back on the other, is a half-wavelength, but if made from open-wire line, only a $\lambda/4$ line is required. The phase reversal stub adds a 180-degree phase shift between the inner $\lambda/4$ section and the outer

$\lambda/2$

section. Because either out-board half-wavelength section is $\lambda/2$ or 180 degrees above or below the center half-wave section, the in-phase currents in the sections cancel in the vertical direction. In any horizontal direction, the current in each outer section is in phase with that of its respective inner section, and the antenna exhibits about 3 dB gain at low elevation angles over a conventional

$\lambda/2$

vertical dipole or $\lambda/4$ ground-plane vertical. The phase reversal stub, even if made out of open-wire line, is not being operated as a transmission line and the usual length correction for velocity factor is not applied.

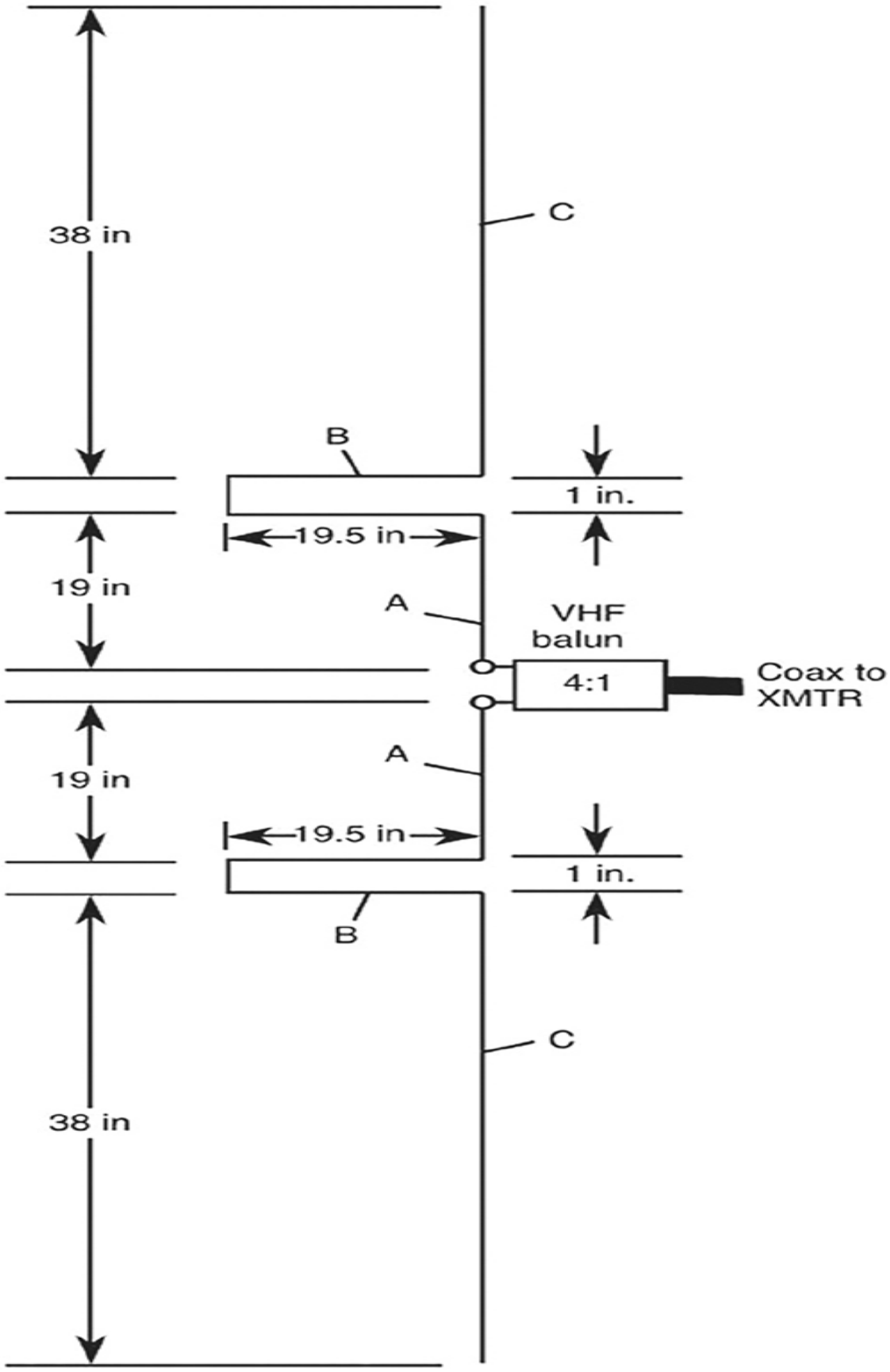
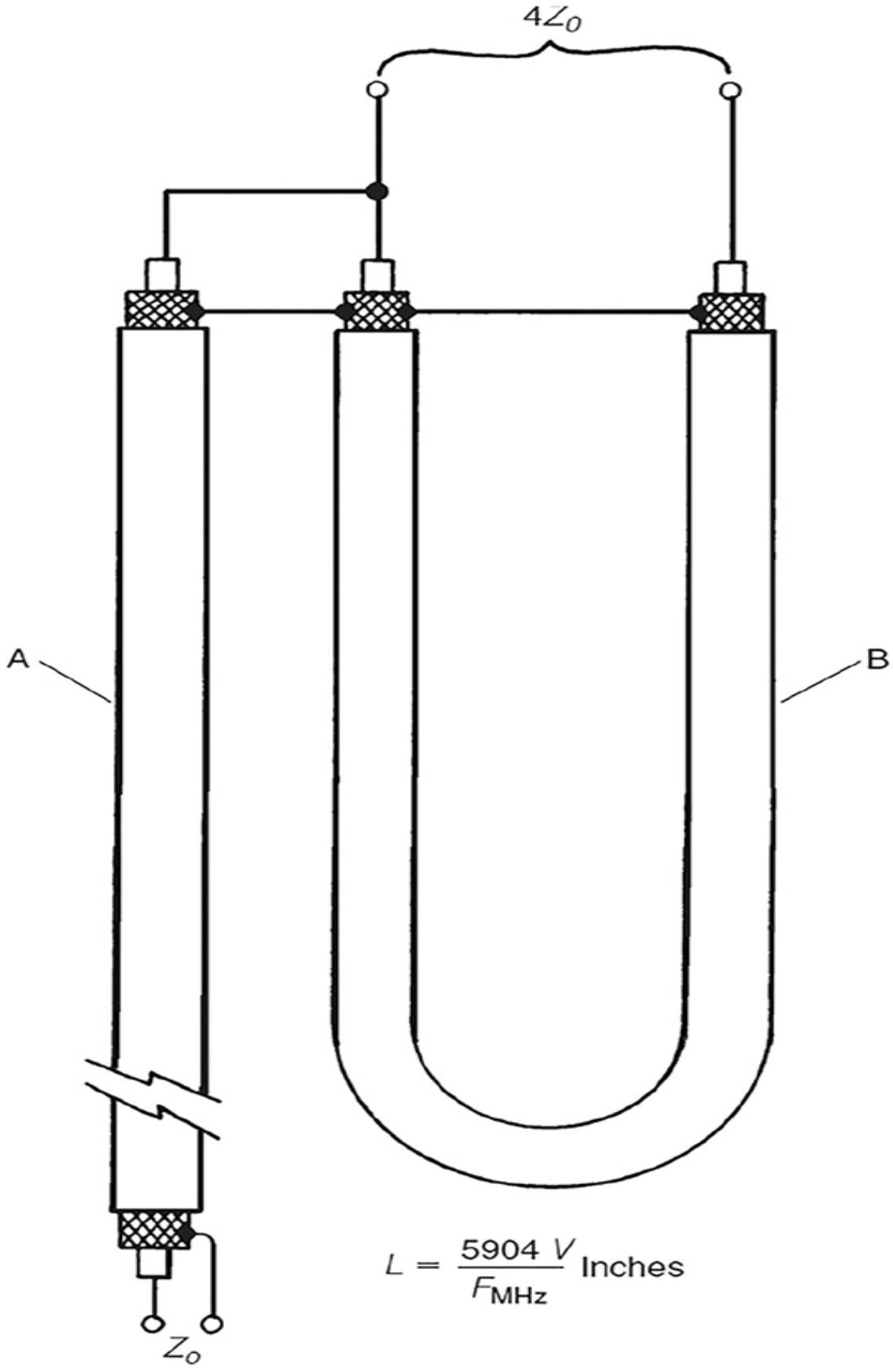

38 in
C
B
1 in.
19.5 in
19 in
A
VHF
balun
4:1
Coax to
XMTR
A
19 in
19.5 in
1 in.
B
C
38 in

FIGURE 21.6 Vertical collinear antenna.

As with an ordinary dipole, the feedpoint is at the midpoint of the array. The free-space feedpoint impedance is around 280 Ω resistive at 146 MHz—a reasonably good match to a 4:1 balun transformer (see Fig. 21.7), especially if fed with 72-Ω coaxial cable. Alternatively, 300-Ω twin-lead can be connected directly to the two sides of the dipole center. If the transmitter lacks the balanced output needed to feed twin-lead, use an appropriate balun at the input end of the twin-lead near the transmitter. For proper performance, keep the feedline at right angles to the dipole axis near the dipole.



$$L = \frac{5904\ V}{F_{MHz}} \text{ Inches}$$

FIGURE 21.7 Coaxial 4:1 balun transformer.

21.6 Grounds for Ground-Plane Verticals

Most of the preceding verticals are monopoles. All will be poor performers if a good ground and return path creating the lower half of a dipole is not provided. When the base of these antennas is elevated more than a small fraction of a wavelength above ground, an artificial ground plane is a necessity. The simplest way to provide this is with three or four radials cut to $\lambda/4$ at the design frequency of the vertical and spaced equally around the compass (120 or 90 degrees apart, respectively). The radials should slope down from the base at about a 45-degree angle. Some experimentation with their exact length and slope may be needed. In contrast to radials laid directly on the earth's surface, elevated radials are resonant elements, and their length will have a strong effect on both the SWR and the radiation efficiency of the vertical.

21.7 Yagi Antennas

The Yagi beam antenna is a highly directional gain antenna and is used in both HF and VHF/UHF systems. The antenna can be easier to build at VHF/UHF than at HF. The basic Yagi was covered in Chap. 14, so we will only show examples of practical VHF devices.

21.7.1 Six-Meter Yagi

A four-element Yagi for 6 m is shown in Fig. 21.8. The reflector and directors can be mounted directly to a metallic boom, but the choice of driven element and its feed system require that the driven element be insulated from the boom.

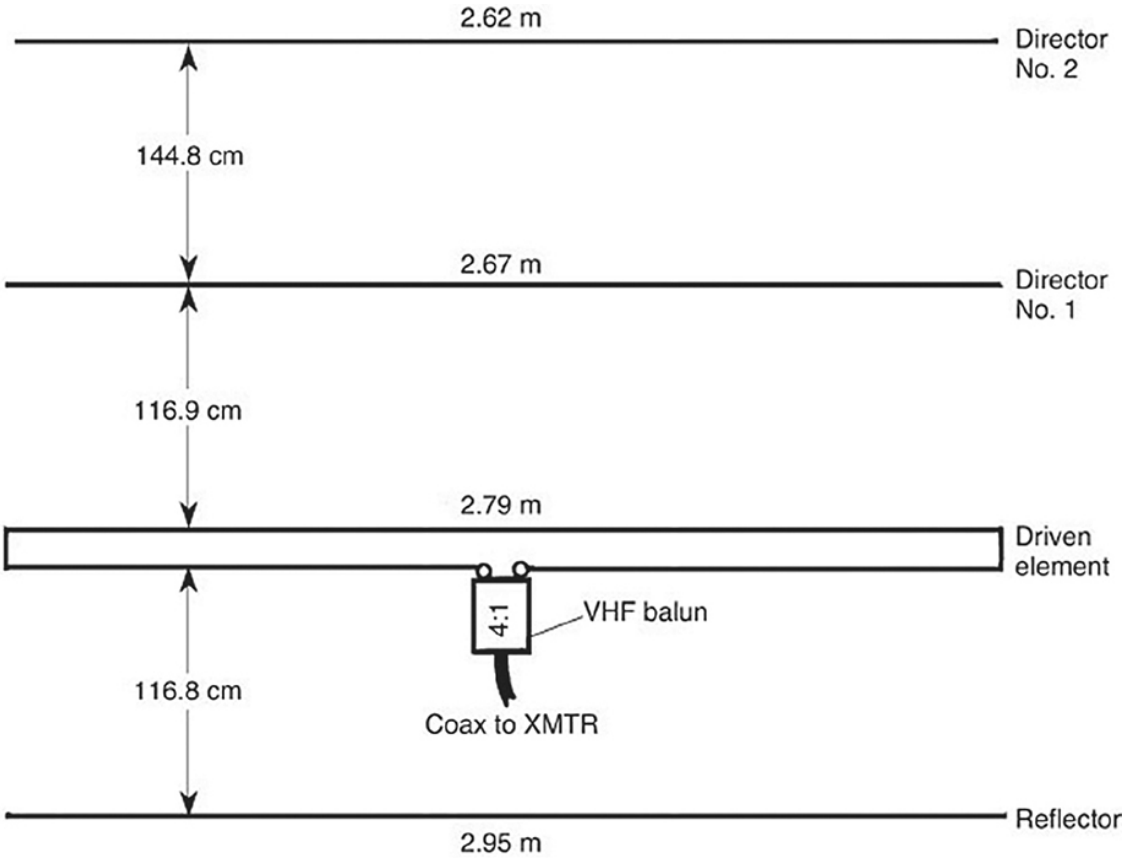


FIGURE 21.8 Six-meter beam antenna.

The driven element shown in Fig. 21.8 is a folded dipole; because both conductors are the same diameter, it provides a 4:1 step-up in feedpoint impedance relative to a conventional dipole that is largely independent of the ratio of conductor spacing to conductor diameter. The use of a folded dipole for the driven element is common practice at VHF because it tends to broaden the SWR bandwidth of the antenna's input impedance and because the native feedpoint impedance of Yagis is often very low. For this construction project, two 3⁄4-inch aluminum tubes, spaced 6.5 in apart, were used.

The antenna can be directly fed with balanced 300-Ω line or from 50-Ω or 75-Ω coax through a 4:1 balun. If broadbanding is not important, the folded dipole of the driven element can be replaced with a conventional dipole and the antenna fed directly from a 52- or 75-Ω feedline.

21.7.2 Two-Meter Yagi

Figure 21.9 shows the construction details for a six-element 2-m Yagi beam antenna. This antenna is built using a 2×2-in wooden boom and elements made of brass or copper rod. Threaded brass rod simplifies assembly, but is not strictly necessary. The job of securing the elements (other than the driven element) is easier with threaded rod because it allows securing of each parasitic element with a pair of hex nuts, one on either side of the boom. Unthreaded elements can be secured with RTV to seal a press-fit. Alternatively, secure the rods with tie wires (see inset to Fig. 21.9): drill a hole through the boom to admit the rod or tubing, secure the element by wrapping a tie wire around the rod on either side of the boom, and solder it in place. Use #14 to #10 solid bare or tinned wire for the tie wires.

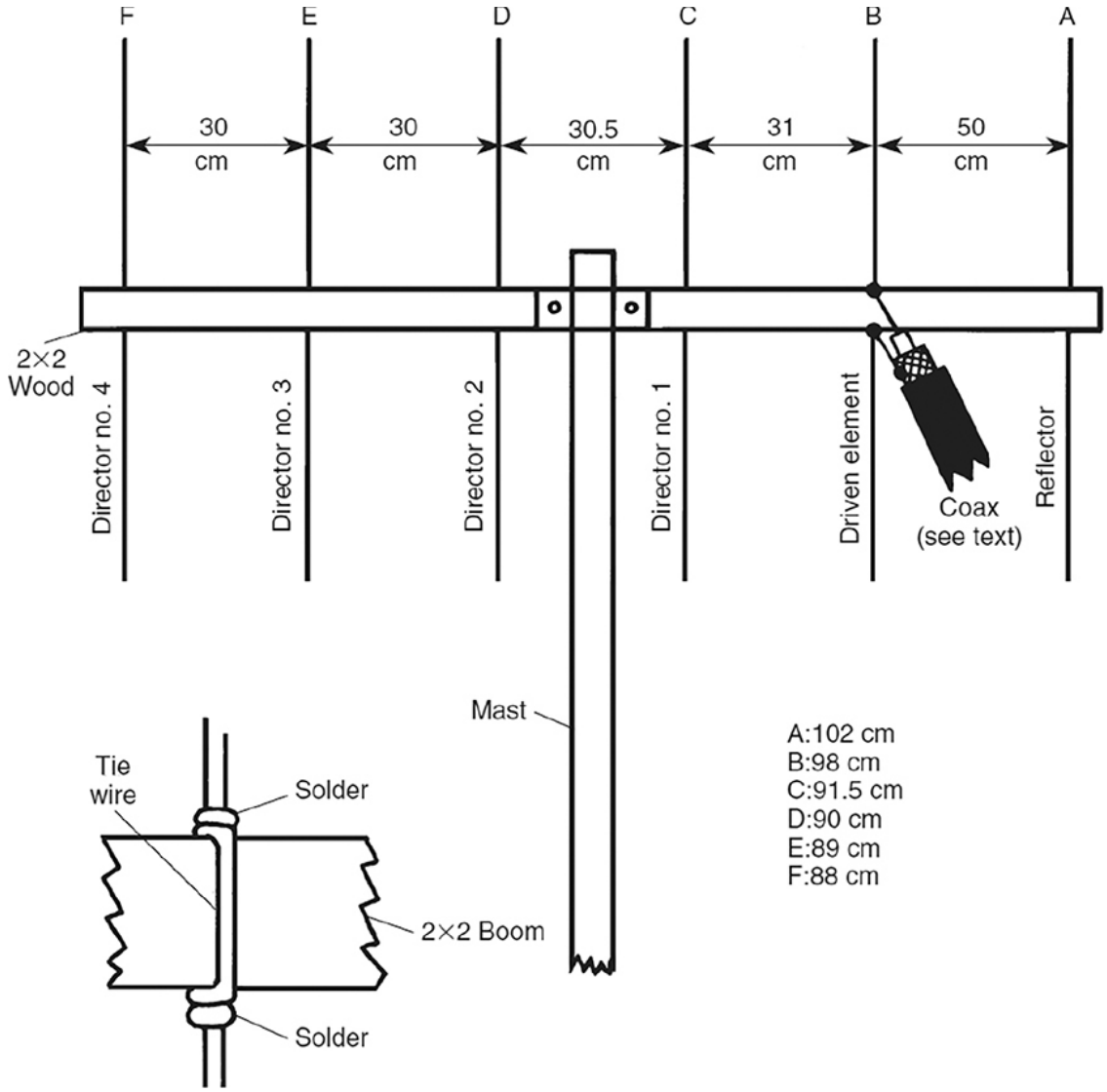


FIGURE 21.9 Two-meter vertical beam.

Mount the mast of the antenna to the boom with a bracket or clamp. One alternative is an end-flange clamp, such as is sometimes used to support pole lamps and the like. The mast should be attached to the boom at the center of gravity. If you try to balance the antenna on one hand unsupported, there is one point at which it is balanced. Attach the mast hardware at or near this point to prevent gravitational torques from tearing the mounting apart.

The antenna is fed with coaxial cable at the center of the driven element. Often, either a matching section of coax or a gamma

match will be needed because the usual effect of parasitic elements on a Yagi's feedpoint impedance is to make it too low to be a good direct match to standard coaxial cable impedances.

21.8 Halo Antennas

Another VHF antenna is the halo (Fig. 21.10), formed by bending a half-wavelength dipole into a circle. The ends of the dipole are connected to the two plates of a capacitor. A transmitting-type mica "button" capacitor or two 3-in disks separated by air or plastic dielectric can be used. In some implementations of the halo, the spacing of the disks is adjustable via a threaded screw. While air is a good dielectric, the use of a plastic spacer provides mechanical rigidity to help stabilize the value of the capacitor.

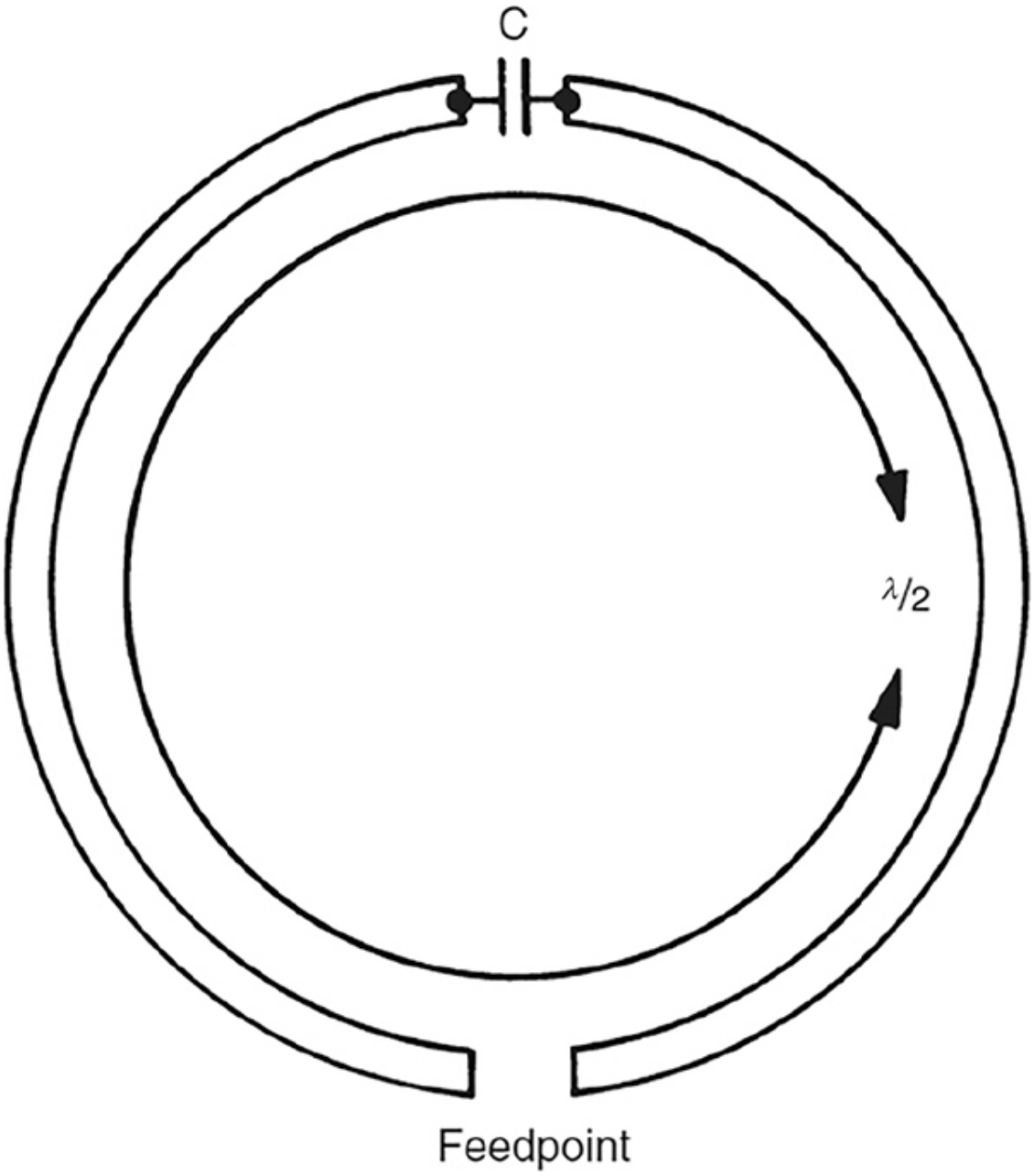


FIGURE 21.10 Halo antenna.

21.9 Quad Beam Antennas

The quad antenna was introduced in Chap. 15. Although the discussion there revolved around implementations for HF, it has emerged as a good VHF/UHF antenna. This antenna is a lot easier to construct at VHF/UHF frequencies than it is at HF frequencies.

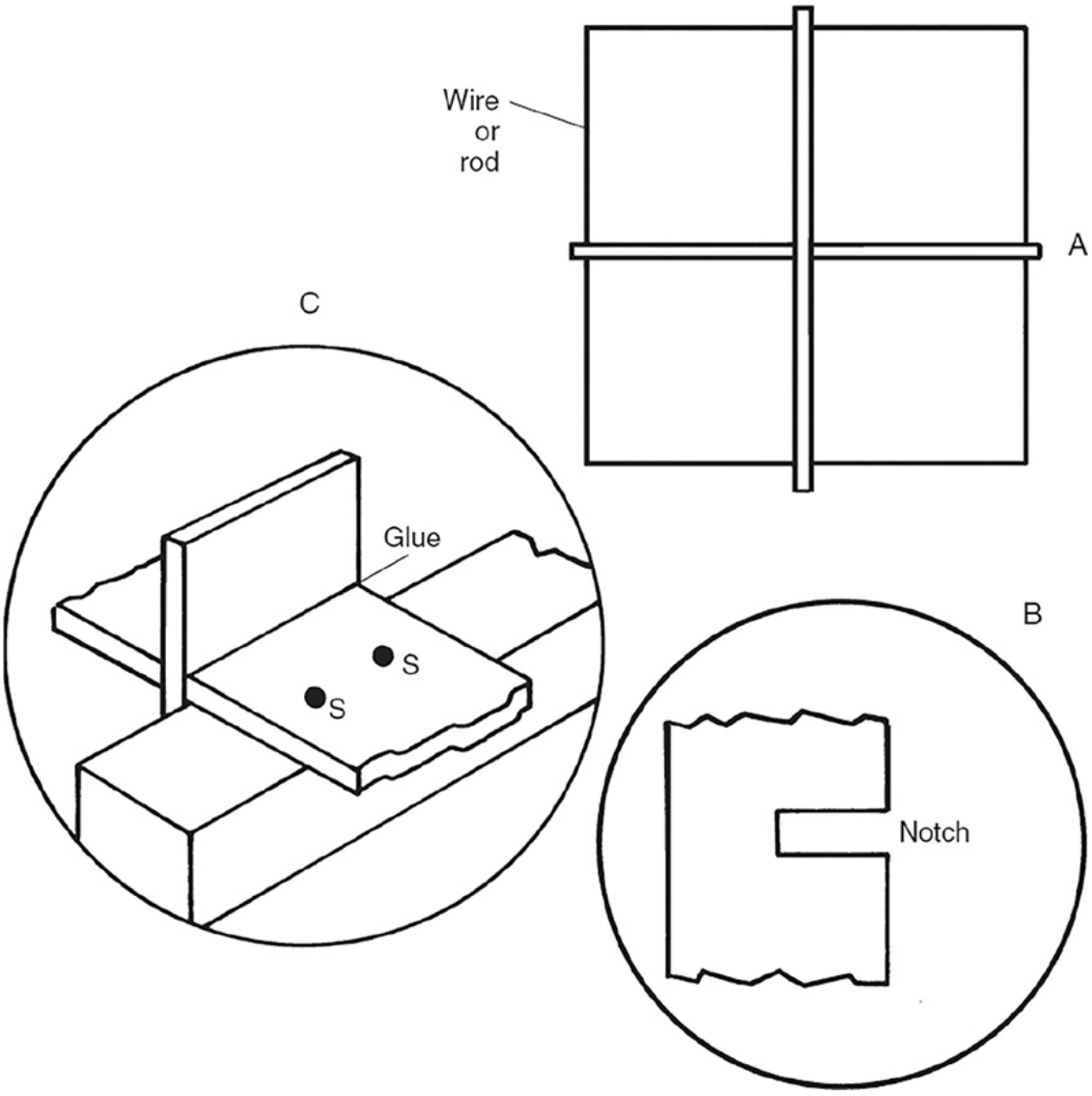
Figure 21.11 depicts one of several methods for building a quad for VHF or UHF. The radiator element can be any of several materials, including heavy solid wire (#8 to #12), tubing, or metal rods. The overall lengths of the elements are given by:

Overall lengths:

$$\text{Reflector: } L = \frac{12{,}562}{F_{\text{MHz}}} \text{ in}$$

$$\text{Director: } L = \frac{11{,}248}{F_{\text{MHz}}} \text{ in}$$

$$\text{Driven element: } L = \frac{11{,}826}{F_{\text{MHz}}} \text{ in}$$



F IGURE 21.11 Quad loop construction on VHF.

Driven element:

$$L(\text{in}) = \frac{11,826}{f(\text{MHz})} \quad (21.5)$$

Reflector:

$$L(\text{in}) = \frac{12,562}{f(\text{MHz})} \quad (21.6)$$

Director:

$$L(\text{in}) = \frac{11,248}{f(\text{MHz})} \quad (21.7)$$

Thanks to this antenna's lightweight construction, there are several alternatives for making the supports for the elements. In Fig. 21.11, detail A, the spreaders are made from furring, trim strips, or even wooden paint-stirring sticks. The sticks are cut to length and half-notched in the center (Fig. 21.11, detail B). The spreaders for each element are joined together at right angles and glued (Fig. 21.11, detail C). The spreaders can be fastened to the wooden boom at points S in detail C. Follow the usual rules regarding element spacing (0.15 to 0.31 wavelength). See the information on quad antennas in Chap. 15 for further details.

Quads have been successfully employed on all amateur bands up to 1296 MHz. One variant of the VHF quad is the Quagi, a Yagi that employs a quad-driven element.

21.10 VHF/UHF Scanner Band Antennas

Many radio enthusiasts monitor emergency response frequencies out of a desire to learn what's going on in the community. Antennas for this purpose cover some or all of the VHF and UHF bands. Although scanner antenna designs might be optimized for VHF or UHF, they are basically the same as others shown in this book, most of which can be used over at least part of the scanner spectrum.

21.10.1 "Scanner-Vision" Antennas

The antennas used by scanner listeners are widely varied. One possibility is an antenna designed for the phased out analog broadcast television bands, which ranged from 54 MHz to around 800 MHz. This encompasses most of the commonly used scanner frequencies. Even if antenna performance is not optimized for the scanner frequencies, performance should be reasonable.

For the scanner user, the directional characteristic of the TV antenna may or may not be useful. If the antenna has a rotator, rotate the antenna in the direction of interest. Since no antenna completely rejects signals arriving from off the sides or the back and local scanner repeaters are quite strong, the rotator may not be needed. An alternative is an indoor set-top TV antenna. These can be used with a wideband TV antenna preamplifier if needed.

21.10.2 Scanner Skyhooks

Some scanners cover much of the HF spectrum. A random-length wire antenna can provide decent performance, and it may even be able to do double duty as a VHF/UHF longwire antenna. Typically, this antenna is a 30- to 150-ft length of #14 wire attached to a distant support.

Gain of around 3 dB can be achieved by stacking VHF/UHF antennas together. Figure 21.12 shows a typical arrangement in which two half-wavelength dipole antennas are connected together through a quarter-wavelength harness of RG-59/U coaxial cable. This harness is physically shorter than an electrical quarter-wavelength by the velocity factor of the coaxial cable:

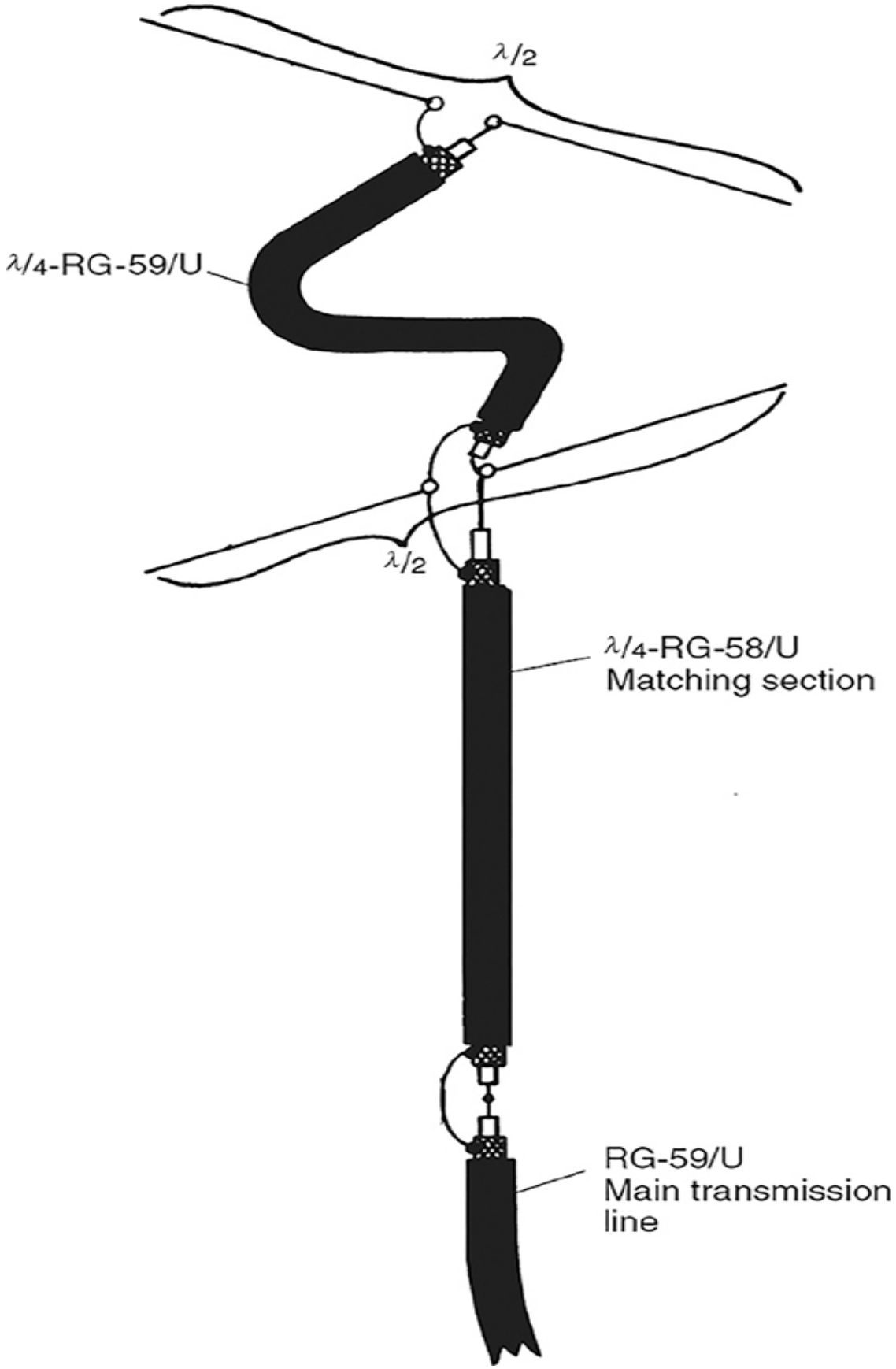
$$L(\text{in}) = \frac{2832\nu_F}{f(\text{MHz})} \quad (21.8)$$

where

v_F

is the velocity factor (typically 0.66 or 0.80 for common coax) and *f* is the frequency in megahertz. The antennas can be oriented in the

same direction to increase gain, or orthogonally (as shown in Fig. 21.12) to obtain a more omnidirectional clover-leaf pattern.


λ/2
λ/4-RG-59/U
λ/2
λ/4-RG-58/U
Matching section
RG-59/U
Main transmission
line

F IGURE 21.12 Stacking VHF antennas.

Because the impedance of two identical dipoles, fed in parallel, is one-half that of a single dipole, it is necessary to have an impedance-matching section made of RG-58/U coaxial cable. This cable is then fed with RG-59/U coax from the receiver.

21.11 VHF/UHF Antenna Impedance Matching

Like HF antennas, VHF/UHF antennas usually require impedance matching to the feed-line. Some methods, like coax baluns and the delta match, are easier at the higher frequencies, while others become difficult or impossible. An example of the latter case is the tuned *LC* impedance-matching network. At 6 m, and even to some extent at 2 m, lumped-component *LC* networks can be used. But at 2 m and above, other methods are often easier to implement, such as stripline tank circuits.

The balun transformer provides an impedance transformation between *bal*anced and *un*balanced circuit configurations. Although both 1:1 and 4:1 impedance ratios are possible, the 4:1 ratio is most commonly used for VHF/UHF antenna work. At lower frequencies it is easy to build broadband transformer baluns, but these become more of a problem at VHF and above.

Figure 21.7 is a 4:1 impedance ratio coaxial balun often used on VHF/UHF frequencies. Two sections of identical coaxial cable are needed. One section (A) can be any convenient length needed to reach from the antenna to the transmitter. Its characteristic impedance is

Z_0

. The other section (B) is cut to be a half-wavelength long at the center of the frequency range of interest. The physical length is found from

$$L(\text{in}) = \frac{5904\nu_F}{f(\text{MHz})} \quad (21.9)$$

where *f* is the operating frequency in megahertz and v_F is the coaxial cable velocity factor. Hard line velocity factors typically lie in the 0.8 to 0.9 range. Alternatively, you can measure the velocity factor of the specific piece of cable you intend to use. See Chap. 30 for more information. The velocity factors of common coaxial cables are shown in Table 21.1.

TABLE 21.1 Coaxial Cable Velocity Factors

Regular polyethylene	0.66
Polyethylene foam	0.80
Teflon	0.72

Example 21.1 Calculate the physical length required of a 146-MHz 4:1 balun made of polyethylene foam coaxial cable.

Solution

$$L = \frac{5904\nu_F}{f} = \frac{(5904)(0.08)}{146} = 32.4 \text{ in}$$

One approach to joining the coaxial cables is shown in Fig. 21.13. Three SO-239 coaxial receptacles are mounted on a metal plate. This arrangement has the effect of shorting together the shields of the three ends of coaxial cable while the center conductors are connected in the manner shown. This method is sometimes provided as part of a commercial antenna. Because the balun of Figs. 21.7 and 21.13 is mounted at the antenna feedpoint and is almost always located outdoors, care must be taken to weatherproof the ends of the transmission line, especially the long run back to the transmitter.

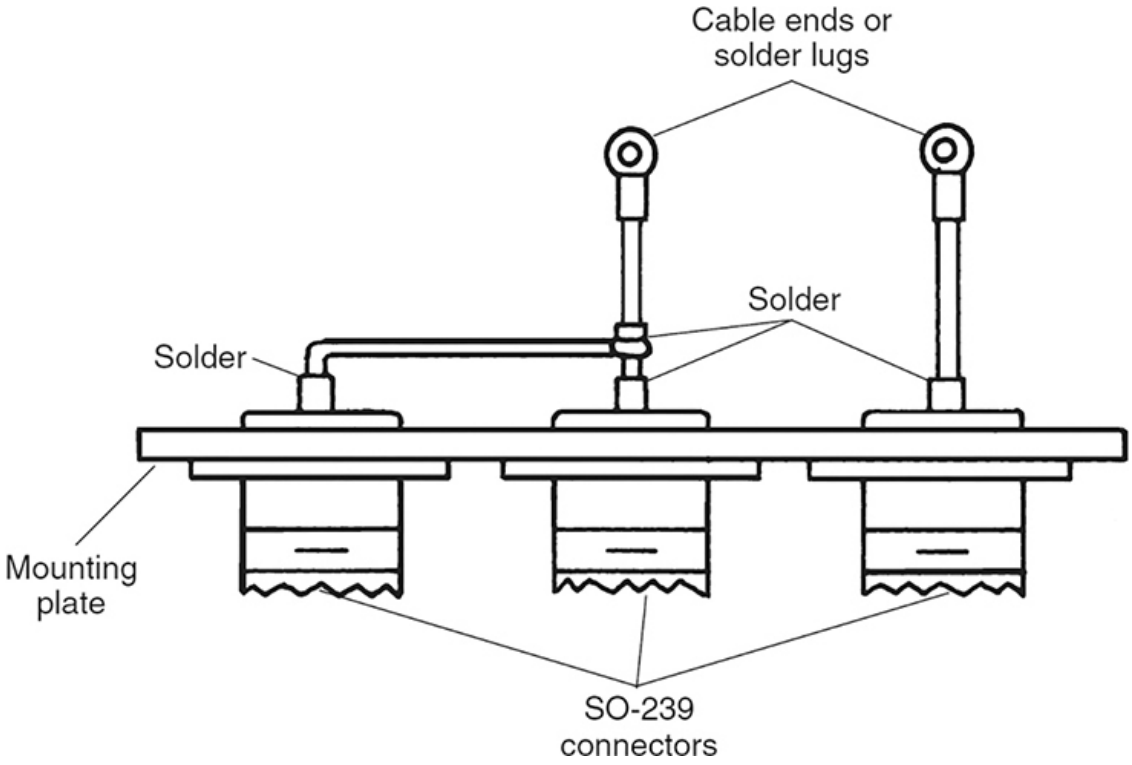


FIGURE 21.13 Practical implementation of 4:1 balun using connectors.

The delta match gets its name from the fact that the physical layout of the matching network used when tapping out on the two sides of a driven element or dipole a distance corresponding to the transmission line impedance has the shape of the Greek letter delta. Figure 21.14 shows the basic delta match scheme. The delta match connections to the driven element are made symmetrically on both sides of the center, and can be made from brass, copper, aluminum tubing, or a bronze brazing rod bolted to the main radiator element.

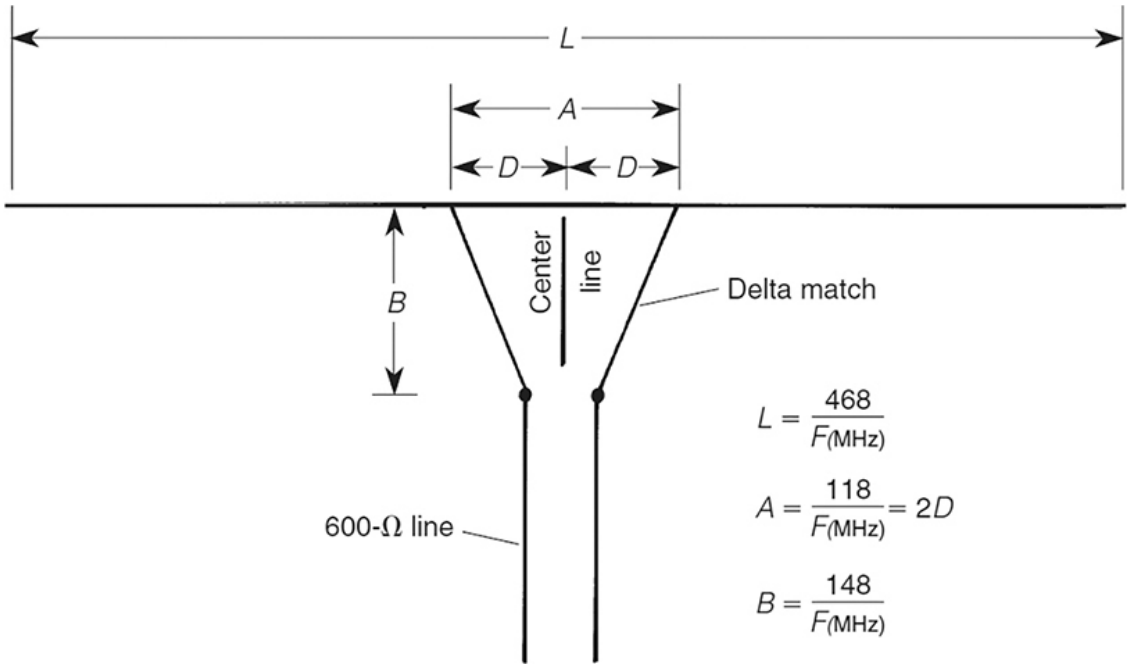


FIGURE 21.14 Delta feed matching system.

The width (*A*) of the delta match is given by

$$A(\text{in}) = \frac{1416}{f(\text{MHz})} \quad (21.10)$$

while the length (*B*) of the matching section is

$$B(\text{in}) = \frac{1776}{f(\text{MHz})} \quad (21.11)$$

The transmission line feeding the delta match is a balanced line, such as a parallel transmission line or twin-lead. The exact impedance is not critical because the dimensions (especially *A*) can be adjusted for a good match. Although the 450-Ω line is commonly used, 300- or 600-Ω lines are good alternatives. Figure 21.15 shows a method for using coaxial cable or a hard line with the delta match. The 50- or 75-Ω coax or hardline impedance is transformed up to the delta match impedance with a 4:1 balun transformer such as the one in Fig. 21.7.

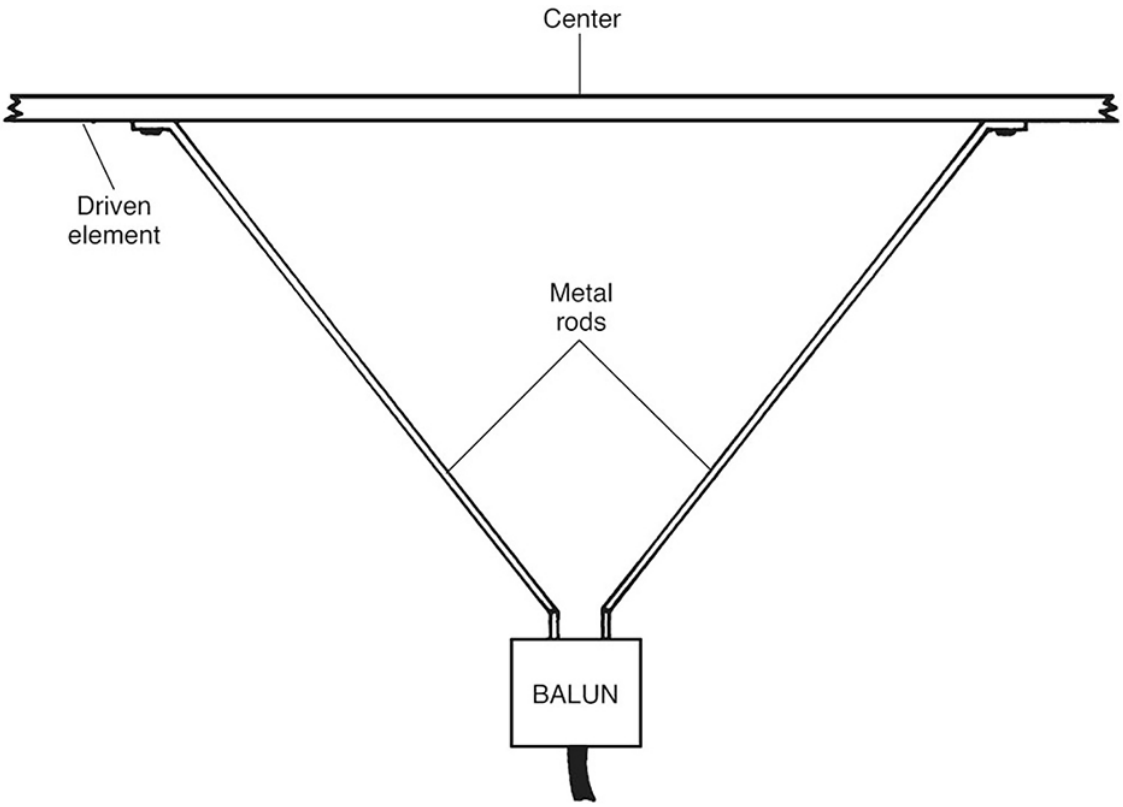
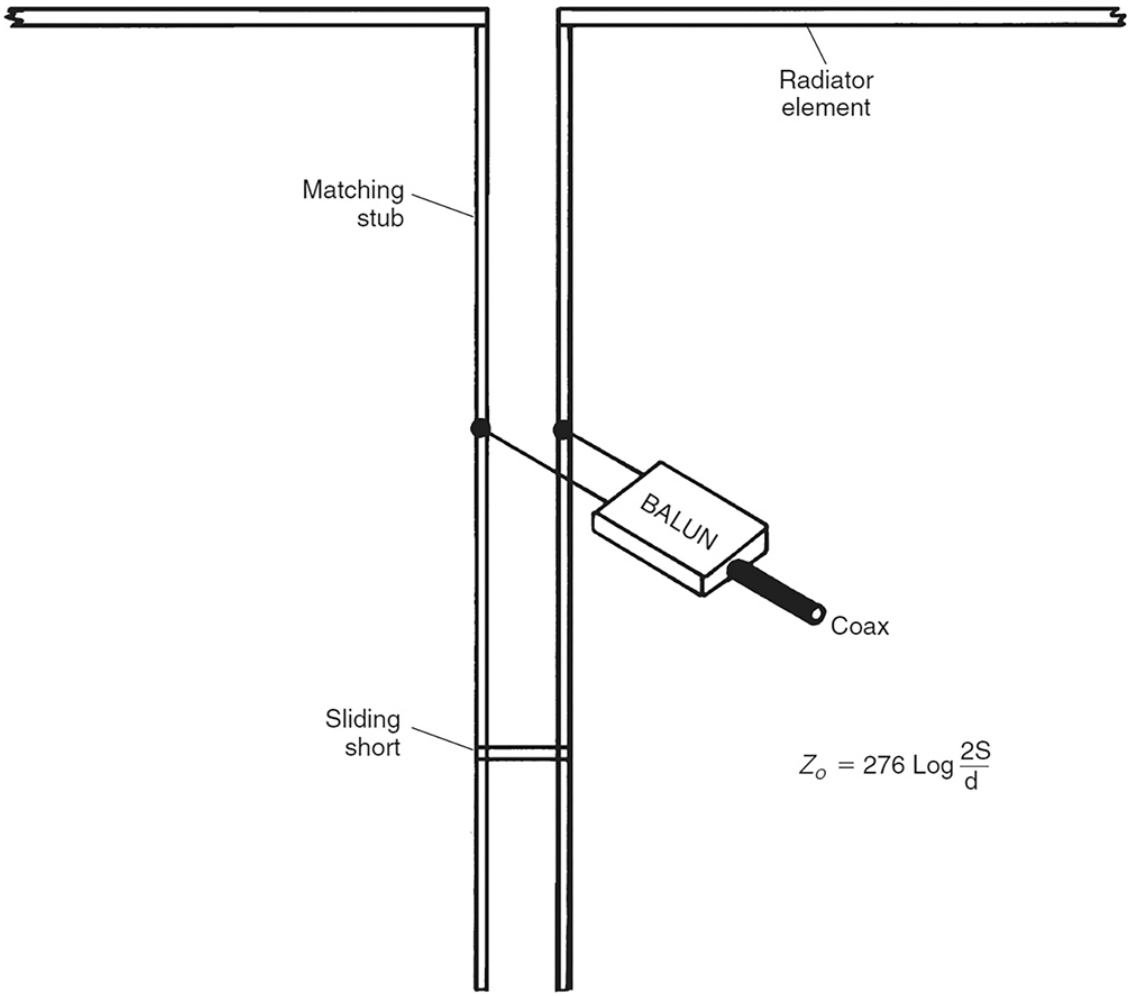


FIGURE 21.15 Practical VHF delta match.

A stub matching system is shown in Fig. 21.16. The exact impedance of the line is not critical and is found from

$$Z_0 = 276 \log_{10}\left(\frac{2S}{d}\right) \quad (21.12)$$

where *S* is the space between the two sides of the line and *d* is the diameter of each conductor.



F IGURE 21.16 Stub matching.

The matching stub section is made from metal elements such as tubing, wire, or rods, since all three are practical at VHF/UHF frequencies. For a $3/16$-in rod, the spacing is approximately 2.56 in to make a 450-Ω transmission line. A sliding short circuit sets the electrical length of the half-wave stub. The stub is tapped at a distance from the antenna feedpoint that matches the impedance of the transmission line. In the example shown, the transmission line is coaxial cable, so a 4:1 balun transformer is used between the stub and the transmission line. The two adjustments to make in this system are: (1) the distance of the short from the feedpoint and (2) the distance of the transmission line tap point from the feedpoint. The two are alternately adjusted for minimum VSWR.

The gamma match is basically a half-delta match and operates according to similar principles (Fig. 21.17). The shield (outer conductor) of the coaxial cable is connected to the exact center of the driven element. The center conductor of the coaxial cable series feeds the gamma element through a variable capacitor.

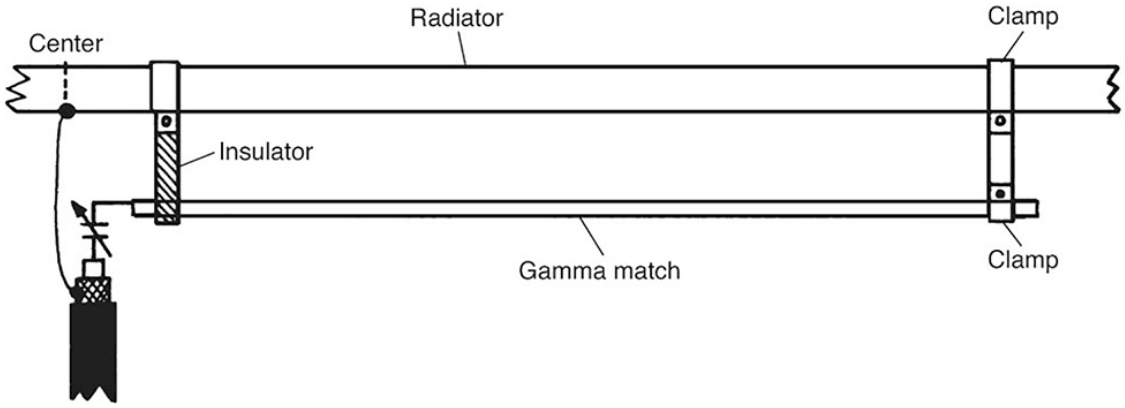


FIGURE 21.17 Gamma matching.

PART V

Microwave Antennas

CHAPTER 22
Microwave Waveguides
CHAPTER 23
Microwave Antennas
CHAPTER 24
Antenna Noise
CHAPTER 25
Radio Astronomy Antennas
CHAPTER 26
Radio Direction-Finding (RDF) Antennas
CHAPTER 27
Antennas for Wireless Devices

CHAPTER 22
Microwave Waveguides

The microwave portion of the radio spectrum covers frequencies from about 900 MHz to 300 GHz, with wavelengths in free space ranging from 33 cm to 1 mm. At microwave frequencies, waveguides and other structures that are essentially conductive pipes for RF energy are widely used. These structures guide RF energy to and from antennas and are used as building blocks for some types of microwave antennas, including the horn antenna that we will cover in Sec. 23.2.

22.1 Waveguide Basics

A waveguide is a generalization of a transmission line. Coaxial cable, parallel wire line, microstrip, and other types of lines were analyzed using transmission line formulas in Chap. 6, but they can also be analyzed as waveguides. All long structures, whether coaxial with two conductors, a hollow waveguide, or an optical fiber, support discrete propagating solutions with different phase velocities and field profiles called modes.

The main difference between these types of lines and the waveguides on which we will focus on in this chapter is that the lines in Chap. 6 have two conductors and support low frequency transmission all the way down to DC, whereas the hollow waveguides we are looking at here only allow energy to propagate above a certain frequency that depends on the dimensions of the structure. We will discuss this more below.

Compared to coaxial cable, which is dielectric filled, and parallel wire line, which is unshielded and loses power through radiation, hollow waveguides can guide RF energy with lower loss. In general, transmission lines exhibit three kinds of losses: ohmic, dielectric, and radiation. Ohmic losses are the result of the skin effect, which

increases a conductor's apparent resistance with increasing frequency. Dielectric losses are caused by the electric field acting on the molecules of the insulator, thereby causing heating through molecular agitation, and tend to be largest when the wavelength is comparable to the interatomic resonances of the dielectric. Radiation losses result from incomplete shielding or from imbalance and incomplete cancellation of RF on the two sides of the line, usually because at some frequency the spacing between the wires becomes too large a percentage of the RF wavelength. These losses are proportional transmission line length and generally grow with frequency. Losses reduce signal strength and limit the maximum power and usable frequency for a given line length.

To circumvent these problems, hollow waveguides can be used for high frequency or high power applications. These types of waveguides are essentially metal pipes and can have circular or rectangular cross sections, though occasionally waveguides with other cross section shapes are used. The internal walls of a waveguide are electrical conductors, typically aluminum, brass, or copper. To further reduce ohmic losses, the internal surfaces of some waveguides are electroplated with gold or silver, both of which have lower resistivities than the other metals mentioned. In a waveguide, RF energy reflects from the inner walls of the waveguide and is guided down the hollow tube. Due to the zigzag propagation path, the forward velocity along the waveguide's axis is less than the speed of light in the air or dielectric inside the waveguide.

22.2 Waveguide Modes

In empty space, there is a continuum of possible directions for a radio wave to travel, but within a waveguide, boundary conditions restrict the possible propagating solutions to a discrete set of modes. Each mode has a different field pattern, propagation velocity, characteristic impedance, and cutoff frequency below which it cannot propagate down the waveguide.

The mode with lowest cutoff frequency is the dominant mode. Waveguides are usually used with only the dominant mode excited.

But a waveguide in general can be viewed as an infinite number of transmission lines operating in parallel and each higher order mode can be analyzed as if it were a separate transmission line. Higher order modes become important in special cases where the waveguide is deliberately operated with multiple modes or a waveguide bend or junction converts power from one desired mode to other unwanted modes.

22.3 The TEM Mode

Waveguides can be divided into two classes, those with two or more separate conductors, like coaxial cable, and those with a single conductive boundary, such as the rectangular or circular waveguide.

The multiconductor lines considered in Chap. 6 support a mode with transverse electromagnetic (TEM) fields. The word transverse indicates that the wave is propagating at right angles to both the electric and magnetic fields. The E- and H-fields are "normal" or "orthogonal" to the direction of travel down the waveguide. The TEM mode has a cutoff frequency of zero, which follows from the fact that a two-conductor waveguide can carry DC current due to the presence of an electrical return path. When the line is small in diameter or cross-sectional size, the cutoff frequencies of the higher order modes are much higher than the operating frequency of signals on the line. We can usually ignore the higher order modes that can potentially exist in the structure, and we analyze propagation along the structure using transmission line formulas for the TEM mode.

On the other hand, hollow, single-conductor waveguides, such as rectangular or circular metallic tubes, do not allow a TEM mode or carry DC current and only modes with a nonzero cutoff frequency can propagate in the waveguide. These waveguides can only be used above a certain frequency that depends on the dimensions of the waveguide.

In view of this practical distinction, we often think of coaxial cable, parallel wire transmission lines, and other types of lines as transmission lines, and hollow-pipe type lines as waveguides, but it is more rigorous to think of waveguides as a broader concept that

encompasses all transmission lines of any type that are long in one direction and have a given cross-sectional shape. Coaxial cable can support higher order modes, although this is almost never used in practice. A sharp bend can cause mode conversion and occasionally produce difficulties in a system.

22.4 Rectangular Waveguide

Of the hollow waveguides, a rectangular shape is most common, followed by circular. Figure 22.1 is a cross-sectional view of a typical rectangular waveguide. We orient the waveguide so that the widest interior dimension *a* is horizontal and *b* is the narrower dimension.

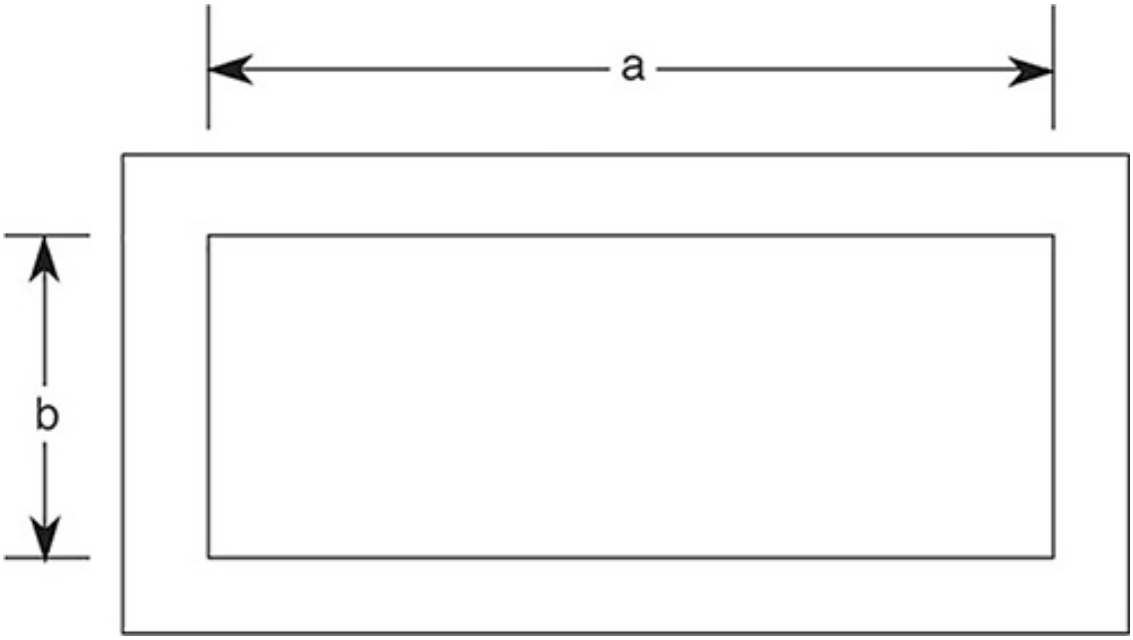


FIGURE 22.1 Rectangular waveguide (end view). The longer dimension is placed in the *x* direction of a coordinate system, with the shorter dimension in the *y* direction, and energy propagates along the waveguide in the $\pm z$ direction.

22.5 Boundary Conditions

A hollow, single-conductor waveguide will not support a propagating TEM mode because such a mode does not satisfy constraints existing

at the metal walls, called boundary conditions. Although the wave in the waveguide propagates through the air (or inert gas dielectric) in a manner similar to free-space propagation, the phenomenon is bounded by the walls of the waveguide. This restricts the fields inside the waveguide to certain modes that satisfy the boundary condition. The boundary conditions are:

- The electric field must be orthogonal to a conducting surface.
- The magnetic field must be orthogonal to a conducting surface.

To satisfy these boundary conditions the waveguide supports two types of propagation modes: transverse electric (TE) and transverse magnetic (TM). The TEM mode for radio waves propagating through free space or on multiconductor transmission lines violates the boundary conditions because the magnetic field is not parallel to the surface.

The transverse electric field requirement means that the E-field must be perpendicular to the conductor wall of the waveguide. To launch a TE mode, a vertically polarized coupling radiator can be introduced through a small hole in the waveguide wall to provide the necessary transverse field.

The second boundary condition implies that the magnetic (H) field may not be orthogonal to the conductor surface. An H-field that is at right angles to the E-field (which is orthogonal to the conductor surface) meets this requirement (see Fig. 22.2). The planes formed by the magnetic field are parallel to both the direction of propagation and the wide surface dimension of the waveguide.

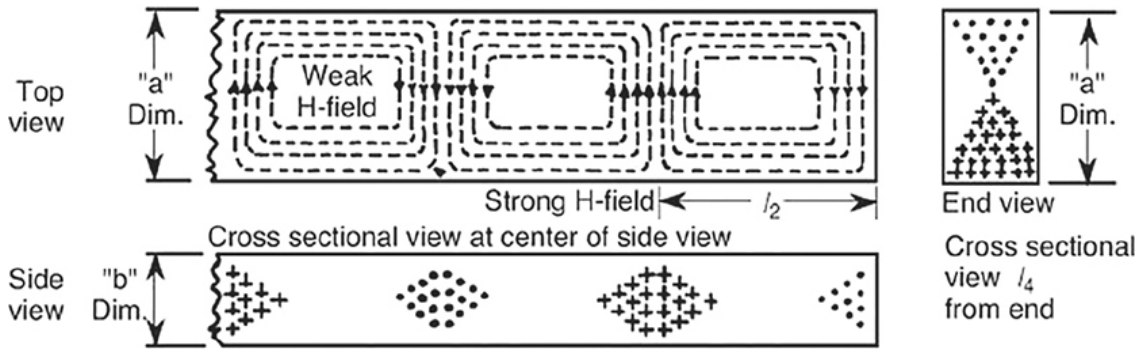


FIGURE 22.2 Magnetic fields in waveguide.

22.6 Rectangular Waveguide Modes

With transmission lines like coaxial cable, we typically model only the TEM mode, but in general in a waveguide, an infinite number of modes can exist. For a rectangular waveguide, we denote the modes as

TE_{mn}

for transverse electric and

TM_{mn}

where *m* is the number of half-wavelengths along the *x* axis (*a* dimension) and *n* is the number of half-wavelengths along the *y* axis (*b* dimension).

For the rectangular waveguide, the

TE_{10}

mode is the dominant mode and is the best mode for low-attenuation propagation along the waveguide. The nomenclature

TE_{10}

indicates that there is one half-wavelength in the *a* dimension and zero half-wavelengths in the *b* dimension. The cutoff frequency of the dominant mode occurs when the longest dimension of the rectangular waveguide cross section is a half wavelength.

Transverse magnetic modes also propagate in waveguides, but the base

TM_{10}

mode is excluded by the boundary conditions. Thus, the

TM_{11}

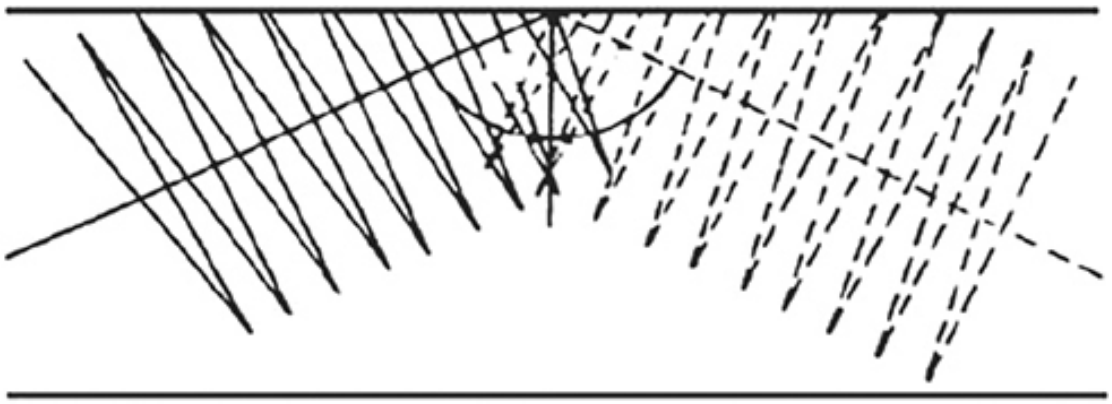
mode is the lowest magnetic mode that will propagate.

22.7 Waveguide Excitation

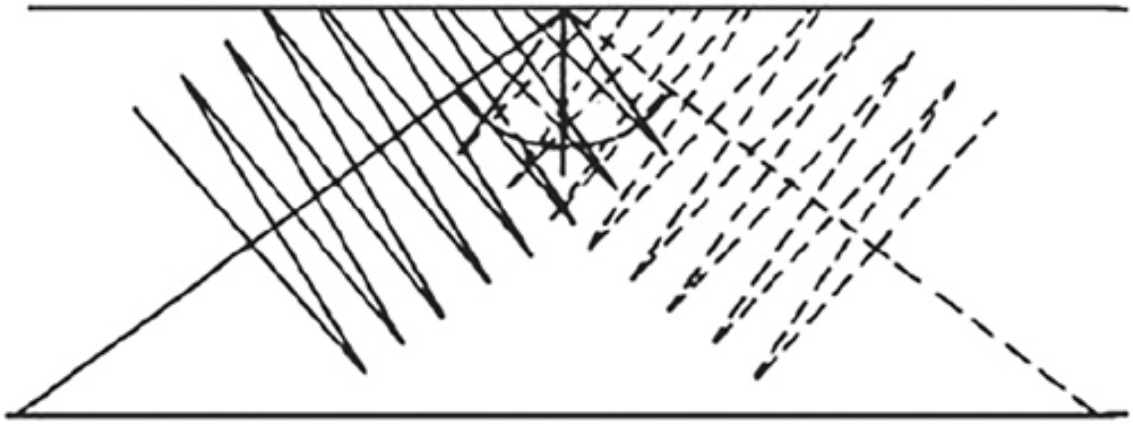
When a probe or other source of energy introduced into the waveguide, many modes are excited near the probe. Most of the modes have cutoff frequencies higher than the operating frequency, so these modes rapidly attenuate. As the wave propagates away from an input radiator, it resolves into the dominant mode (or in cases of multimode operation, a combination of a few low order modes).

22.8 Bouncing Wave Interpretation

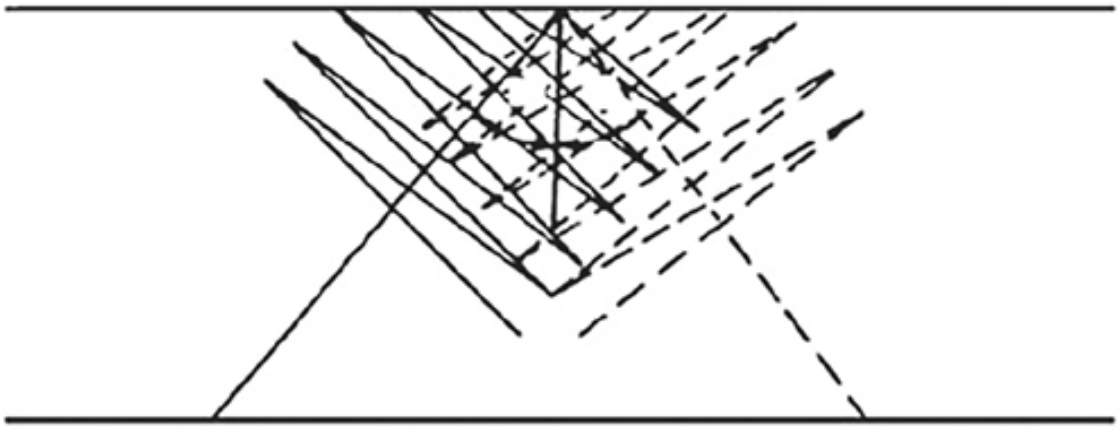
In a rectangular waveguide, a propagating mode is a plane wave that reflects from wall to wall. This is shown in Fig. 22.3. Three cases are shown the figure, high, medium, and low frequency. The angle of incidence with the waveguide wall increases as frequency drops. The angle rises toward 90 degrees as the cutoff frequency is approached from above. Below the cutoff frequency the angle is 90 degrees, so the wave bounces back and forth between the walls without propagating and the fields decay exponentially away from the power source that excites the waveguide.



A. High frequency



B. Medium frequency



C. Low frequency
(approaching cutoff)

= of incidence
= of reflection

FIGURE 22.3 Frequency effect on propagating wave.

22.9 Velocity and Wavelength in Waveguides

Figures 22.4 and 22.5 show the geometry for two wave components, simplified for the sake of illustration. There are three different wave velocities to consider with respect to waveguides:

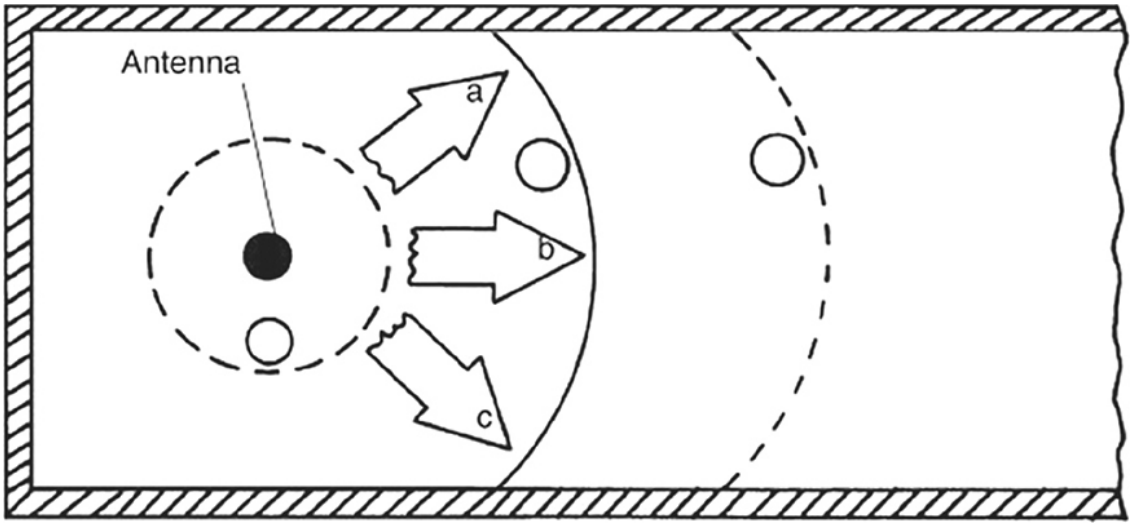


FIGURE 22.4 Antenna radiator in capped waveguide.

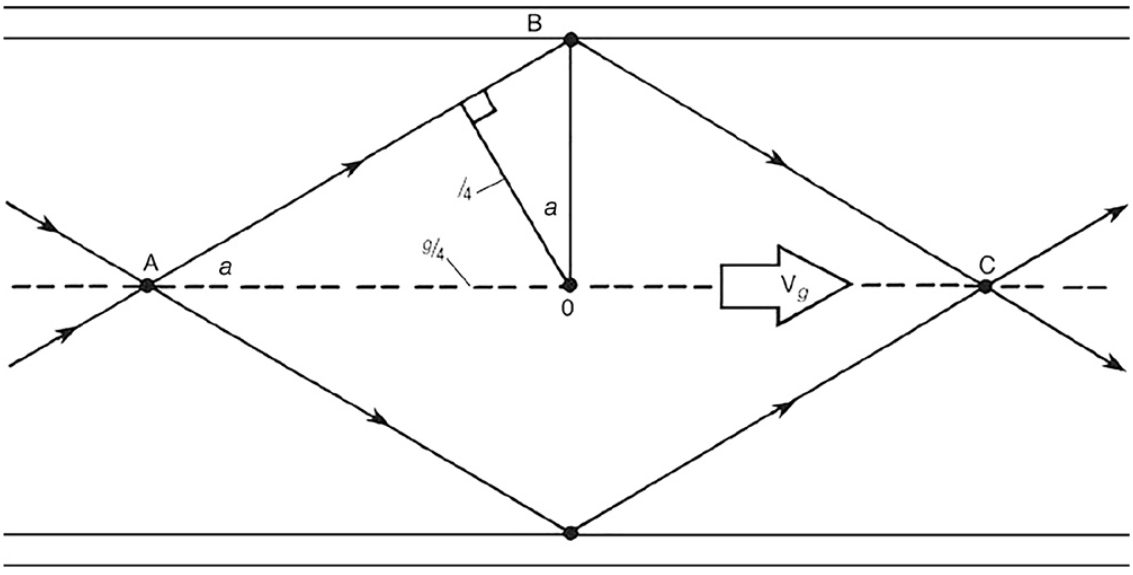


FIGURE 22.5 Wave propagation in waveguide.

- Free-space velocity *c*
- Group velocity v_g
- Phase velocity v_p

The group velocity is the straight-line velocity of propagation of the wave down the centerline (*z* axis) of the waveguides. The value of V_g is always less than c, because the actual path length taken, as the wave bounces back and forth, is longer than the straight-line path (i.e., path ABC in Fig. 22.5 is longer than path AC). The relationship between *c* and v_g is

$$v_g = c \sin \theta \quad (22.1)$$

where *c* is the free-space velocity ($3 \times 10^8 \text{ m/s}$) and *θ* is the angle of incidence of the mode in the waveguide. The phase velocity is the velocity of propagation of the spot on the waveguide wall where the wave impinges (point B in Fig. 22.5). Depending upon the angle of incidence, this velocity can actually be faster than both the group velocity and the speed of light.

To understand how phase velocity can be faster than the speed of light, let's look at an analogy. Consider an ocean beach, on which the waves arrive from offshore at an angle other than 90 degrees—in other words, the arriving wavefronts are not parallel to the shore. The arriving waves have a group velocity v_g. But as a wave hits the shore, it will strike a point far down the beach first. Then the "point of strike" races up the beach at a much faster phase velocity v_p, which is even faster than the group velocity. Similarly, in a microwave waveguide, the phase velocity can be greater than *c*, as can be seen from

$$v_p = \frac{c}{\sin\theta} \quad (22.2)$$

The three velocities are related by

$$c = \sqrt{\nu_p \nu_g} \quad (22.3)$$

Example 22.1 Calculate the group and phase velocities for an angle of incidence of 33 degrees.

Solution

Group velocity:

$$v_g = c\sin\theta = \left(3\times10^8\right)\left(\sin 33^\circ\right) = 1.6\times10^8 \text{ m/s}$$

Phase velocity:

$$v_p = c/\sin\theta = \frac{3\times10^8 \text{ m/s}}{\sin 33^\circ} = 5.51\times10^8 \text{ m/s}$$

For any wave the ratio of frequency to velocity is the wavelength. In free space, the wavelength is

$$\lambda_0 = \frac{c}{f} \quad (22.4)$$

In a waveguide, the wavelength increases to

$$\lambda = \frac{v_p}{f} \quad (22.5)$$

or alternatively, the wavelength in the waveguide can be found from the phase velocity and free space wavelength using

$$\lambda = \frac{v_p \lambda_0}{c} \quad (22.6)$$

Example 22.2 A 5.6-GHz microwave signal is propagated in a waveguide. Assume that the internal angle of incidence to the waveguide surfaces is 42 degrees. Calculate phase velocity, wavelength in unbounded free space, and wavelength in the waveguide.

Solution

(a) Phase velocity:

$$\nu_p = \frac{c}{\sin a} = \frac{3 \times 10^8 \text{m/s}}{\sin 42^\circ} = 10^8 \text{ m/s}$$

(b) Wavelength in free space:

$$\lambda_0 = \frac{c}{f} = \frac{3 \times 10^8 \text{ m/s}}{5.6 \times 10^9 \text{ Hz}} = 0.054 \text{ m}$$

(c) Wavelength in waveguide:

$$\lambda = \frac{\nu_p \lambda_0}{c} = \frac{\left(4.5 \times 10^8 \text{ m/s}\right)/(0.054 \text{ m})}{3 \times 10^8 \text{ m/s}} = 0.08 \text{ m}$$

Comparing, we find that the free-space wavelength is 0.054 m, while the wavelength inside of the waveguide has increased to 0.08 m.

22.10 Cutoff Frequency (f_c)

The propagation of signals in a waveguide is a function of the frequency of the applied signal. As the frequency drops, the angle of incidence increases toward 90 degrees. When the frequency drops to a point where the angle of incidence is 90 degrees, then group velocity is meaningless and propagation ceases. This occurs at the mode cutoff frequency. The mode still exists in the waveguide, but its amplitude decays exponentially away from a probe or junction that excites it and it is said to be an evanescent mode.

The mode equation for a rectangular waveguide is

$$\frac{1}{(\lambda_c)^2} = \left(\frac{m}{2a}\right)^2 + \left(\frac{n}{2b}\right)^2 \quad (22.7)$$

where
λ_c
is the longest wavelength that will propagate in the waveguide, *a* and *b* are the waveguide dimensions (see Fig. 22.1), and *m*, *n* are integers defining the number of half-wavelengths that will fit in the *a* and *b* dimensions for the *mn*th mode.

For the dominant mode (
TE_{10}
), the longest wavelength that will propagate in the waveguide is

$$\lambda_c = 2a \quad (22.8)$$

from which we can write an expression for the cutoff frequency:

$$f_{c,10} = \frac{c}{2a} \quad (22.9)$$

where *a* is the wider of the two waveguide dimensions.

Example 22.3 A rectangular waveguide has dimensions of
$3 \times 5 \text{ cm}$
. Calculate the
TE_{10}
mode cutoff frequency.

Solution

$$f_{c,10} = \frac{c}{2a} = \frac{(3 \times 108 \text{ m/s})}{(2)\left(5 \text{ cm} \times \dfrac{1 \text{ m}}{100 \text{ cm}}\right)} = 3 \text{ GHz}$$

Equation (22.7) assumes that the dielectric inside the waveguide is air. A more generalized form, which can accommodate other dielectrics, is

$$f_{c,mn} = \frac{1}{2\sqrt{\mu\epsilon}}\sqrt{\left(\frac{m}{a}\right)^2 + \left(\frac{n}{b}\right)^2} \quad (22.10)$$

where ϵ is the dielectric constant of the material inside the waveguide and μ is the permeability constant. One further expression for air-filled waveguide calculates the actual wavelength in the waveguide from a knowledge of the free-space wavelength and operating frequency:

$$\lambda_g = \frac{\lambda_0}{\sqrt{1 - \left(\frac{f_c}{f}\right)^2}} \quad (22.11)$$

where
f_c
is the waveguide mode cutoff frequency and f is the operating frequency.

Example 22.4 A waveguide with a 4.5 GHz cutoff frequency is excited with a 6.7-GHz signal. Find (a) the wavelength in free space and (b) the wavelength in the waveguide.

Solution

(a)

$$\lambda_0 = \frac{c}{f} = \frac{3 \times 10^8 \text{ m/s}}{6.7 \text{ GHz} \times \dfrac{10^9 \text{ Hz}}{1 \text{ GHz}}} = 0.0448 \text{ m}$$

(b)

$$\lambda_g = \frac{\lambda_0}{\sqrt{1-\left(\frac{f_c}{f}\right)^2}} = \frac{0.0448\ \text{m}}{\sqrt{1-\left(\frac{4.5\ \text{GHz}}{6.7\ \text{GHz}}\right)^2}} = 0.136\ \text{m}$$

22.11 Waveguide Impedance

Each waveguide mode has a characteristic impedance associated with it. The characteristic impedance of two-conductor transmission lines was developed in Chap. 6. For a waveguide, the characteristic impedance is approximately equal to the ratio of the electric and magnetic fields and converges as frequency increases to the intrinsic impedance
$\eta = \sqrt{\mu/\epsilon}$
of the dielectric material in the waveguide (Fig. 22.6). The TE mode characteristic impedance is

$$Z_{\text{TE}} = \frac{\eta}{\sqrt{1-\left(\frac{f_c}{f}\right)^2}} \quad (22.12)$$

and the TM mode characteristic impedance is

$$Z_{\text{TM}} = \eta\sqrt{1-\left(\frac{f_c}{f}\right)^2} \quad (22.13)$$

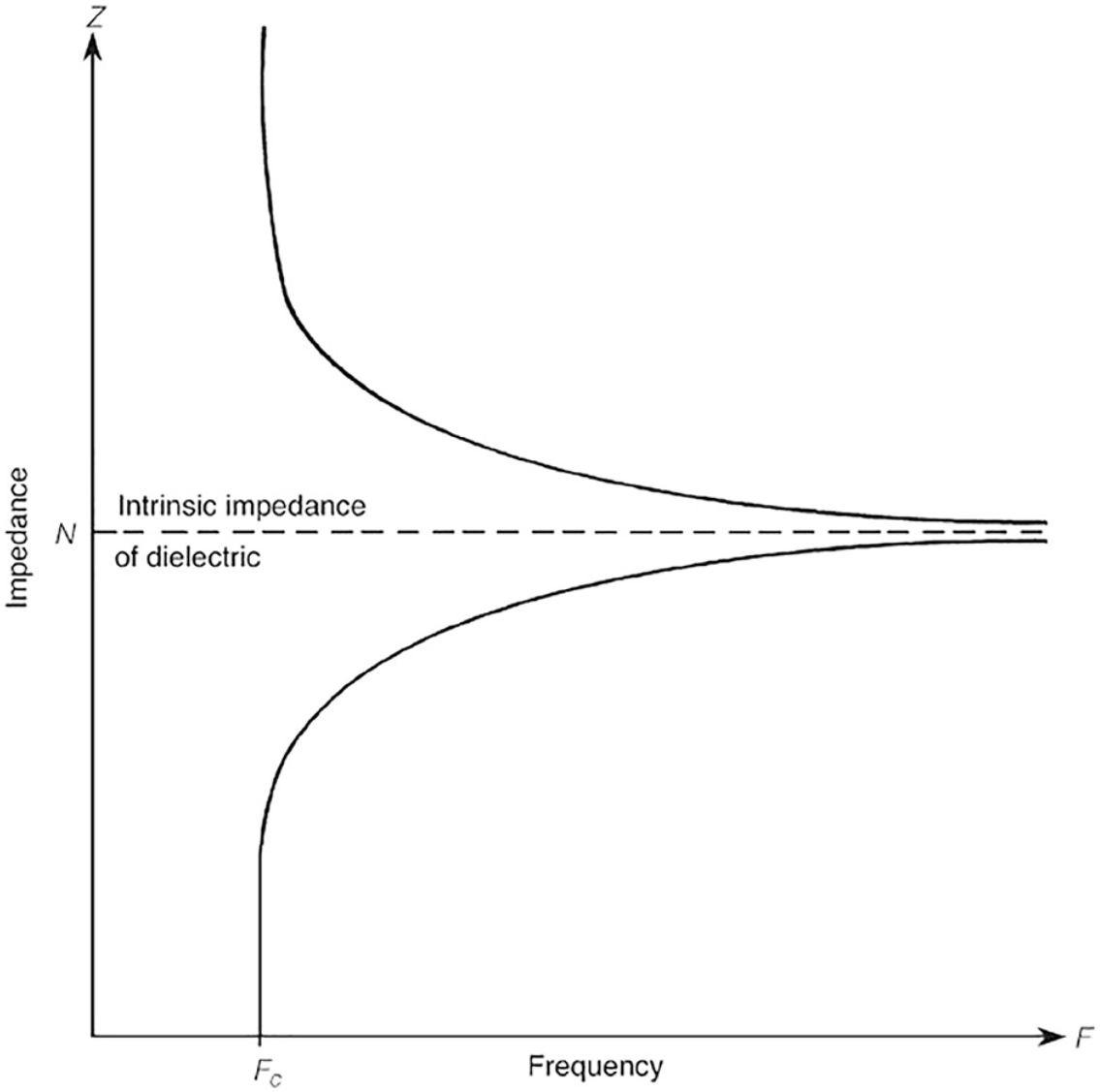
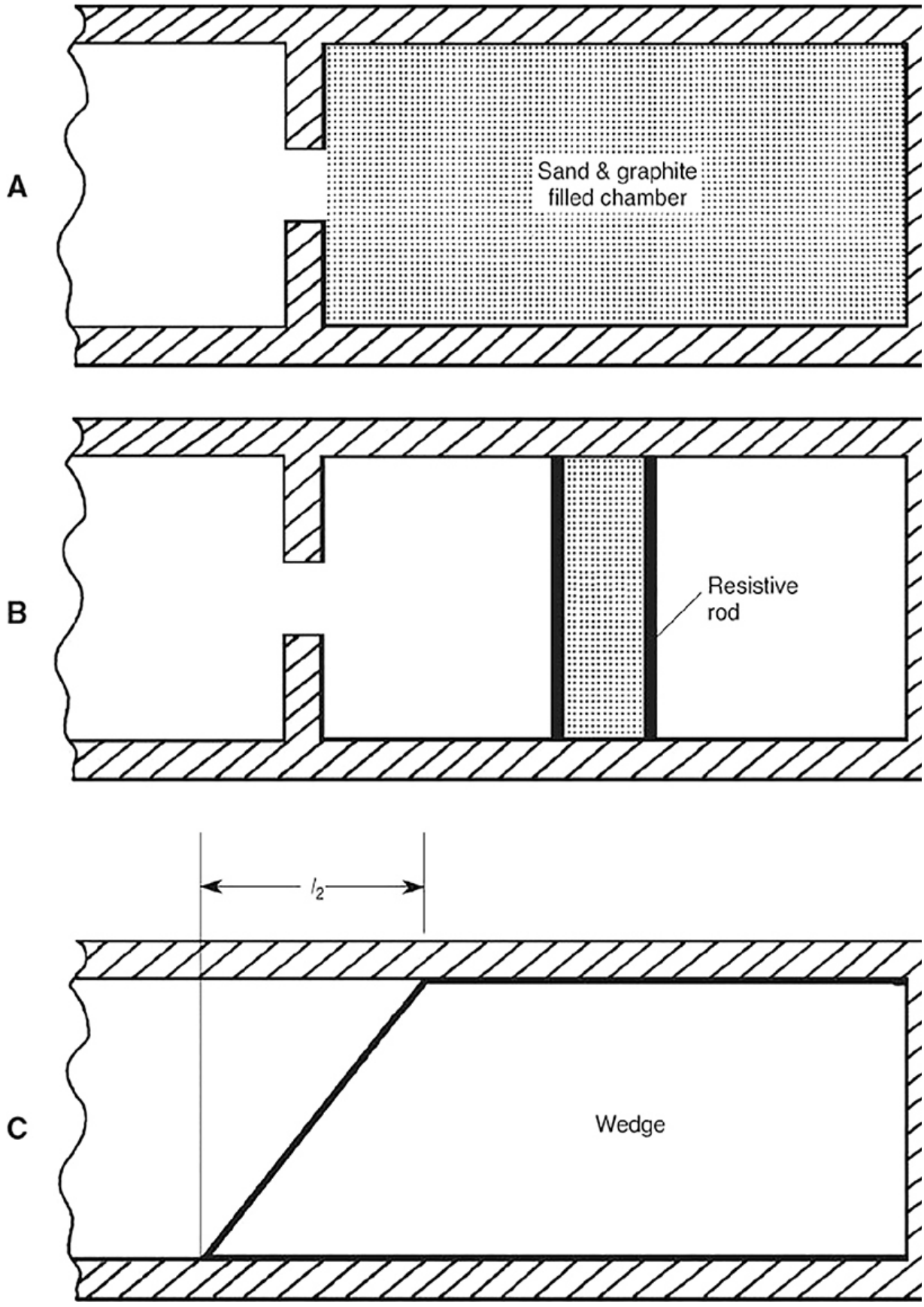


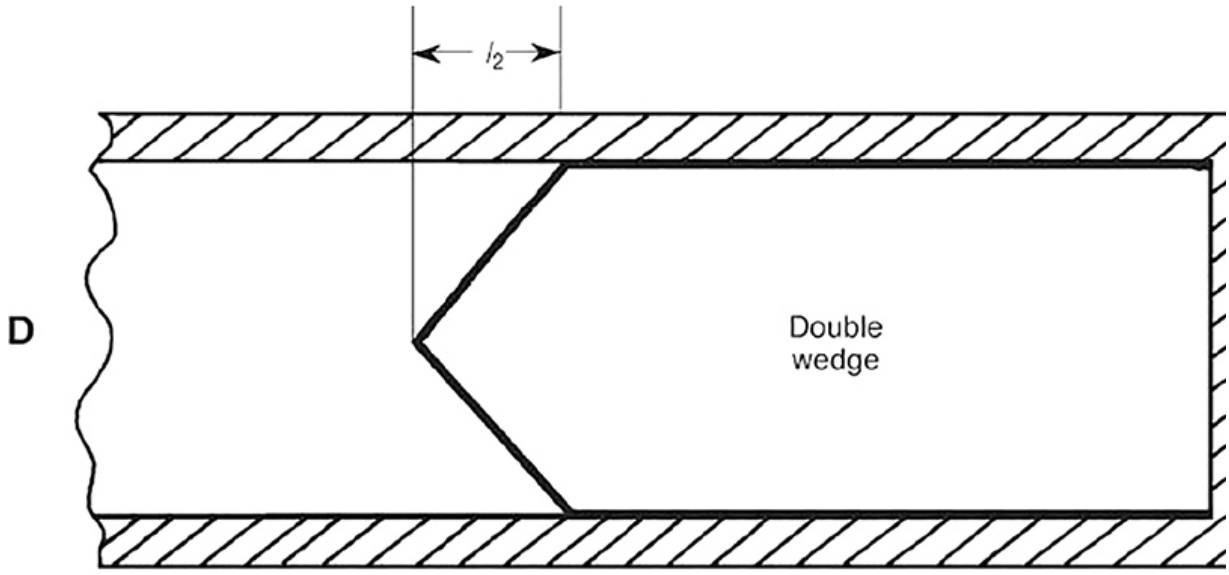
FIGURE 22.6 Impedance versus frequency.

22.12 Waveguide Terminations

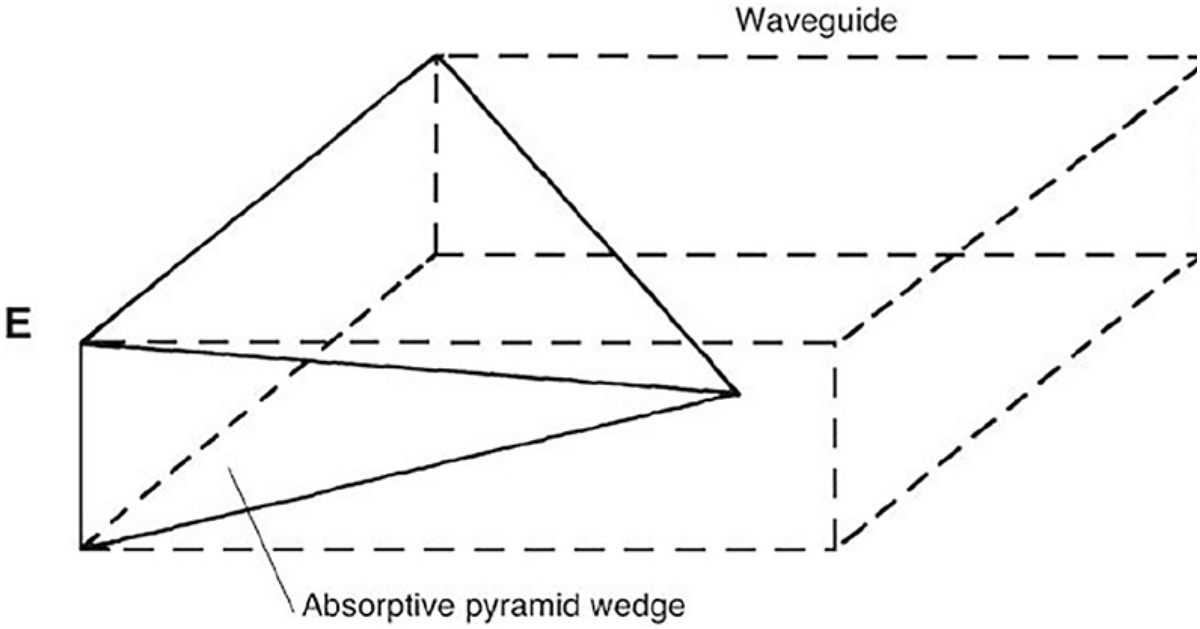
When an electromagnetic wave propagates along a waveguide, it must eventually reach the end of the guide. If the end is open, then the wave will propagate into free space. The horn radiator is an example of an unterminated waveguide. If the waveguide terminates in a metallic wall, the wave reflects back down the waveguide. The interference between incident and reflected waves forms standing waves (see Chap. 6). As the fields in the waveguide vary in time, the envelope of the amplitude of the field has a fixed pattern.

In order to prevent standing waves or, more properly, the reflections that give rise to standing waves, the waveguide must be terminated with a matching impedance. When a properly designed antenna is used to terminate the waveguide, it forms the matched load required to prevent reflections. Otherwise, a dummy load must be provided. Figure 22.7 shows several types of dummy loads.

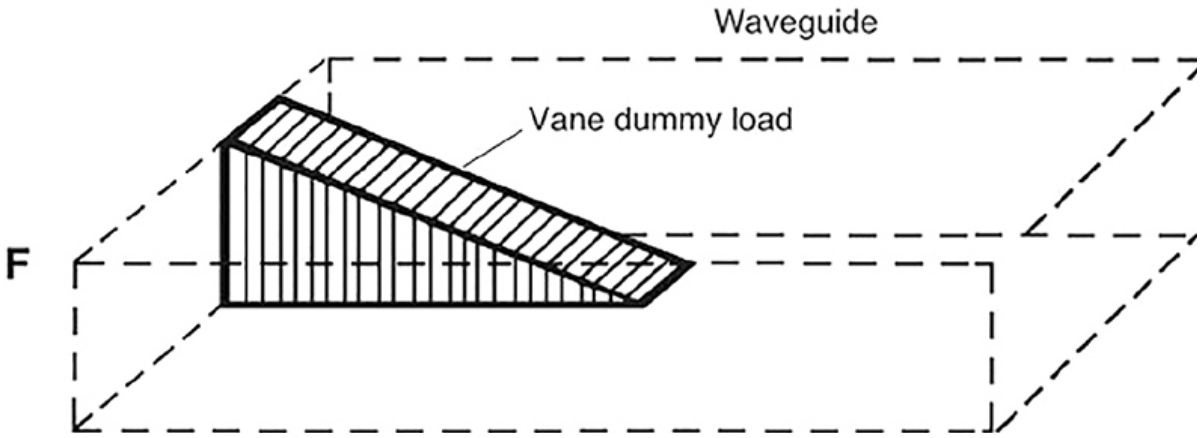

A
Sand & graphite
filled chamber
B
Resistive
rod
l_2
C
Wedge


l_2
Double
wedge

D


Waveguide
Absorptive pyramid wedge

E


Waveguide
Vane dummy load

F

FIGURE 22.7 Dummy loads. (A) Sand-graphite chamber. (B) Resistive rod. (C) Wedge. Dummy loads. (D) Double wedge. (E) Pyramid wedge. (F) Vane dummy load. (*Continued*)

The classic termination is shown in Fig. 22.7A. The resistive material is a mixture of sand and graphite. When the fields of the propagated wave enter the load, they cause currents to flow, which in turn cause heating. RF power dissipates in the sand-graphite mixture rather than being reflected back down the waveguide.

A second type of dummy load is shown in Fig. 22.7B. The resistor element is a carbonized rod critically placed at the center of the electric field. The E-field causes currents to flow, resulting in

I^2R

losses that dissipate the power.

Bulk loads, similar to the graphite-sand chamber, are shown in Fig. 22.7 C, D, and E. Using bulk material such as graphite or a carbonized synthetic material, these loads are used in much the same way as the sand load (i.e., currents set up, and

I^2R

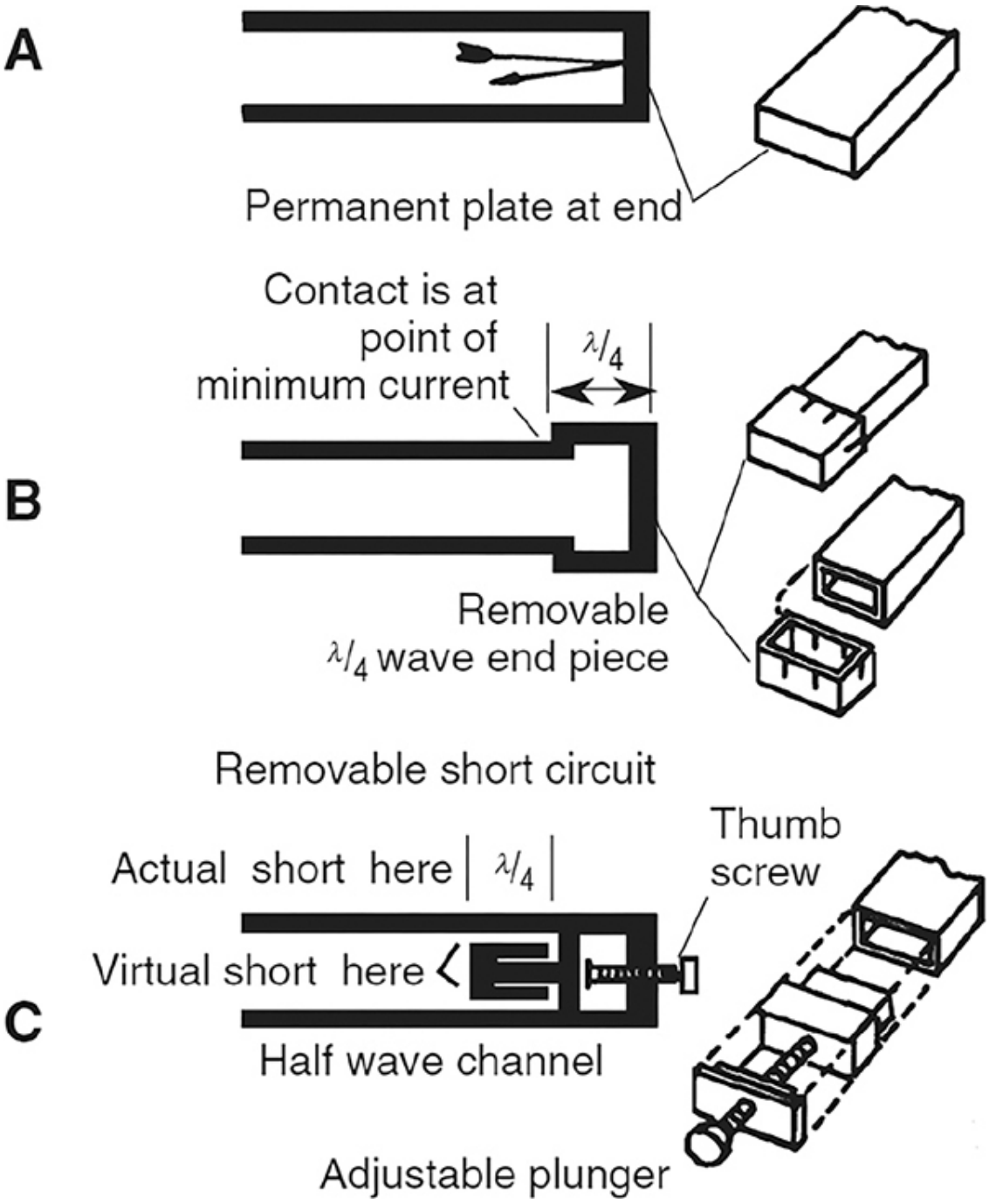
losses dissipate the power).

The resistive vane load is shown in Fig. 22.7F. The plane of the element is orthogonal to the magnetic lines of force. When the magnetic lines cut across the vane, currents are induced, which gives rise to the

I^2R

losses. Very little RF energy reaches the metallic end of the waveguide, so there is little reflected energy and a low VSWR.

There are situations where it isn't desirable to terminate the waveguide in a dummy load. Reflective terminations are shown in Fig. 22.8. The simplest is the permanent end plate shown in Fig. 22.8A. The metal cover must be welded or otherwise affixed through a very low resistance joint. At the high power levels often handled in transmitter waveguides, even small resistances can be important.



F IGURE 22.8 End terminations.

The end plate (shown in Fig. 22.8B) uses a quarter-wavelength cup to reduce the effect of joint resistances. The cup places the contact joint at a point that is a quarter-wavelength from the end. This point is a minimum-current node, so

I^2R

losses in the contact resistance become less important.

The adjustable short circuit is shown in Fig. 22.8C. The walls of the waveguide and the surface of the plunger form a half-wavelength channel. Because the metallic end of the channel is a short circuit, the impedance reflected back to the front of the plunger is 0 Ω, or nearly so. Thus, a virtual short exists at the points shown. By this means, the contact (or joint) resistance problem is overcome.

22.13 Waveguide Joints and Bends

Joints and bends in any form of transmission line or waveguide are seen as impedance discontinuities, and so are points at which reflection and mode conversion occur. Improperly formed bends and joints are substantial contributors to a poor VSWR.

Extensive runs of waveguide are sometimes difficult to make in a straight line. Although some installations do permit a straight waveguide, many others require directional change. This possibility is especially likely on shipboard installations. Figure 22.9A shows the proper ways to bend a waveguide around a corner. The radius of the bend must be at least two wavelengths at the lowest frequency that will be propagated in the system.

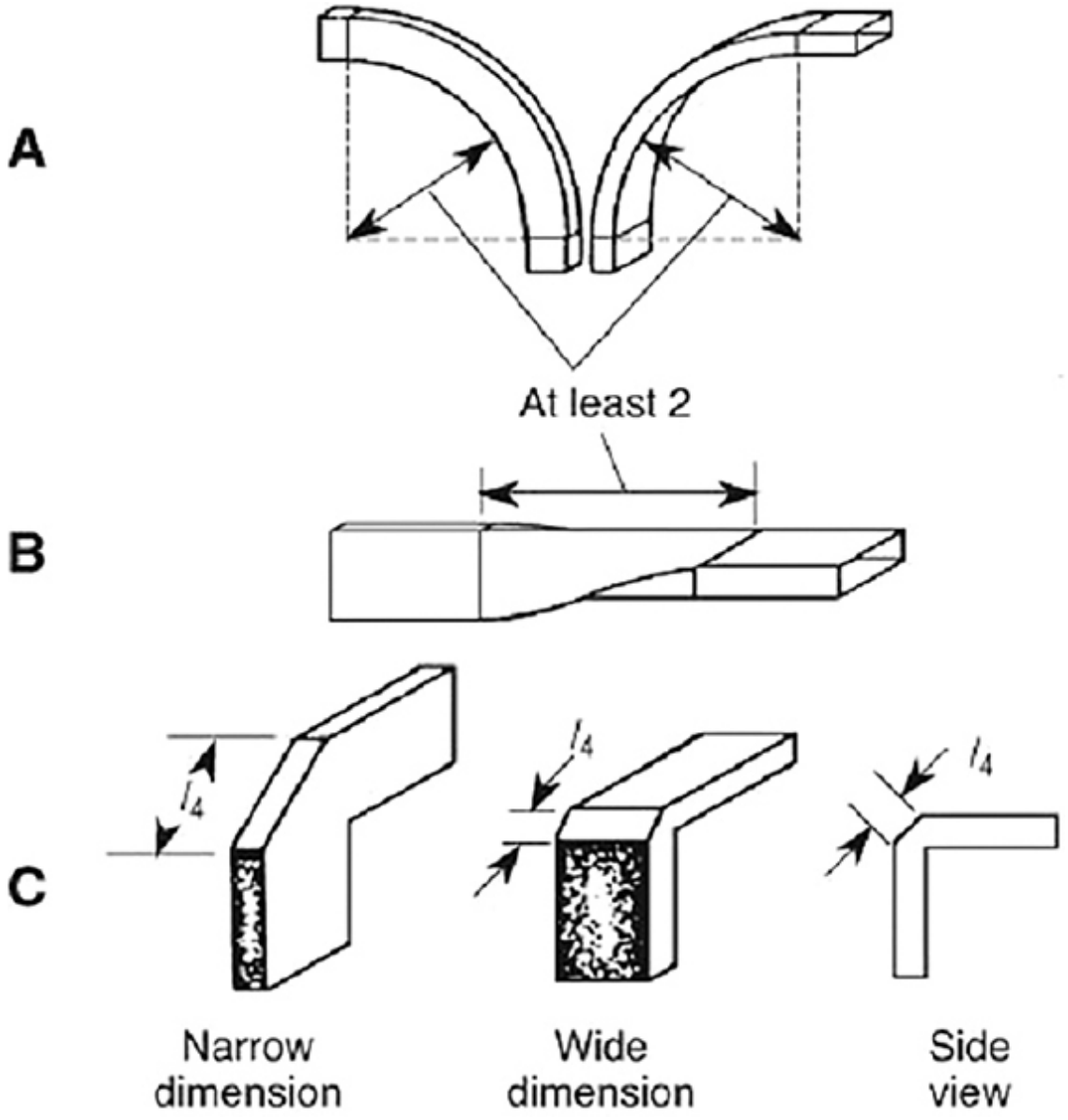


FIGURE 22.9 Bends in waveguide must be gentle.

The twist shown in Fig. 22.9B is used to rotate the polarity of the E- and H-fields by 90 degrees. This type of section is sometimes used in antenna arrays for phasing the elements. As in the case of the bend, the twist must be made over a distance of at least two wavelengths.

When an abrupt 90-degree transition is needed, it is better to use two successive 45-degree bends spaced one quarter-wavelength apart (see Fig. 22.9C). The interference between the direct reflection

of one bend and the inverted reflection of the other tends to cancel the overall reflected wave.

When two waveguide sections come together or a waveguide passes through a moving structure, joints are needed. Three types of common joints are used: permanent, semipermanent, and rotating.

To make a permanent joint, the two waveguide ends must be machined flat so that they can be butt-fitted together. A welded or brazed seam bonds the two sections together. To minimize reflections and VSWR, the interior surfaces are milled flat and then polished to a mirrorlike finish.

A semipermanent joint allows the joint to be disassembled for repair and maintenance, as well as allowing easier on-site assembly. The most common example of this class is the choke joint, as shown in Fig. 22.10. One surface of the choke joint is machined flat and is a simple butt-end planar flange. The other surface is the mate to the planar flange, but it has a quarter-wavelength circular slot cut at a distance of one quarter-wavelength from the waveguide aperture. The two flanges are shown in side view in Fig. 22.10A, and the slotted end view is shown in Fig. 22.10B. The method for fitting the two ends together is shown in the oblique view in Fig. 22.10C.

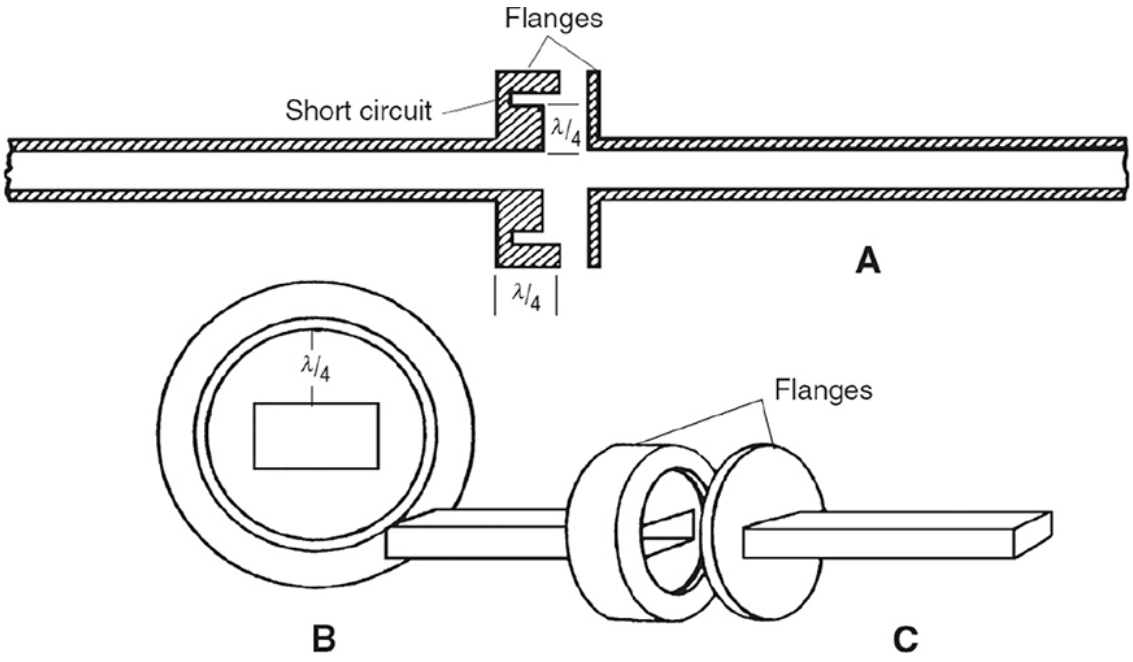


FIGURE 22.10 Choke joint.

Rotating joints are used in cases where the antenna has to point in different directions at different times. A common example is a rotating radar antenna.

The simplest form of rotating joint is shown in Fig. 22.11. A circular waveguide operating in the

TM_{01}

mode is used. The selected mode is symmetrical about the rotating axis. The waveguide rotates but the internal fields do not, thereby minimizing reflections. When a rectangular waveguide is required, a joint with transitions is used. Figure 22.12 shows a rotating joint consisting of two circular waveguide sections inserted between segments of a rectangular waveguide.

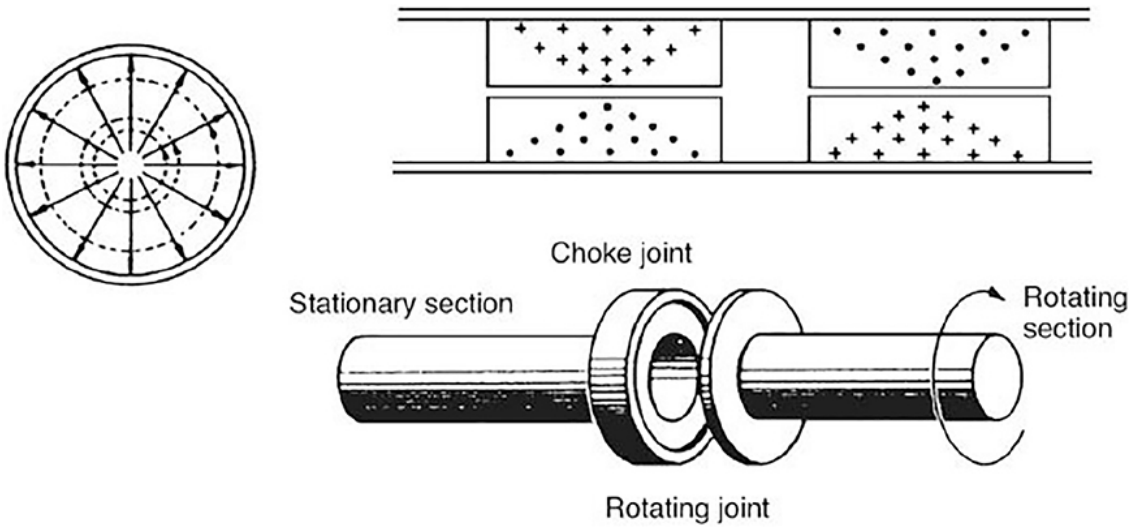


FIGURE 22.11 Basic rotating joint.

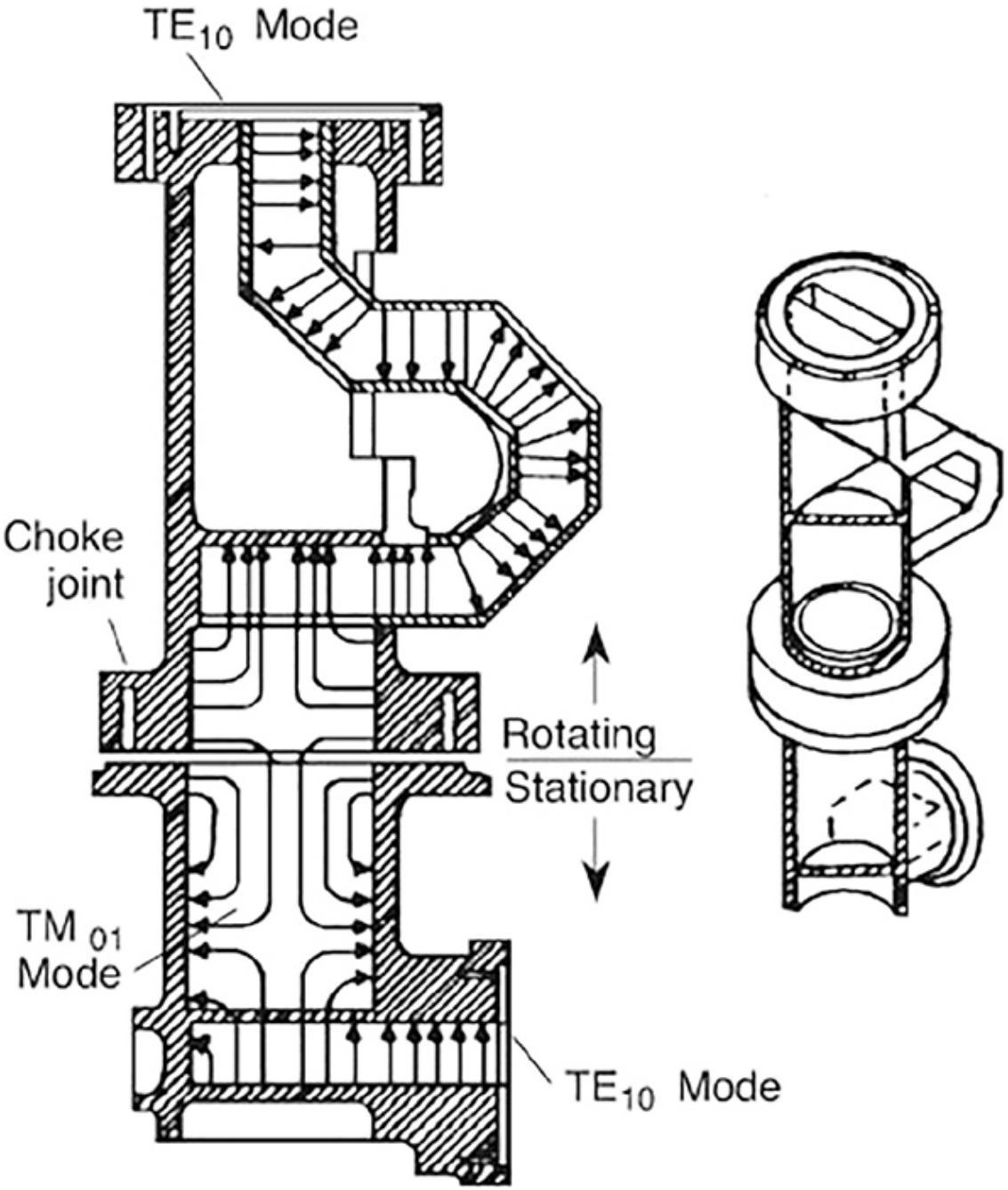


FIGURE 22.12 Representative practical rotating joint.

The rectangular input waveguide in Fig. 22.12 operates in the TE_{10} mode that is most efficient for rectangular waveguide. The E-field lines of force couple with the circular segment, thereby setting up a TM_{01}

mode wave. The
TM_{01}
mode has the required symmetry to permit coupling across the junction, where it meets another transition zone and is reconverted to
TE_{10}
mode.

22.14 Waveguide Coupling Methods

Most waveguide systems require an input or output coupling. There are three basic types of coupling used in a microwave waveguide: capacitive (or probe), inductive (or loop), and aperture (or slot).

Capacitive coupling is shown in Fig. 22.13. A vertical radiator is inserted into one end of the waveguide. For a narrowband system, the probe is a quarter wavelength rod analogous to the vertical monopoles employed at lower frequencies. A characteristic of this type of radiator is that the E-field is parallel to the waveguide top and bottom surfaces, satisfying the E-field boundary condition.

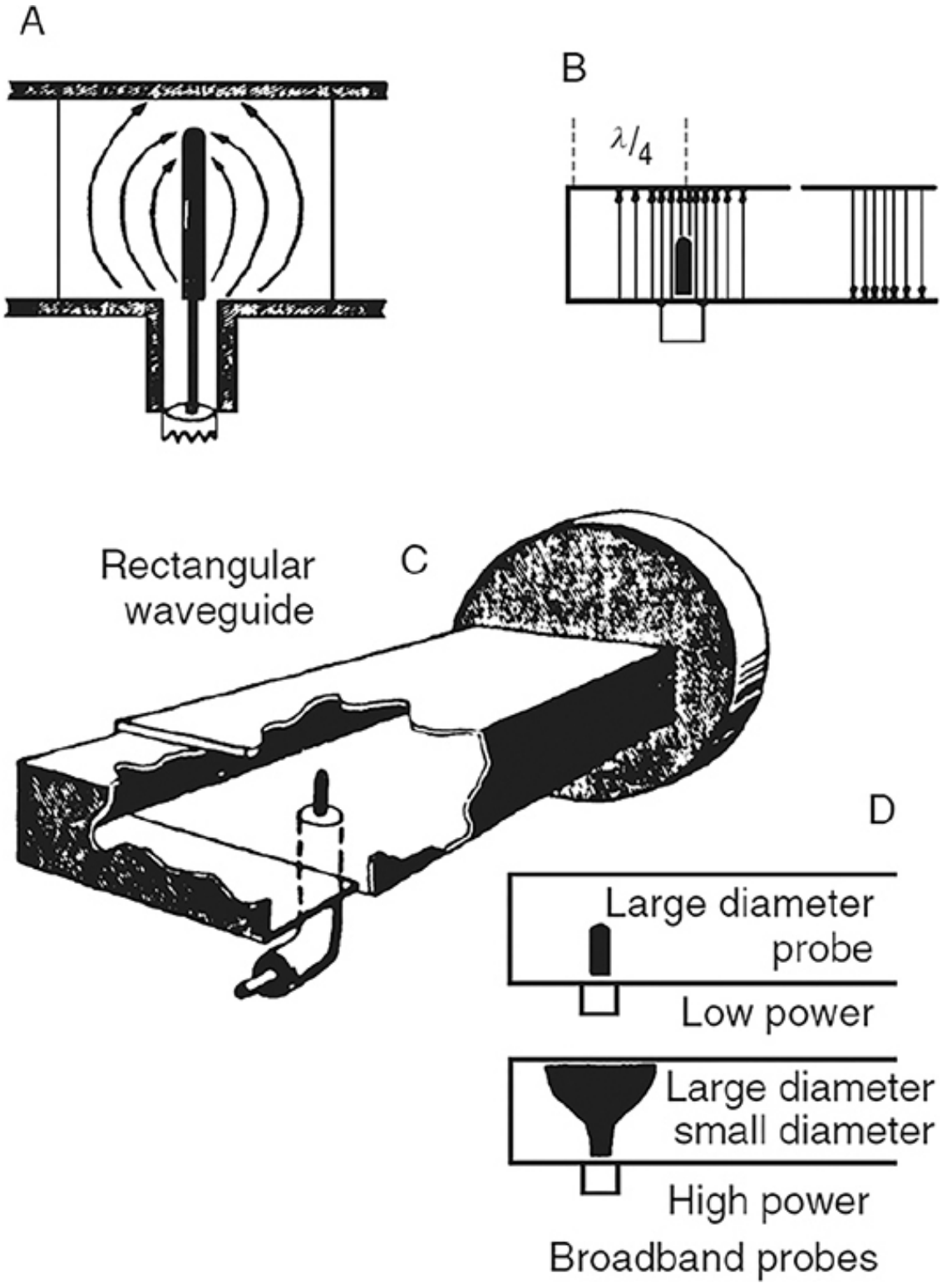


FIGURE 22.13 Probe (capacitive) coupling.

The radiator is a quarter-wavelength from the rear wall (Fig. 22.13B). By traversing the quarter-wave distance (90-degree phase shift), reflecting off the rear wall (180-degree phase shift), and back (another 90-degree phase shift), the wave undergoes a total phase shift of one complete cycle, or 360 degrees. The reflected wave arrives back at the radiator in phase to reinforce the outgoing wave. Some waveguides have an adjustable end cap (Fig. 22.13C) to accommodate multiple frequencies. The end cap position is varied as required for signals of different wavelengths.

Figure 22.13D shows high- and low-power broadband probes. Broadbanding is accomplished by attention to the diameter-to-length ratio. The degree of coupling can be varied by adjusting the probe length, position, or adding a shield.

Inductive, or loop coupling, is shown in Fig. 22.14. A small loop of wire or other conductor is placed such that the number of magnetic flux lines it cuts is maximized. This form of coupling is popular on microwave receiver antennas as a way of making a waveguide-to-coaxial cable transition. In some cases, the loop is formed by the pigtail lead of a detector diode that, when combined with a local oscillator, downconverts the microwave signal to an intermediate frequency (IF).

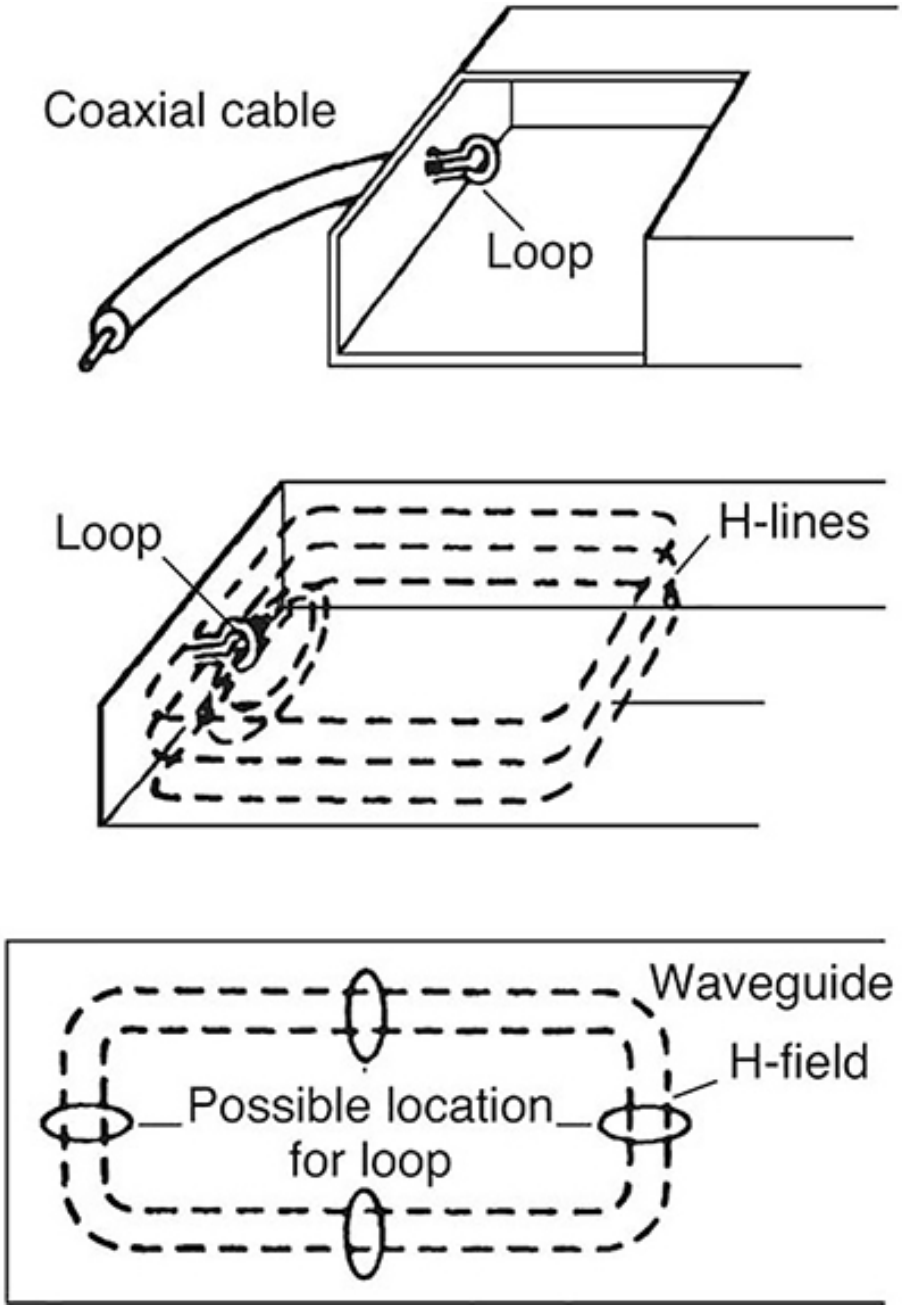


FIGURE 22.14 Loop (inductive) coupling.

Aperture, or slot, coupling is shown in Fig. 22.15. This is used to couple two sections of waveguide, as on an antenna feed system. Slots can be designed to couple electric, magnetic, or electromagnetic fields. In Fig. 22.15, slot A is placed at a point where the E-field peaks, so it allows electrical field coupling. Slot B is at a

point where the H-field peaks, so it allows magnetic field coupling. Slot C allows electromagnetic field coupling.

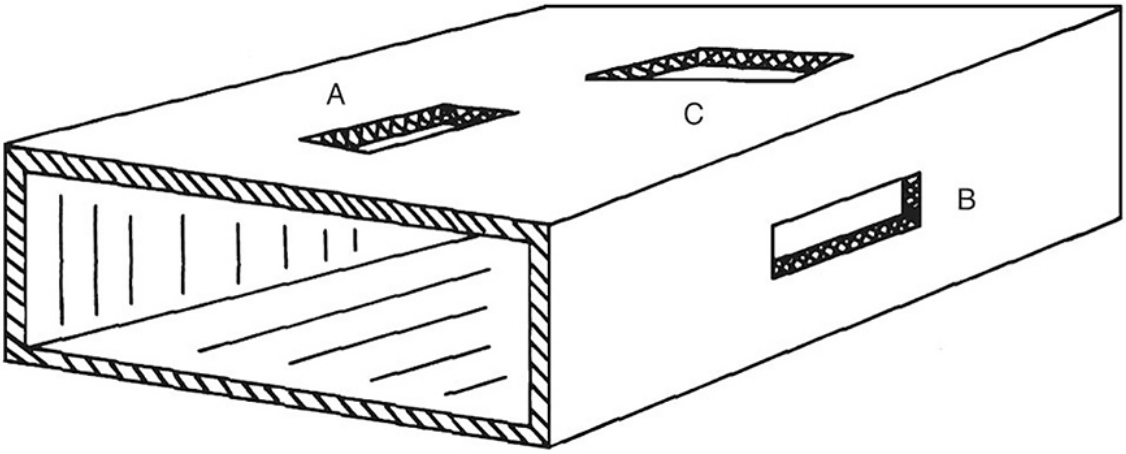


FIGURE 22.15 Slot coupling.

Slots can be characterized according to whether they are radiating or nonradiating. A nonradiating slot is cut at a point that does not interrupt the flow of currents in the waveguide walls. The radiating slot interrupts currents in the walls and is used for some types of waveguide antennas.

CHAPTER 23
Microwave Antennas

Antennas are used in communications and radar systems over a wide range of radio frequencies. Antennas are used until operating frequencies reach infrared and visible light, at which point optical techniques take over. Microwaves begin the transition between radio waves and visible light, so as might be expected microwave technology makes use of techniques from both worlds. Both dipoles and parabolic reflectors, for example, are used in microwave systems.

While lower frequencies are well covered for many applications by the wire dipole antenna and its derivatives, at microwave frequencies, a wide variety of antennas are available. Smaller wavelengths mean that electrically large antennas with high gain become practical and can be used to overcome blockage effects and high losses for non-line-of-sight transmissions that occur at high frequencies.

Like RF energy at other frequencies, microwave RF energy at high enough power can be dangerous. The principal problem is tissue heating. Eyes are especially sensitive to microwave energy. Because of their relatively high gain, microwave antennas can produce hazardous field strengths in close proximity. As a safety precaution, avoid being close to a radiating antenna when power levels may be above safe exposure limits. Operating antennas should have radiation hazard warnings as required by applicable radiations.

23.1 Dipole Antenna

A dipole operating at microwave frequencies is shown in Fig. 23.1. The antenna radiating element consists of a short conductor at the

end of a section of waveguide. Although most low-frequency dipoles are a half-wavelength, microwave dipoles might be exactly a half-wavelength, less than a half-wavelength, or greater than a half-wavelength, depending upon the application.

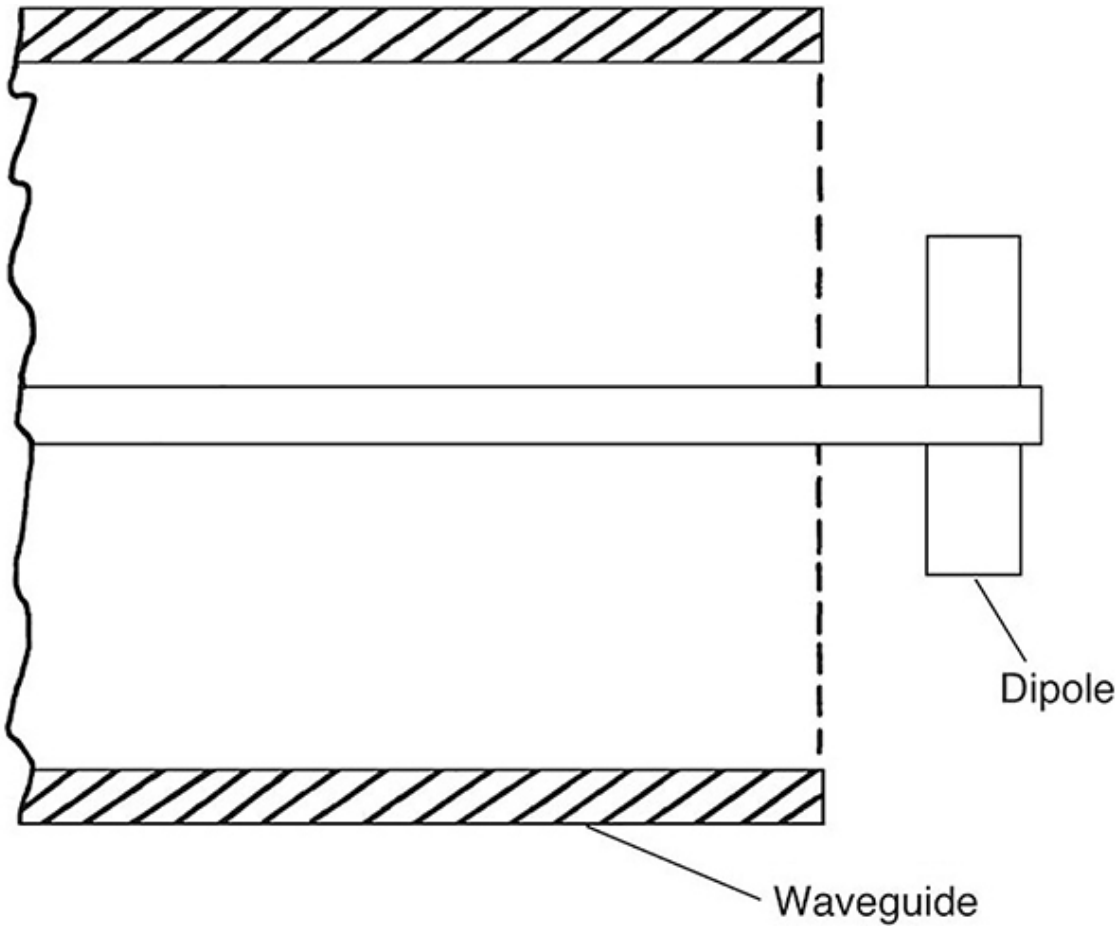


FIGURE 23.1 Microwave dipole radiator.

23.2 Horn Antenna

The horn antenna is a tapered termination of a length of waveguide that provides the impedance transformation between the waveguide and free space (see Fig. 23.2A–C). Horn radiators are used both as antennas in their own right and as illuminators for reflector antennas. Horn antennas are usually matched to the waveguide with an SWR of around 1.5:1.

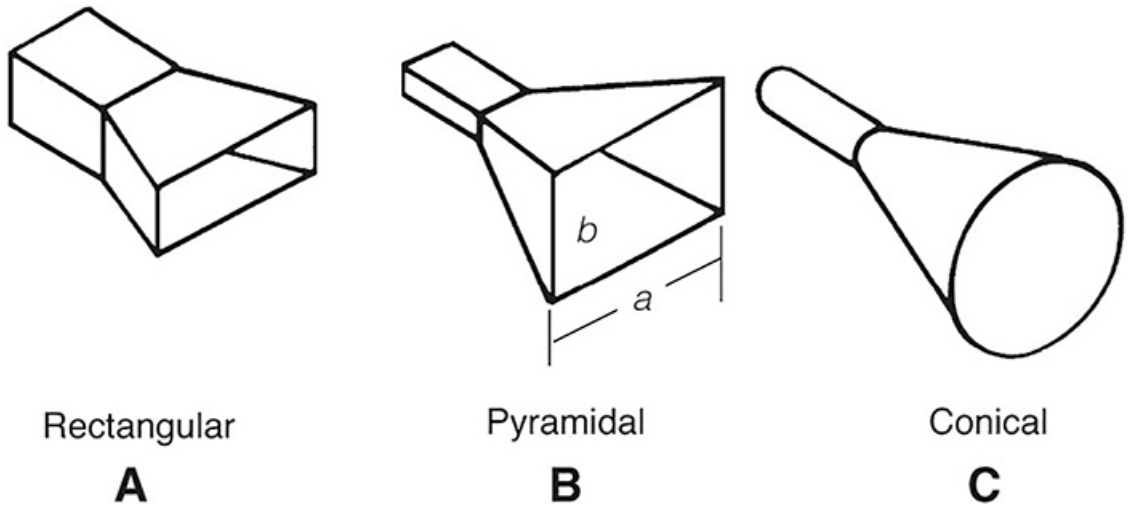


FIGURE 23.2 Horn radiators.

The gain of a horn radiator is proportional to the area of the flared open flange (

$A = ab$

in Fig. 23.2B), and inversely proportional to the square of the wavelength:

$$G \simeq 0.8A\frac{4\pi}{\lambda^2} \quad (23.1)$$

where *A* is the area of the open face of the flange. The 3-dB beamwidth for vertical and horizontal extents can be approximated from

$$\Phi_V = \frac{51\lambda}{b} \text{ degrees} \quad (23.2a)$$

$$\Phi_H = \frac{70\lambda}{a} \text{ degrees} \quad (23.2b)$$

where *a* and *b* are the dimensions of the flared flange.

Related to the horn is the cavity antenna of Fig. 23.3. A quarter-wavelength radiating element extends from a waveguide or transmission line connector into a resonant cavity. The radiator element is placed a quarter-wavelength into a resonant cavity,

spaced a quarter-wavelength from the rear wall of the cavity. A tuning disk provides a limited tuning range for the antenna by altering cavity dimensions. Gains to about 6 dB are possible with this arrangement.

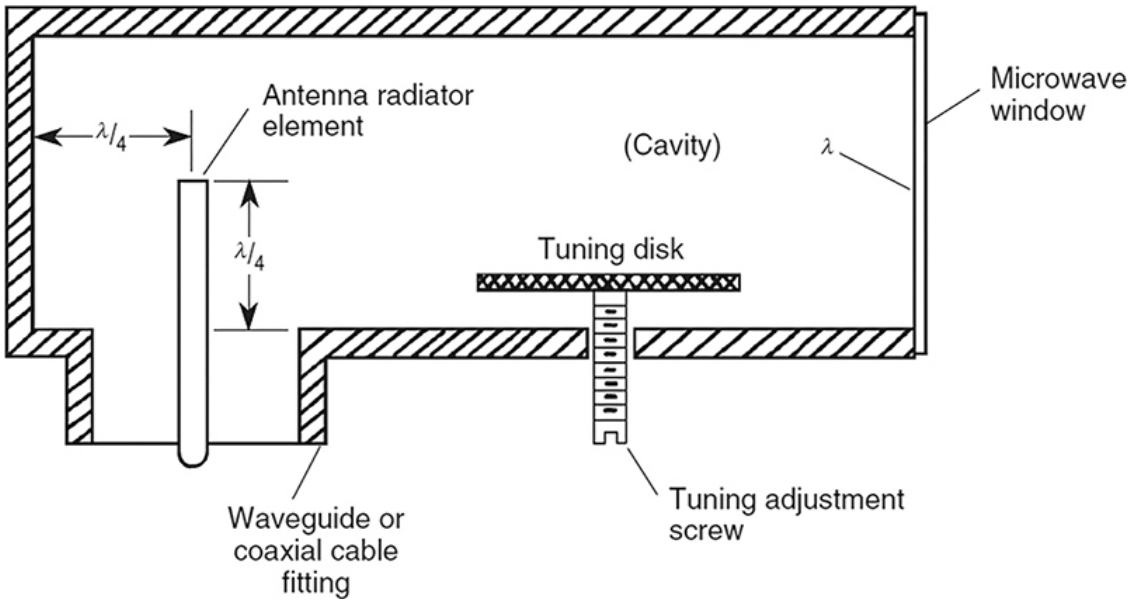


FIGURE 23.3 Cavity antenna.

23.3 Reflector Antennas

At microwave frequencies, it becomes possible to use reflector antennas. Reflectors are possible at lower frequencies, but the large physical dimensions that the longer wavelengths require, making them impractical. Reflector antennas are a type of aperture antenna, where the radiation emanates from an electrically large illuminated region.

Reflectors are used for radar, satellite links, point-to-point microwave links, and radio astronomical telescopes. An assortment of reflector shapes can be seen in Fig. 23.4. Solid metallic reflector surfaces or wire mesh may be used. When mesh is used, the holes in the mesh must be about $\lambda/12$ or smaller.

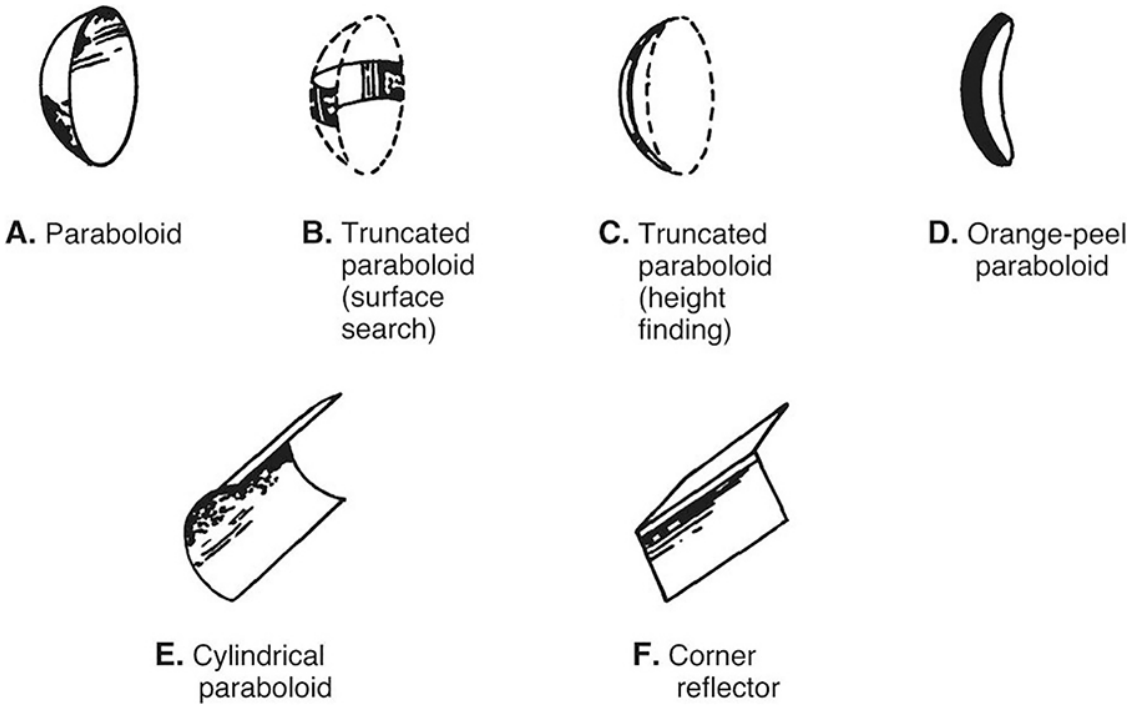


FIGURE 23.4 Reflector antennas.

In Fig. 23.5 we see the corner reflector antenna, used primarily in the high-UHF and low-microwave region. A dipole element receives reflected wavefronts from many regions of the reflector surface. The corner reflector is a simplified version of the common parabolic reflector.

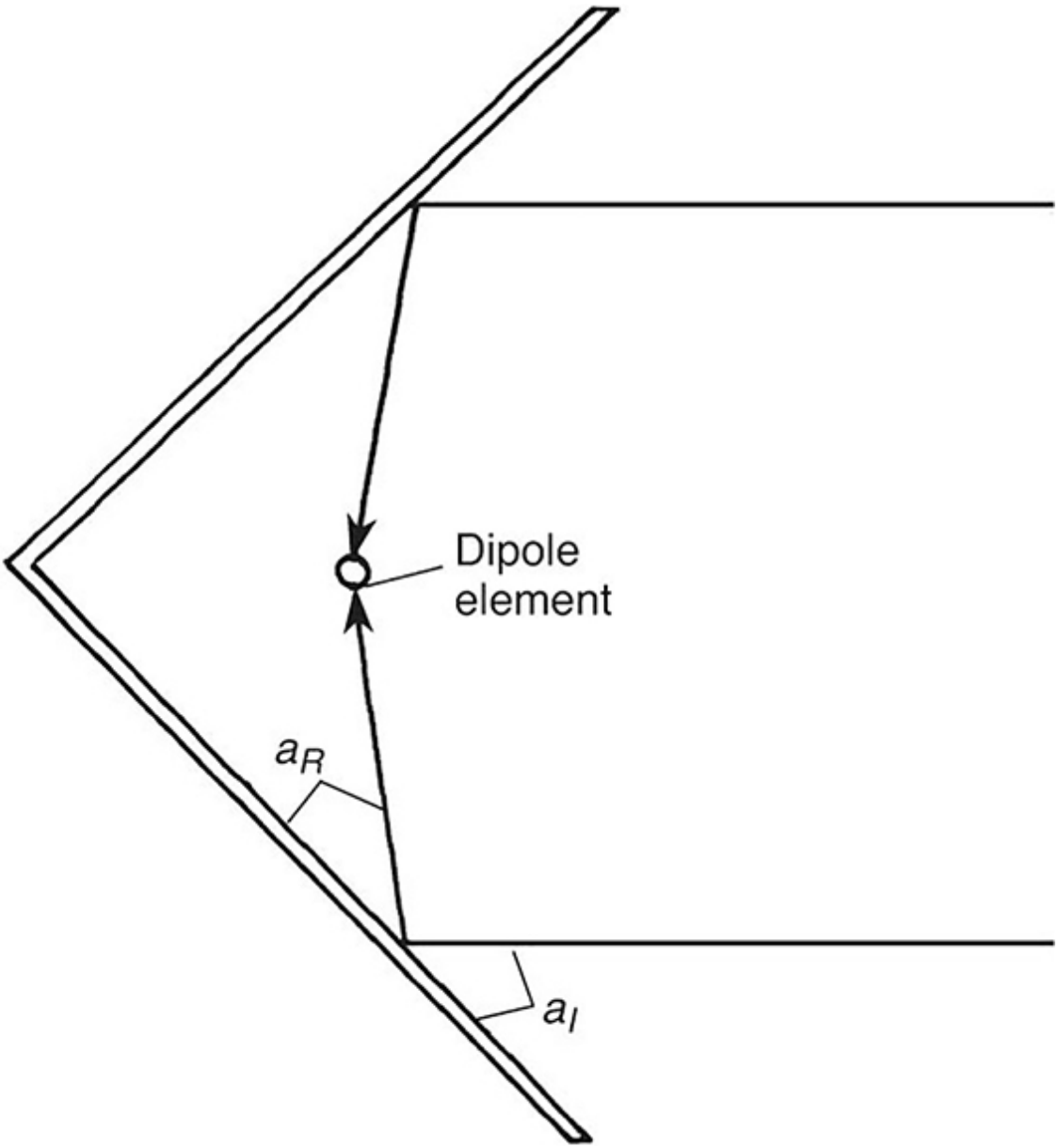
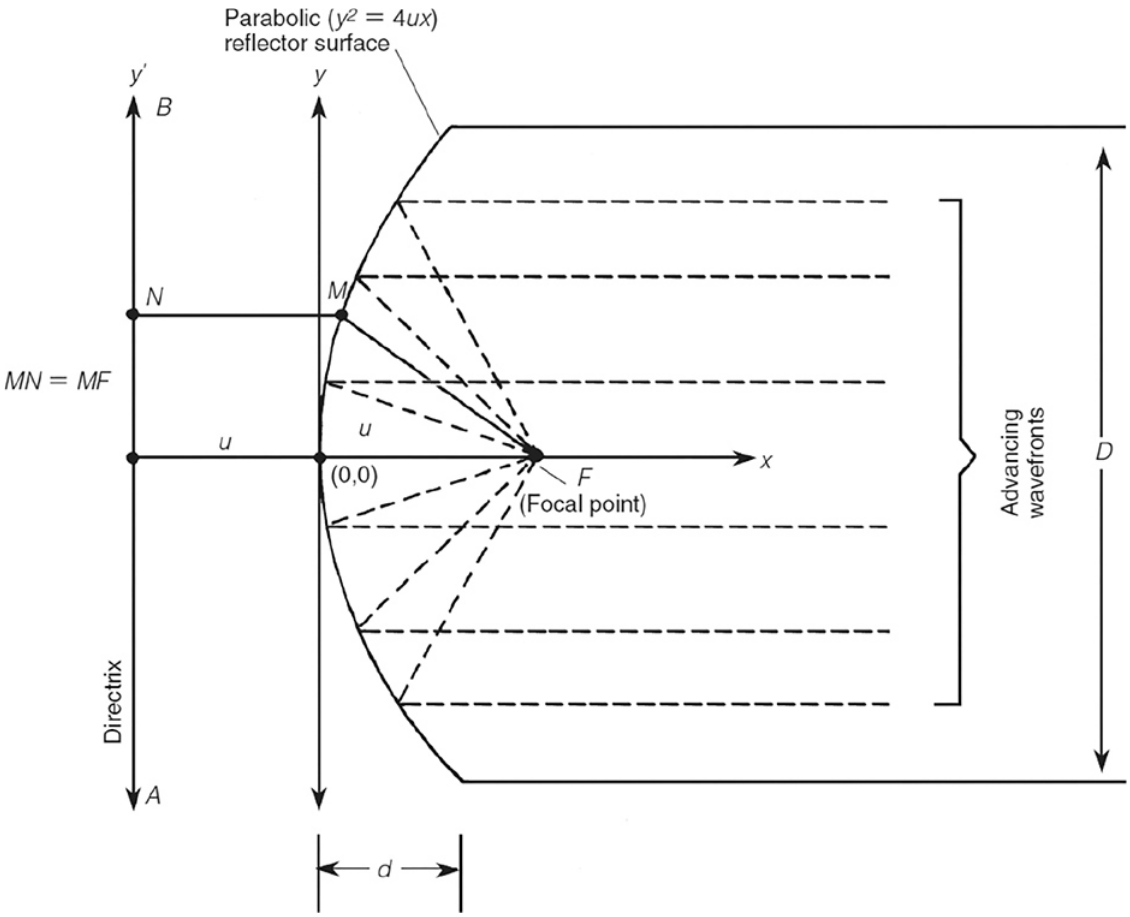


FIGURE 23.5 Corner reflector.

23.4 Parabolic Dish Antennas

The parabolic reflector antenna is one of the most widely used microwave antennas, and one that normally comes to mind when thinking of microwave systems. The dish antenna has a paraboloid shape, as defined by Fig. 23.6.



F IGURE 23.6 Ray tracing shows operation of parabolic antenna.

In the figure, the dish surface is positioned such that the center is at the origin (0,0) of an $x - y$ coordinate system. We place a second vertical axis called the directrix (y') a distance behind the surface equal to the focal length (u). The paraboloid surface follows the function $y^2 = 4uX$, and has the property that a line from the focal point *F* to any point on the surface is the same length as a line from that same point to the directrix. If a radiator element is placed at the focal point *F*, it will illuminate the reflector surface, causing wavefronts to be propagated away from the surface in phase. Similarly, wavefronts intercepted by the reflector surface are reflected to the focal point.

The gain of a parabolic antenna is a function of several factors, including dish diameter, feed illumination, and surface accuracy. The key parameters are the dish diameter *D* and the *F/D* ratio, where *F* is the distance from the dish apex to the focal point, or the focal length. The dish diameter *D* is large compared with the depth. Surface accuracy refers to the degree of surface irregularities. For commercial antennas, 1⁄8-wavelength surface accuracy is usually sufficient, although on certain radar antennas the surface accuracy specification is tighter.

Feed illumination refers to how evenly the feed element radiates to the reflector surface. For circular parabolic dishes, a circular waveguide feed produces optimum illumination. Rectangular waveguides are not as good a match. The circular waveguide TE_{11} mode is preferred. For best performance, the illumination should drop off evenly from the center to the edge, with the edge being around 10 dB down from the center. The diameter, length, and beamwidth of the radiator element (or horn) must be optimized for the specific *F/D* ratio of the dish. The cutoff frequency of the circular waveguide used as the horn feed is approximated from

$$f_c = \frac{175,698}{d} \quad (23.3)$$

where *d* is the inside diameter of the circular feedhorn in mm.

The gain of a lossless parabolic dish antenna is found from

$$G = \frac{\eta_{\mathrm{ap}} (\pi D)^2}{\lambda^2} \quad (23.4)$$

where *D* is the dish diameter, *λ* is the wavelength in the same units as *D*, and η_{ap} is the aperture efficiency (0.4 to 0.7).

For receiving antennas, the effective aperture is the relevant specification and is found from

$$A_e = \eta_{\text{ant}} \pi (D/2)^2 \quad (23.5)$$

where antenna efficiency is the product of aperture efficiency and radiation efficiency and in the lossless case
η_{ant}
is equal to the aperture efficiency
η_{ap}
. For low-noise applications like satellite communications, we tighten the feed illumination pattern to reduce the pickup of thermal noise from warm ground behind the dish, and use a feed illumination pattern taper higher than 10 dB. In this case, the aperture efficiency drops, reducing the signal level slightly, but the signal-to-noise ratio increases due to lower noise.

The 3-dB beamwidth of the parabolic dish antenna is approximated by

$$\text{BW} = \frac{70\lambda}{D} \quad (23.6)$$

The focal length is related to the dish diameter *D* and the depth *d* by

$$F = \frac{D^2}{16d} \quad (23.7)$$

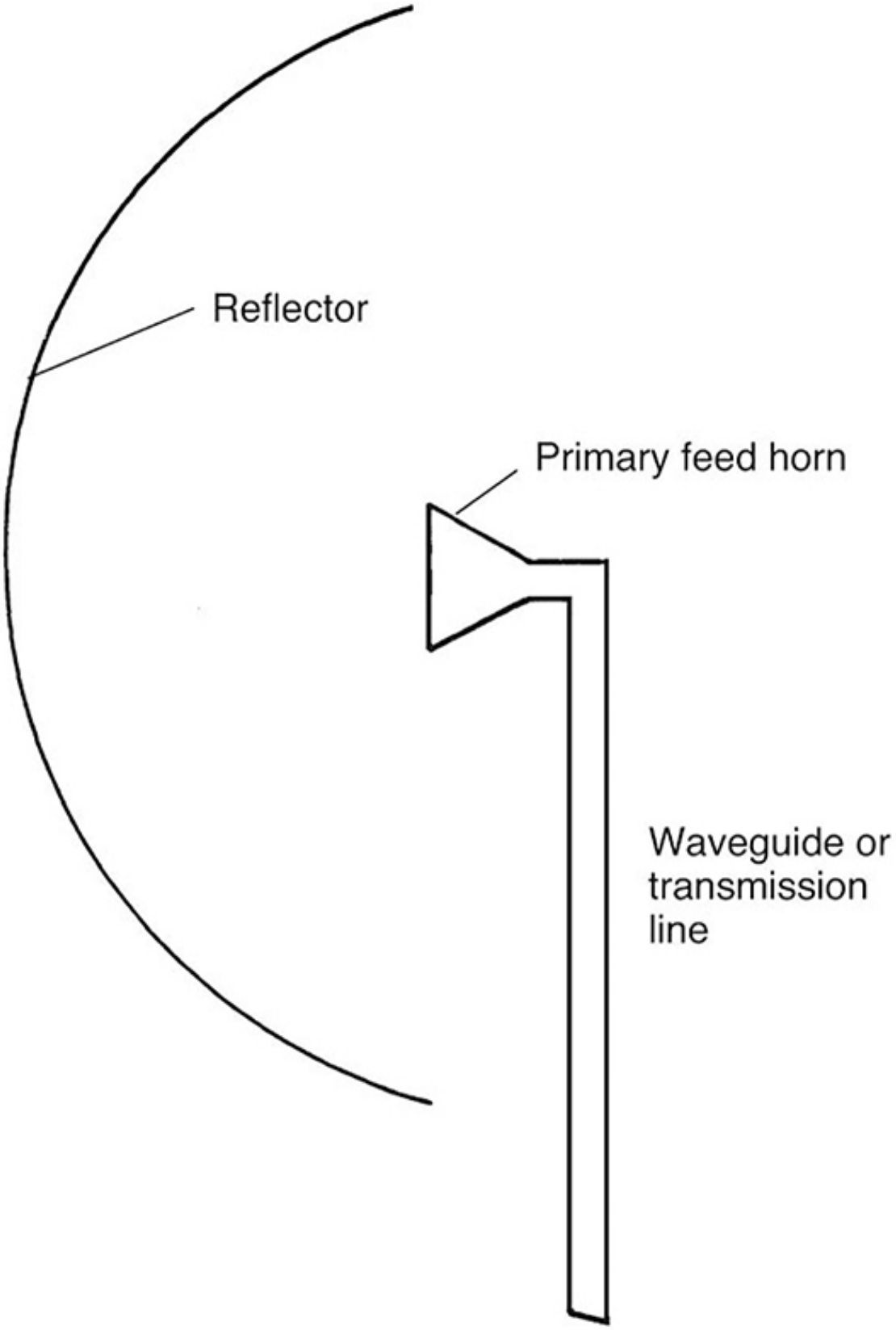
The antenna pattern radiated by the antenna is similar to Fig. 5.16. With horn illumination, the sidelobes are around 23 to 28 dB below the main lobe, or 10 to 15 dB below isotropic. Of the energy radiated by the parabolic dish, 50 percent is within the 3-dB beamwidth, and 90 percent is between the first nulls bracketing the main lobe.

If a dipole element is used for the feed device, then a splash plate is placed a quarter-wavelength behind the dipole in order to improve illumination. The splash plate must be several wavelengths in diameter and is used to reflect the dipole's back lobe toward the reflector surface. When added to the half-wave phase reversal

inherent in the reflection process, the two-way quarter-wavelength adds another half-wavelength so that the reflected wave is in phase with the front lobe.

23.4.1 Parabolic Dish Feed Geometries

Figures 23.7 and 23.8 show two methods for feeding parabolic dish antennas, independent of the choice of radiator (horn, dipole, etc.). In Fig. 23.7, the radiator element is placed at the focal point, and a waveguide or transmission line is routed to it. This method is used in low-cost installations such as home satellite TV receive-only (TVRO) antennas.



F IGURE 23.7 Parabolic antenna conventional feed.

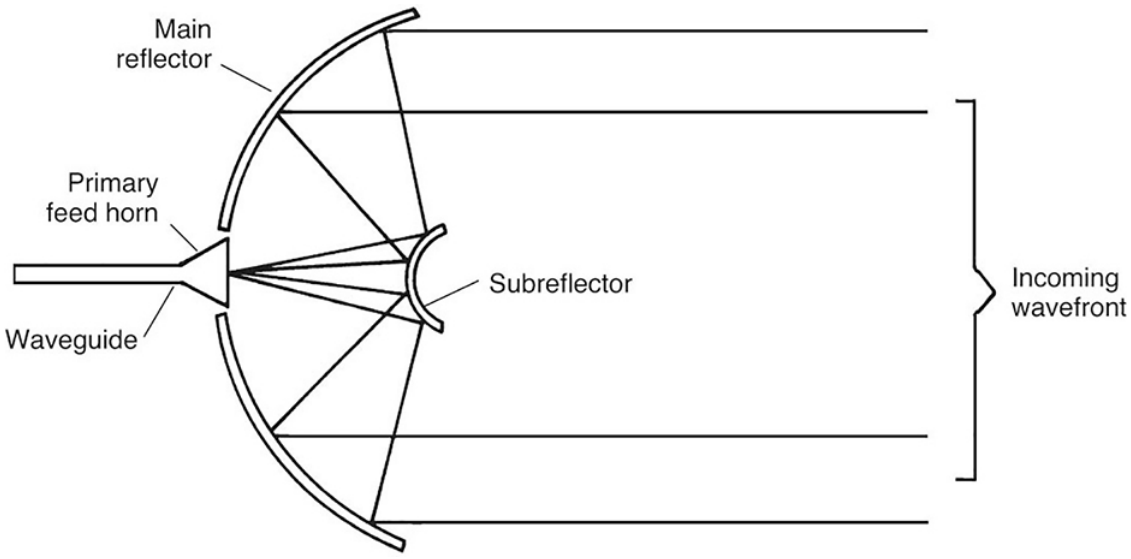
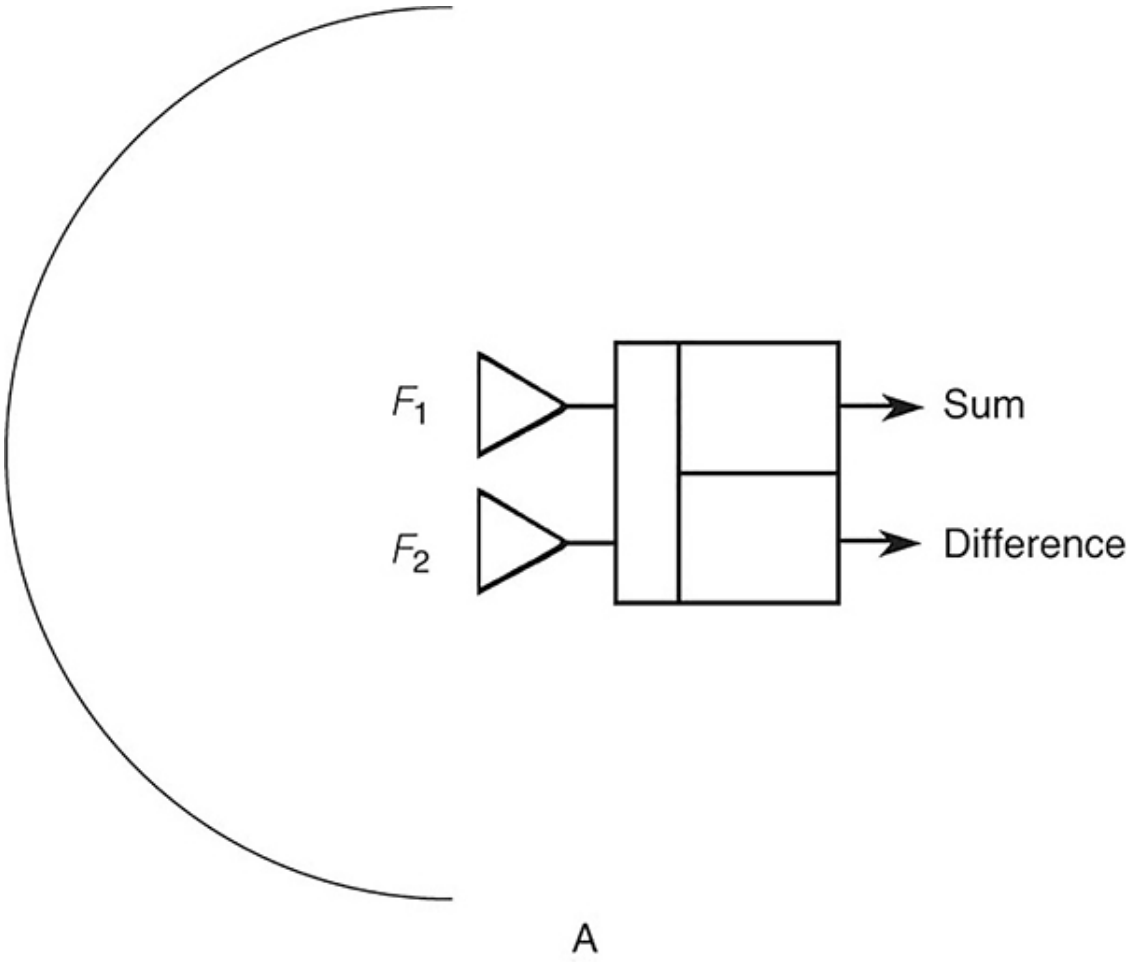


FIGURE 23.8 Parabolic antenna Cassegrain feed.

Figure 23.8 shows the Cassegrain feed system modeled after the Cassegrain optical telescope. The radiator element is placed at an opening at the center of the dish. A hyperbolic subreflector is placed at the focal point and reflects the wavefront to the radiator element. The Cassegrain system results in lower-noise operation for several reasons. The transmission line length to the feed is shorter, sidelobes are generally lower, and the horn sees the sky instead of earth, reducing spillover noise from the warm ground below the dish. On the negative side, galactic and solar noise might be slightly higher on a Cassegrain dish, depending on where it's aimed.

Figure 23.9 shows the monopulse feed geometry. Two radiator elements (or four for a two-dimensional tracking system) placed at the focal point are fed to a power splitter network that produces both sum and difference signals. When these are combined, the resultant beam shape (Fig. 23.10) has an improved 3-dB beamwidth as a result of the algebraic summation of the two. The difference signal can be used in a closed-loop tracking control system that maintains zero response to a signal source in the difference beam and keeps the dish pointed at the desired source.



F IGURE 23.9 Monopulse feed.

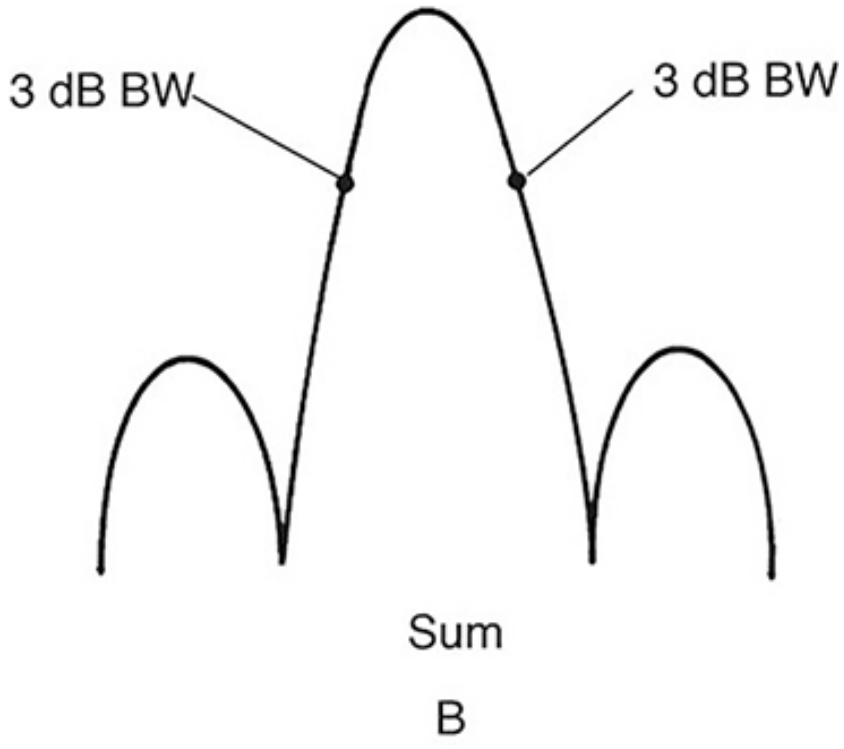


FIGURE 23.10 Monopulse patterns.

23.5 Array Antennas

Microwave arrays can be formed from dipole elements, as in the broadside array of Fig. 23.15 (which is used in the UHF region), or a series of waveguide slots, horns, patch antennas on PCB substrate, or other radiators.

The gain of an array antenna is roughly equal to the number of elements multiplied by the gain of one element. The precise gain depends on the geometry of the array, particularly the spacing between elements, as well as mutual coupling between the elements, losses in the array feed network, and any objects around the array that perturb the radiated fields.

Some simple types of phased array antennas have fixed beams. The phases of the signals at each array element are controlled by transmission lines such as microstrip with different lengths so that signals arrive at the elements with phases that produce a beam in a

desired direction. Since the transmission lines have fixed lengths, the beam cannot be steered.

More complex antenna arrays include electronic control of the signal amplitudes, phases, or time delays. This allows for beam scanning and tracking of moving sources such as aircraft or satellites. More sophisticated types of control of the array pattern are also possible, including formation of simultaneous beams to track multiple targets, side-lobe control, and dynamic nulling of interfering sources.

23.5.1 Slot Array Antennas

Open ended or flared waveguide antennas are common, but these are single-aperture type antennas with a fixed beam pattern. Rather than having an open end, holes or slots can be cut into the wall of a waveguide section at locations and spacings designed to achieve a controllable beam pattern. Slot arrays tend to be highly efficient and are used for marine navigation radars, telemetry systems, and the reception of microwave television signals in the Multipoint Distribution Service (MDS) on 2.145 GHz.

A slot in a waveguide wall radiates in a way that is similar to a dipole. This can be used to analyze or design the locations of slots to achieve a particular goal for the beam pattern.

Figure 23.11 shows a simple slot antenna used in telemetry applications. A slotted section of rectangular waveguide is mounted to a right-angle waveguide flange. An internal wedge (not shown) is placed at the top of the waveguide and serves as a matching-impedance termination to prevent internal reflected waves. Directivity is enhanced by attaching flanges to the slotted section of waveguide parallel to the direction of propagation (see end view of Fig. 23.11).

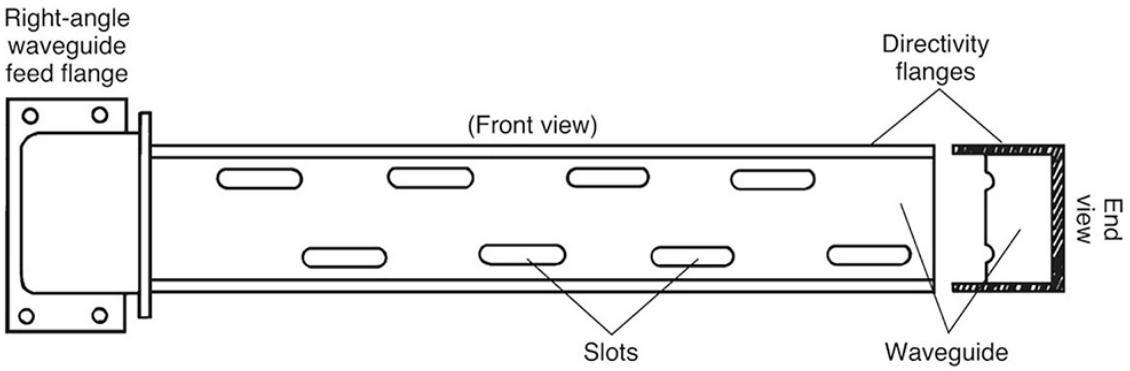


FIGURE 23.11 Slot antenna (front view).

Figures 23.12 and 23.13 show two forms of flatplate array antennas constructed from slotted waveguide radiator elements (shown as insets). Figure 20.34A depicts a rectangular array and Fig. 23.13 shows a circular array. These flatplate arrays are used extensively in microwave communications and radar applications.

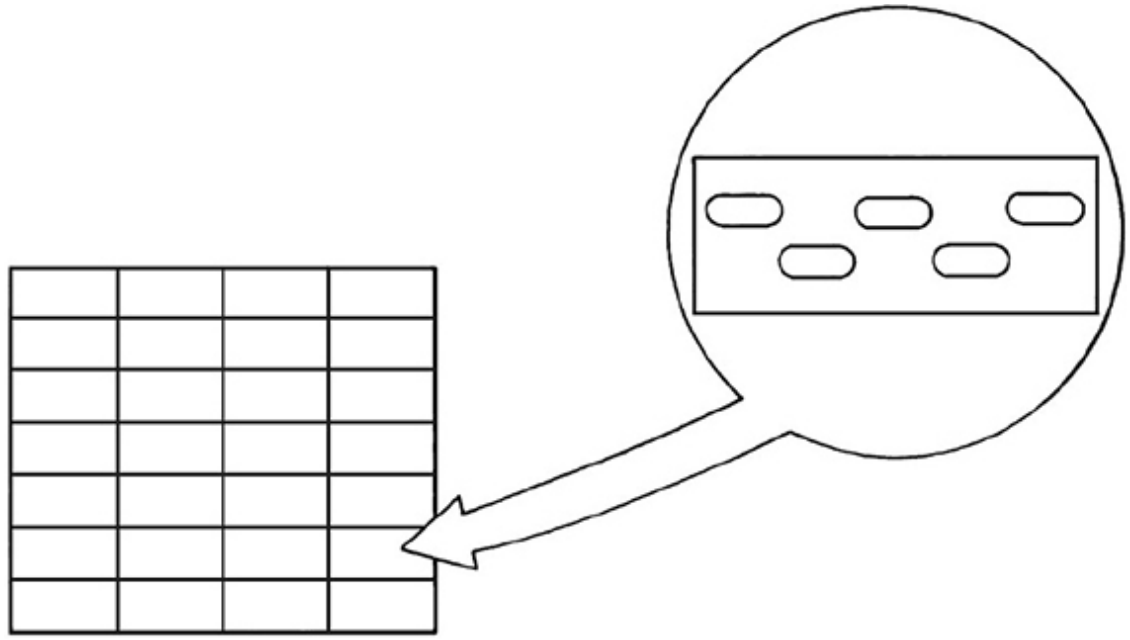


FIGURE 23.12 Flatplate slot array.

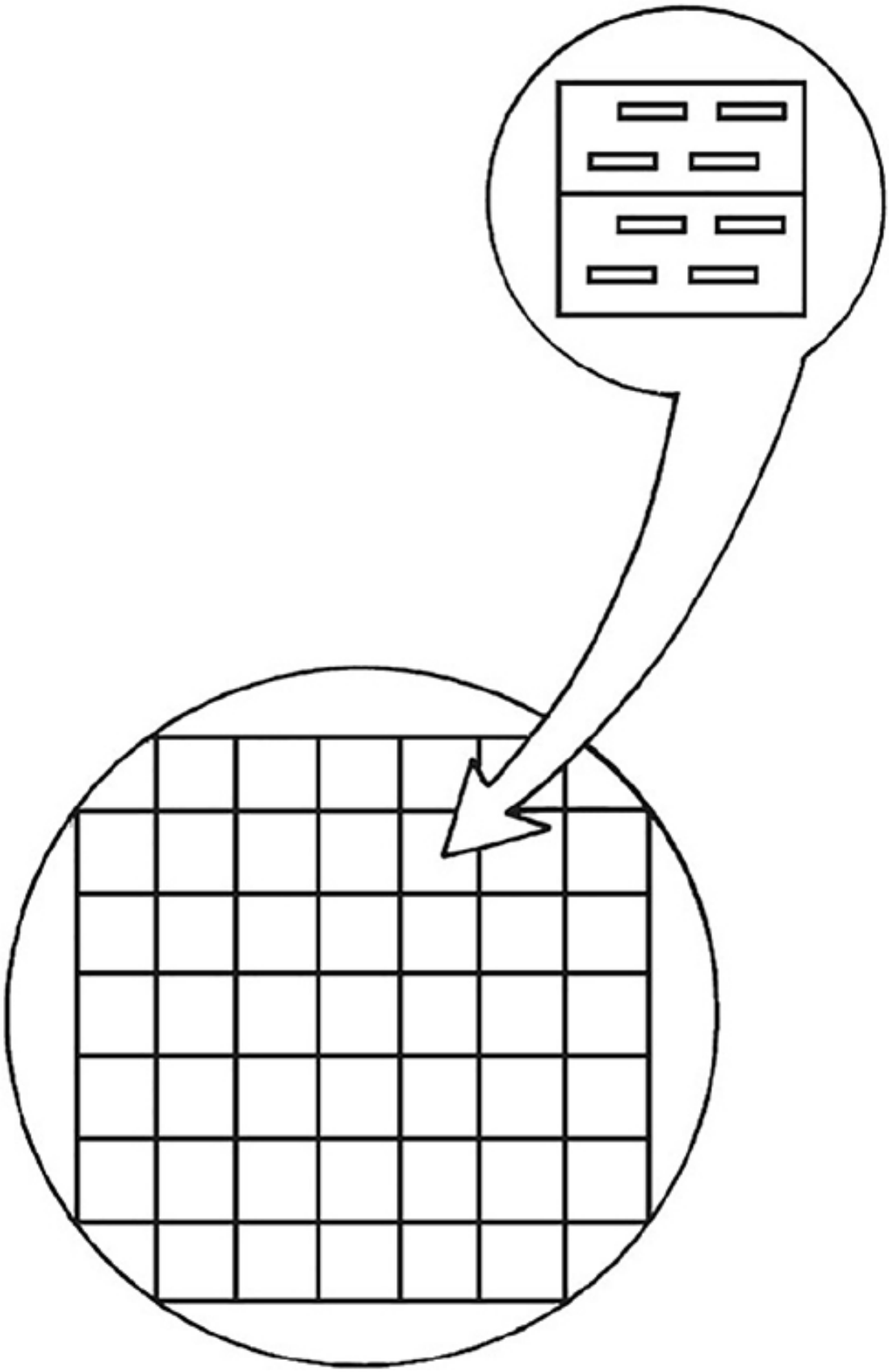


FIGURE 23.13 Flatplate antenna.

The feed structure for a flatplate array is shown in Fig. 23.14. A distribution waveguide is physically mated with the element, and a coupling slot is provided between the two waveguides so that energy propagating in the distribution system waveguide can couple into the antenna radiator element. In some cases, metallic or dielectric phase-shifting stubs are also used to fine-tune the antenna radiation pattern.

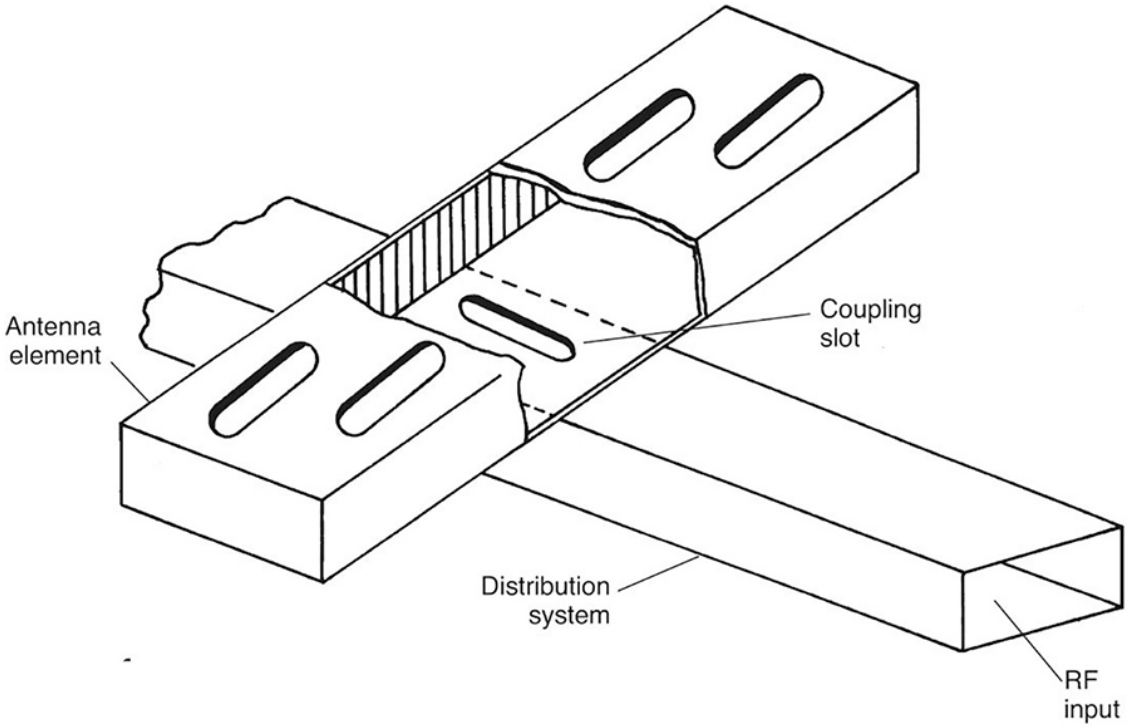


FIGURE 23.14 Flatplate antenna feed coupling.

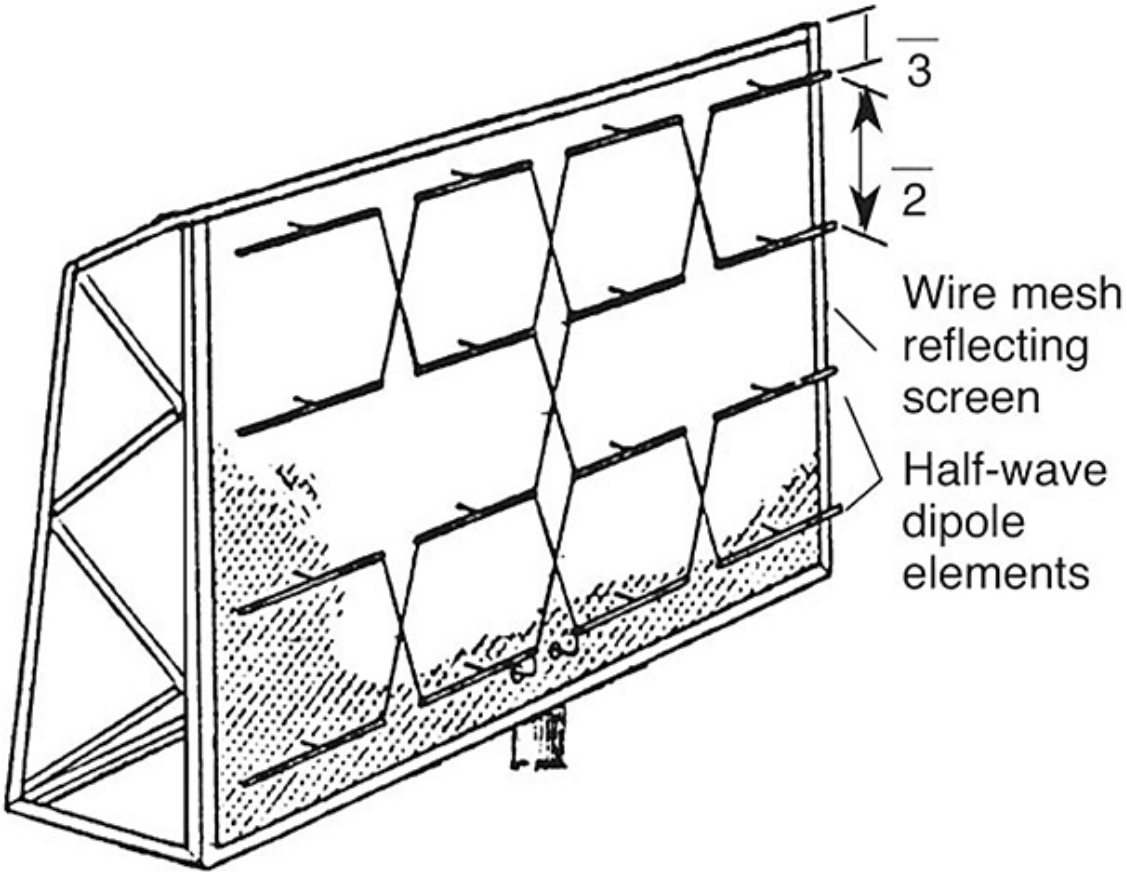


FIGURE 23.15 Reflector array antenna.

23.5.2 Electronically Steered Arrays

For dynamic control of the beam pattern, electronically steered arrays (ESAs) are used. Applications include satellite communication terminals, phased array radars, multiple antenna transceivers with controllable beams for wireless networks and cellular systems, imaging arrays for radio astronomy, and automotive collision avoidance sensors.

To electronically steer the main beam, array antennas require a method of phase shifting. Electronically controlled phase shifters are expensive compared to other components in an antenna system, so electronically steered arrays in the past have been used mainly for military applications. In recent years, phase shifters have been integrated on silicon chips, which lowers the cost of phase control, and consumer electronically steered phased arrays for

automotive applications, cellular systems, and satellite communications have become cost effective.

Two methods of feeding an array are shown in Fig. 23.16. The corporate feed method connects all elements and their phase shifters in parallel with the source. The branch feed method breaks the waveguide network into separate paths.

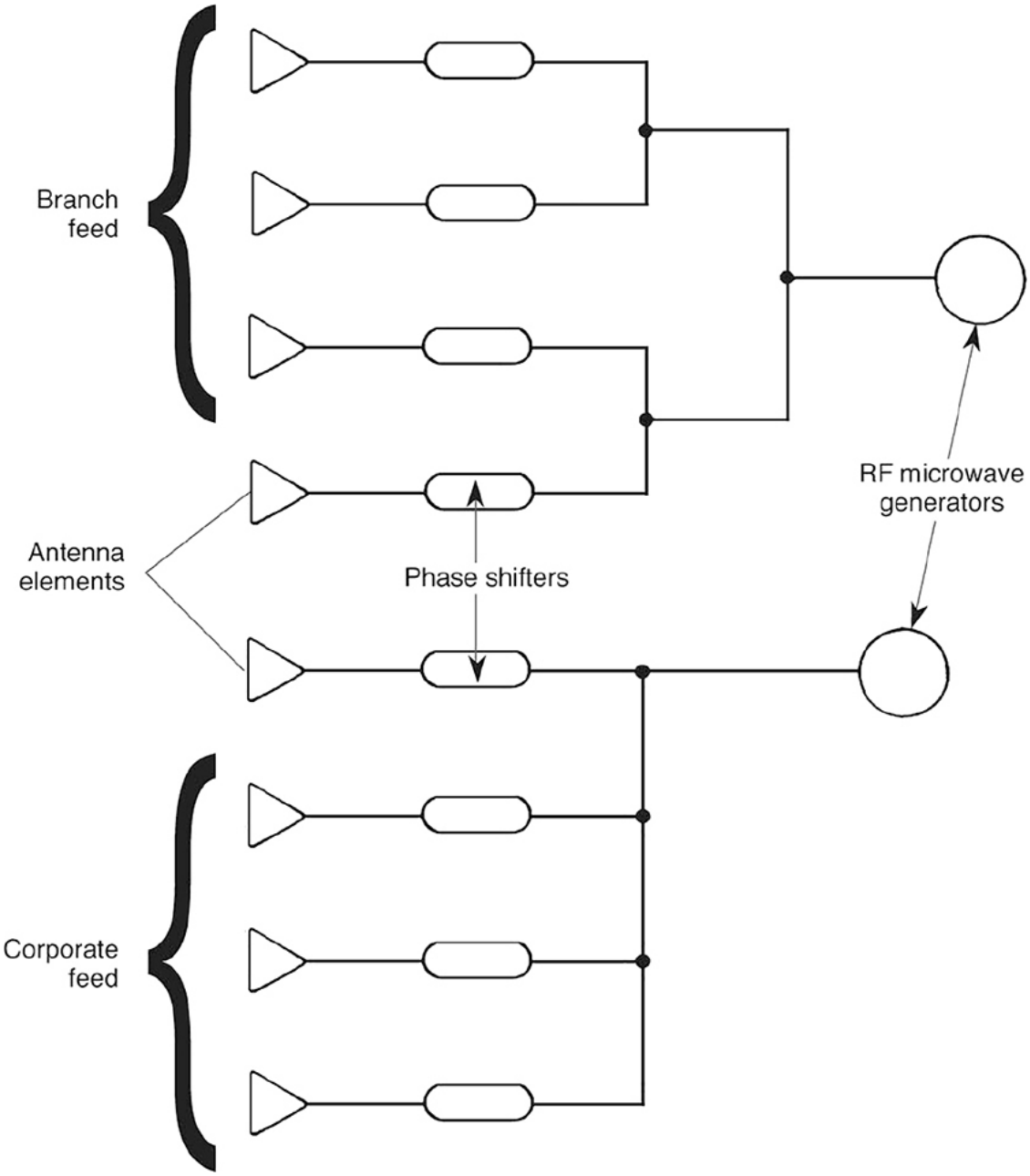


FIGURE 23.16 Branch feed and corporate feed.

Phase shifters come in various types. Figure 23.17 shows a passive phase shifter that is controllable but has no external energy source. Some implementations use a ferrite or ferroelectric material or a PIN diode placed in the transmission path between the RF input and the antenna element. By changing the phase of the RF signal selectively in each feed line leading to the elements in an antenna array, it is possible to form and steer the beam on transmit or on receive. A 3-bit phase shifter is capable of eight distinct states and thus allows the phase to shift in 45-degree increments, while a 4-bit phase shifter with 16 states allows 22.5-degree increments of phase shift. In general, the larger the array in terms of the number of elements, the lower the number of states required for the phase shifters to steer high-quality beams in reasonably small increments.

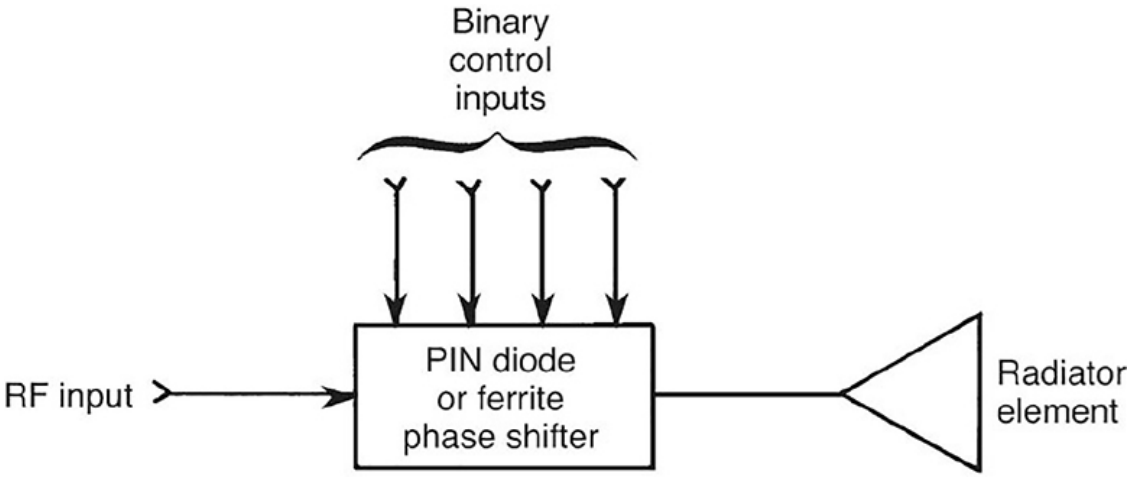


FIGURE 23.17 Phase shifter.

The active module of Fig. 23.18 contains a transmit power amplifier (typically 1 or 2 W) and a low-noise amplifier (LNA) for receiving. A transmit/receive (T/R) switch isolates the LNA input from the high-power RF emitted by the power amplifier, while a second T/R switch selects whether the phase shifter is in line with the transmitted or the received RF. The output power of the antenna is the sum of all output powers from all elements in the array, so that an array of a thousand 2W elements forms a 2000 W system.

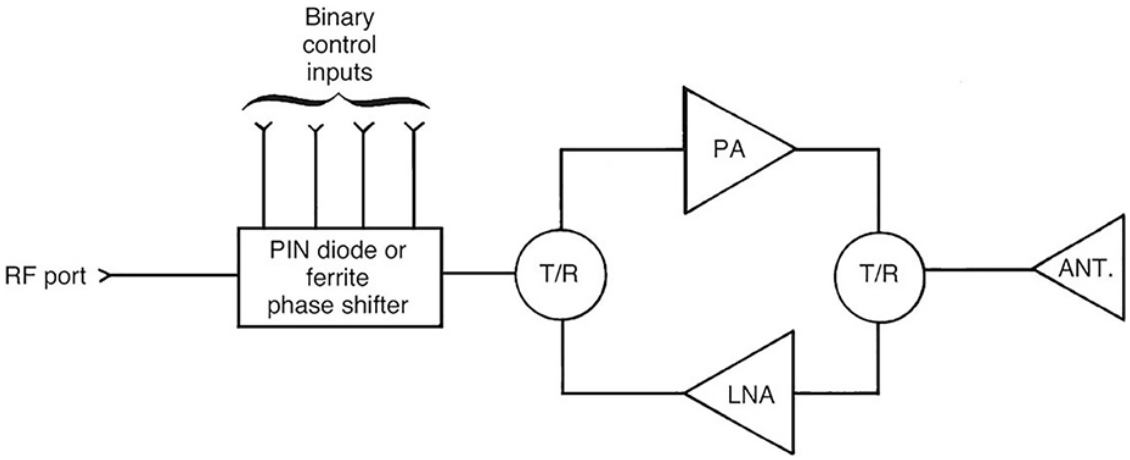


FIGURE 23.18 Phase-shift T/R element.

In some applications, digitally beamformed arrays are used. This is typically a higher cost, larger, and more complex solution, but offers greater flexibility in controlling the beam pattern of the array than analog beamforming. Signals from array elements are sampled and time delays or phase shifts are applied in digital signal processing (DSP). Digitally beamformed arrays can form more than one transmit beam, dividing power among different main lobes, and on receive can distinguish multiple signals with different arrival angles.

Mobile wireless devices often use multiple input multiple output (MIMO) arrays to improve performance in environments with strong reflections from objects leading to multiple propagation paths from the base station to device, or multipath. With a single antenna, multipath can cause destructive interference and fades, leading to bit errors and poor quality of service. A MIMO system transmits a training sequence of known signals and the receiver measures the state of the propagation channel. Beams associated with different multipaths are encoded with parallel information streams that are received and aggregated by the receiver in such a way that the data rate is higher than with a conventional single antenna transceiver.

CHAPTER 24
Antenna Noise

The most important figure of merit in radio communication is the signal-to-noise ratio (SNR). Signals must be at or above some amplitude relative to the noise floor of the system in order to be detected properly. Noise is emitted by warm objects, sky sources, interfering sources, antennas, and electronics. To maximize SNR, we need to receive as high a signal level as possible and minimize the noise added by the antenna and receiver electronics.

At HF and below, the external noise due to galactic radiation and human-made systems is extremely high. The noise from external sources is larger than the noise added by most common receivers, and no special design attention to front-end amplifiers or minimizing antenna losses is needed. Antenna efficiency is unimportant, as a more efficient antenna receives more external noise as well as signals of interest. At low frequencies, SNR does not increase much with the lower losses in a better designed antenna or the decreased noise added by a higher quality receiver system.

At the VHF band and microwave bands, levels of noise from the external environment are much lower. Near L band, the microwave sky is so cool that its equivalent temperature is nearly absolute zero in degrees Kelvin. At these bands, taking special measures to improve antenna efficiency and receiver quality can significantly increase the SNR. The chapter is primarily for the benefit of those who are concerned with weak signal reception above 30 MHz or so.

In electronic systems, there are various forms of noise. These include shot noise, flicker noise, quantization noise, rounding error in digital signal processing, nonlinearity due to out-of-band signals above the receiver dynamic range, in-band noise generated by the

transmitter of interest, interference from other transmitters, and thermal noise. Of these noise sources, at microwave frequencies interference and thermal noise are typically dominant. We are concerned here mainly with minimizing thermal noise.

24.1 Thermal Noise

Thermal noise, also called Johnson noise, is produced by the random motion of electrons in warm objects around an antenna, in the antenna itself, and in any lossy components or transmission lines between the antenna and active electronics in the receiver. At all temperatures above absolute zero (about $-273.16^{\circ}\mathrm{C}$), the electrons in any material are in constant motion, leading to detectable noise voltages at the terminals of any lossy device or component.

A lossless object produces no thermal noise, so thermal noise is associated with losses in a system. A metallic reflector antenna, for example, produces so little thermal noise that we neglect it when analyzing the noise budget of a parabolic satellite dish receiver.

Thermal noise in a resistor or other lossy object can be modeled as a voltage source, $\overline{V}_{\mathrm{N}}$ in series with an ideal, noise-free resistor with value *R*, where $\overline{V}_{\mathrm{N}}$ is the root mean square (rms) value or standard deviation of the nondeterministic, thermally generated noise voltage. When the resistor is connected across a matched load, the noise power transferred to the load is

$$P_N = k_B T B \quad (24.1)$$

where P_N is the noise power in Watts, k_B

is Boltzmann's constant (
$1.38 \times 10^{-23}\ \text{joules/K}$
), *T* is the temperature of the resistor or lossy material in degrees Kelvin (K), and *B* is the receiver bandwidth in Hertz.

Degrees Kelvin (K) is the international way of defining temperatures relative to absolute zero. To express temperature in degrees Kelvin (K) we add 273.16 to a temperature expressed in Celsius. The formula is
$T(K) = T(^\circ C) + 273.16$
. For Fahrenheit temperatures, the conversion is
$T(K) = (5/9)T(^\circ \text{F}) + 255.38$
. Absolute zero corresponds to
-459.67°F
,
-273.16°C
, or 0 K.

An important point on the scale is room temperature, typically taken to be 290 K. By convention, terrestrial components and system elements are assumed to be the reference temperature
$T_0 = 290 \text{K}$
unless otherwise stated. A commonly used reference value is

$$10\ \log_{10}(1000\ k_B T_0) = -174 \text{dBm/Hz} \quad (24.2)$$

Starting with this value, one can add the receiver bandwidth converted to dB Hz, the receiver noise figure, and the receiver gain in dBm to quickly estimate the noise power dB relative to 1 mW (dBm) at the receiver output.

Example 24.1 A terrestrial antenna is connected to a receiver with a 1 MHz bandwidth. The noise power delivered by the antenna to the receiver input stage is
$\left(1.38 \times 10^{-23}\ \text{J/K}\right) \times \left(290\ \text{K}\right) \times \left(1,000,000\ \text{Hz}\right) = 4 \times 10^{-15}\ \text{W}$
or
$-174\ \text{dBm/Hz}\ + 60\ \text{dBHz}\ = -114\ \text{dBm}$

.

To improve weak-signal detection at microwave frequencies and above, receiver input stages can be cooled by liquid nitrogen or helium. The lower physical temperature of the amplifier stages results in their contributing a substantially reduced noise power to the overall system noise figure and a higher SNR. Cryogenic receivers are used for high-sensitivity applications like deep space communications, remote sensing, and radio astronomy.

24.2 Noise Figure and Equivalent Noise Temperature

The noise performance of a receiving system can be defined using the noise figure *F* or equivalent noise temperature
T_e
.

24.2.1 Noise Figure (*F*)

For components such as resistors, the noise figure is the ratio of the noise produced by a real resistor to the simple thermal noise of an ideal resistor. The noise figure of a radio receiver or any other RF or microwave system is the ratio of output noise power to input noise power:

$$F = \left. \frac{P_{\mathrm{n,out}}}{P_{\mathrm{n,in}}} \right|_{T_{\mathrm{in}} = 290\,\mathrm{K}} \tag{24.3}$$

By convention, the input noise when computing the noise figure is always
$k_B T_0 B$
, where
$T_0 = 290\ \mathrm{K}$
is the standard reference temperature of
$T_0 = 290\ \mathrm{K}$

. The noise figure can be expressed in terms of input and output signal-to-noise ratios:

$$F = \frac{\mathrm{SNR}_{\mathrm{in}}}{\mathrm{SNR}_{\mathrm{out}}} \quad (24.4)$$

where again the input SNR is computed or measured with the noise floor
$k_B T_0 B$
. The noise figure in natural numbers as defined in (24.3) is sometimes called the noise factor. The noise figure can also be expressed in decibel notation:

$$\mathrm{NF} = 10 \ \log_{10} F \quad (24.5)$$

24.2.2 Noise Temperature (T_{e})

The noise performance of a receiver can also be specified in term of equivalent noise temperature or more concisely noise temperature. The noise temperature of a component, device, or system can be different from its physical temperature. The noise temperature is the temperature
T_e
in Kelvin of a matched resistor at the input to a device that produces the same noise power as is added by the device.

Noise temperature is related to the noise figure by

$$T_e = (F - 1)T_0 \quad (24.6)$$

Noise temperature and noise figure are equivalent and one can be found from the other using (24.6). Noise figure is commonly used in microwave systems and terrestrial communications, whereas noise temperature is common in satellite communication systems.

The system noise is a combination of external noise from warm objects or sky around the antenna, noise due to losses in the antenna, and receiver noise due to electronics after the antenna. The external noise depends on the antenna and the environment around the antenna. It may consist of a sky noise temperature that depends on where the antenna main lobe is pointed and a ground noise temperature that consists of components reflected from the sky as well as noise from the warm ground below the antenna.

In a typical system (Fig. 24.1), the main lobe will be pointed toward the sky noise source, while the sidelobes pick up noise from the ground. The equivalent noise temperature of the antenna is

$$T_a = \eta_{\text{rad}}\eta_{\text{sp}}T_{\text{sky}} + \eta_{\text{rad}}(1-\eta_{\text{sp}})T_g + \underbrace{(1-\eta_{\text{rad}})T_p}_{T_{\text{loss}}} \quad (24.7)$$

where

η_{rad}
is the antenna radiation efficiency,

η_{sp}
is the spillover efficiency or the fraction of power from the antenna that radiated to the sky divided by the total power,

T_{sky}
is the sky noise temperature,

T_g
is the brightness temperature of the ground around the antenna,

T_{loss}
is the equivalent temperature of noise due to losses in the antenna, and

T_p
is the antenna physical temperature. The antenna noise temperature can also be found by integrating the product of external temperature as a function of angle by the average antenna gain pattern over a full sphere of solid angles, so that

$$T_a = \frac{1}{4\pi} \int G(\theta, \phi) T(\theta, \phi) \sin\theta d\theta d\phi \quad (24.8)$$

where $T(\theta, \phi)$ is the brightness temperature of the environment around the antenna at the spherical angles θ, ϕ relative to the location of the antenna.

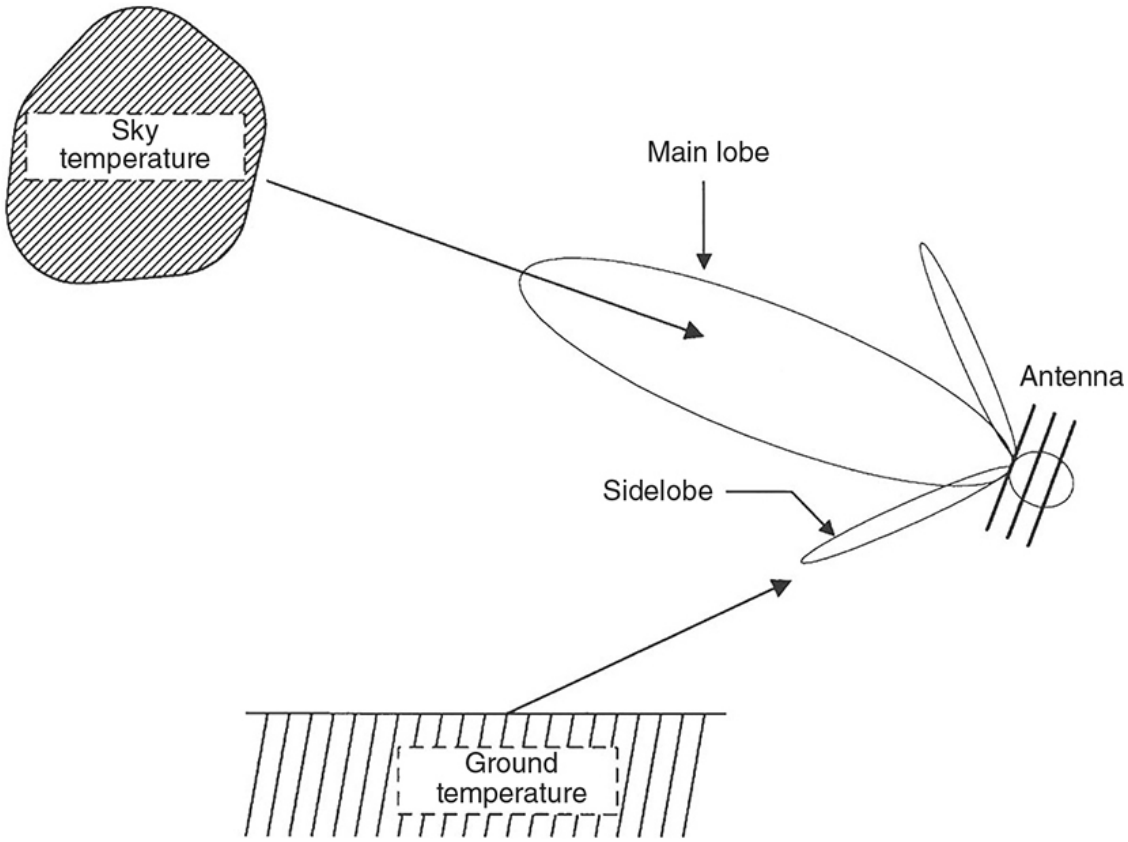


FIGURE 24.1 Contributors to antenna noise temperature.

The system noise temperature is a combination of the antenna noise and receiver noise due to electronics after the antenna:

$$T_{\text{sys}} = T_a + T_{\text{rec}} \quad (24.9)$$

For a satellite dish receiver operating in the 1 to 10 GHz range, the sky noise might be 5 K at the L band or 10 to 15 K at the K band in dry sky conditions, and the receiver noise is around 50 to 60 K for a system noise temperature of 55 to 75 K. For a terrestrial

communication system or handheld wireless device, the external noise is somewhat higher, and the requirements for the antenna radiation efficiency and receiver noise figure can be relaxed, so the system noise temperature may be around 400 K range. Below around 30 MHz, the sky noise is far higher, and the noise temperature may be
$10^5\ \mathrm{K}$
or larger.

CHAPTER 25
Radio Astronomy Antennas

For centuries, astronomers have scanned the heavens with optical telescopes. The sky can also be studied using radio waves emitted by deep space objects. The field of radio astronomy emerged in the 1930s and 1940s through the work of Grote Reber and Carl Jansky who first observed radio signals from the Milky Way.

The main issue for an astronomical antenna is achieving sufficient effective area to detect extremely weak signals from deep space. This can be done with large dish antennas like the 100-m Green Bank Telescope (GBT) in West Virginia, the decommissioned 300-m Arecibo telescope, or the Five-hundred-meter Aperture Spherical Telescope (FAST) in China. With a traditional horn feed, a dish antenna is essentially a wideband, single-pixel radio camera. To expand the field of view, dish antennas can be outfitted with cluster feeds or phased array feeds that can digitally synthesize many simultaneous beams and image a larger patch of sky.

As large dishes are expensive, interferometric arrays of medium-sized dishes like the Very Large Array (VLA) in New Mexico can be used to achieve a large collective aperture area. The advantage of interferometric arrays is that resolution is inversely related to the largest dimension of the array, so they can provide higher resolution than single-dish instruments.

High resolution images of distant sources of current scientific interest require large dishes or long baseline arrays that are beyond the capabilities of the typical hobbyist, but there are astronomical sources that can be detected using more practical setups. For amateur radiation astronomy, Yagis, ring Yagis, cubical quads, and other antennas can be used at lower frequencies (18–1200 MHz). Reflectors of reasonable size can be used for higher frequencies.

Amateur radio astronomers have employed TV receive-only (TVRO) satellite dishes other readily available small or mid-size reflector antennas. In this chapter, we limit our coverage mainly to antennas that are not discussed in other chapters.

25.1 Signals from the Sun

The strongest readily detectable extraterrestrial radio source is the sun. The sun's radio emissions are broadband and receivers typically detect only a small portion of the total radiated spectrum at any one time. When we listen on a receiver, what we hear is noise that varies in amplitude slowly or in bursts.

As a rule, we are restricted to monitoring the sun and other astronomical bodies at frequencies higher than those blocked by earth's ionosphere. The low end of the useable frequency range varies with the sunspot cycle, time of day, etc., but frequencies above 18 MHz should nearly always work. Many amateur radio astronomers concentrate on the region around 20.1 MHz, but 28 MHz also works with just a three-element Yagi. If these bands are open for HF skip propagation, or if many local ground-wave signals are present, interfering terrestrial signals may be too strong and other bands may be better for solar observations.

Needed equipment is a receiver that works well over the range from 18 to 30 MHz and a relatively simple antenna. Automatic gain control (AGC) or automatic volume control (AVC) should be turned off, the audio volume control set to a comfortable level, and the RF gain high enough that external noise overrides the internal noise of the receiver. Set the receiver to the single sideband (SSB) mode and select the widest filter bandwidth available. Observatories also make available spectrograms and other displays of received solar radio signals online.

Orient the antenna with the peak response of its pattern in the direction of the source. Since the sun's path for many of us covers a broad range of azimuth and elevation angles, aim the antenna at the sun's location during parts of the day when observations are

made. For dipoles at 18 MHz and above, rotating them is a relatively simple task, whether done manually or with an antenna rotator.

To improve your ability to pull extraterrestrial signals out of your background noise environment, a higher gain antenna may be used, especially if you are in an urban neighborhood. Rotating a two- or three-element Yagi at 20 MHz and higher can be done. Antenna height is not a factor, particularly during listening periods when the sun is well above the horizon.

25.2 Signals from Jupiter

Second only to the sun, Jupiter is a strong radio source. It produces noiselike signals from VLF to 40 MHz, with peaks between 18 and 24 MHz. The radio signals are likely produced by charges moving through the planet's magnetic field, and are influenced by the transit of the Jovian moons. The signals are plainly audible on the upper HF bands any time Jupiter is above the horizon, day or night. To eliminate the possibility of terrestrial skip signals from interfering, Jupiter DXers prefer to listen when the ionospheric maximum useable frequency (MUF) drops significantly below 18 MHz, which is normally the darkness hours. In preparation, listen to the amateur 17- or 15-m bands; if you hear no skip-distance activity, then the MUF has likely dropped enough to make listening worthwhile. During the day, it is possible to hear Jovian signals, but differentiating them from other signals and solar noise can be difficult.

Jupiter emits two distinctly different types of radio noises. Listening in SSB mode, one form can be heard as "swooshing" noises that rise and fall in amplitude over a relatively long interval. The second type of noise from the planet is heard as a more rapid-fire "popping" sound.

Unlike the sun, Jupiter does not often rise high in our sky. Its transit, as viewed at mid-latitudes in the northern hemisphere, is typically close to the horizon, rising in the southeast and setting in

the southwest. Jupiter is an excellent target for a typical HF Yagi or cubical quad whose height has been optimized for low elevation angles. Jupiter is visible (and audible) at times totally unrelated to our day and night periods. There may be periods when Jupiter's emissions are difficult to hear because the sun is also radiating.

For monitoring Jupiter, a good beginning antenna can be a simple dipole cut for the middle of the 18- to 24-MHz band, which happens to coincide with the 15-m amateur radio band. The antenna should be installed in the normal manner for any dipole, except that if it is fixed in one position the wire should run east-west in order to maximize pickup from this southerly rising planet.

Figure 25.1 shows a broadband dipole that covers the entire frequency span of interest (18–24 MHz) by paralleling three different dipoles: one cut for 18 MHz, one cut for 21 MHz, and one cut for 24 MHz. The dimensions are

$$A = 19.5 \text{ ft}$$

for 24 MHz,

$$B = 22.3 \text{ ft}$$

for 21 MHz, and

$$C = 26 \text{ ft}$$

for 18 MHz.

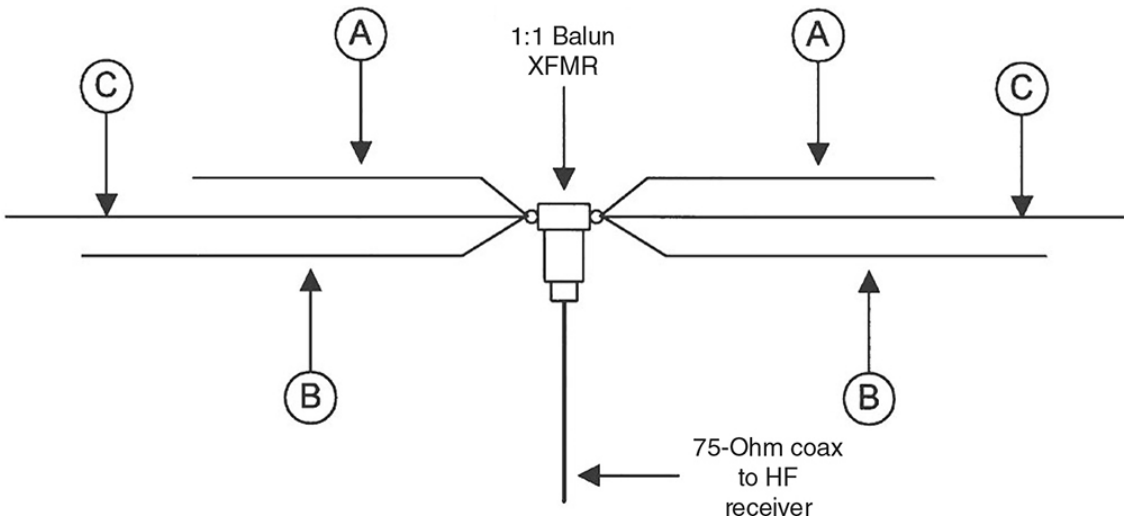


FIGURE 25.1 Wideband HF dipole for Jupiter reception.

As discussed in the chapter on multiband wires (Chap. 10), there are several approaches to making this type of antenna. One is to use three-conductor wire and cut the wires to the lengths indicated here. Another is to use a homemade spacer to spread the wires apart.

25.2.1 Ring Antenna

Another antenna that can be used for receiving signals from Jupiter is the ring radiator. Two versions are shown in Figs. 25.2 and 25.3. This antenna is made of a 5-ft diameter loop of 1/2-in diameter copper plumbing pipe. The single-ended version is shown in Fig. 25.2. The far end of the loop is unterminated (i.e., open-circuited). The center conductor of the coaxial cable feedline is connected to the near end of the ring radiator, while the coax shield is connected to a chicken wire ground plane. The balanced version (Fig. 25.3) has an RF transformer (T_1) at the feedpoint.

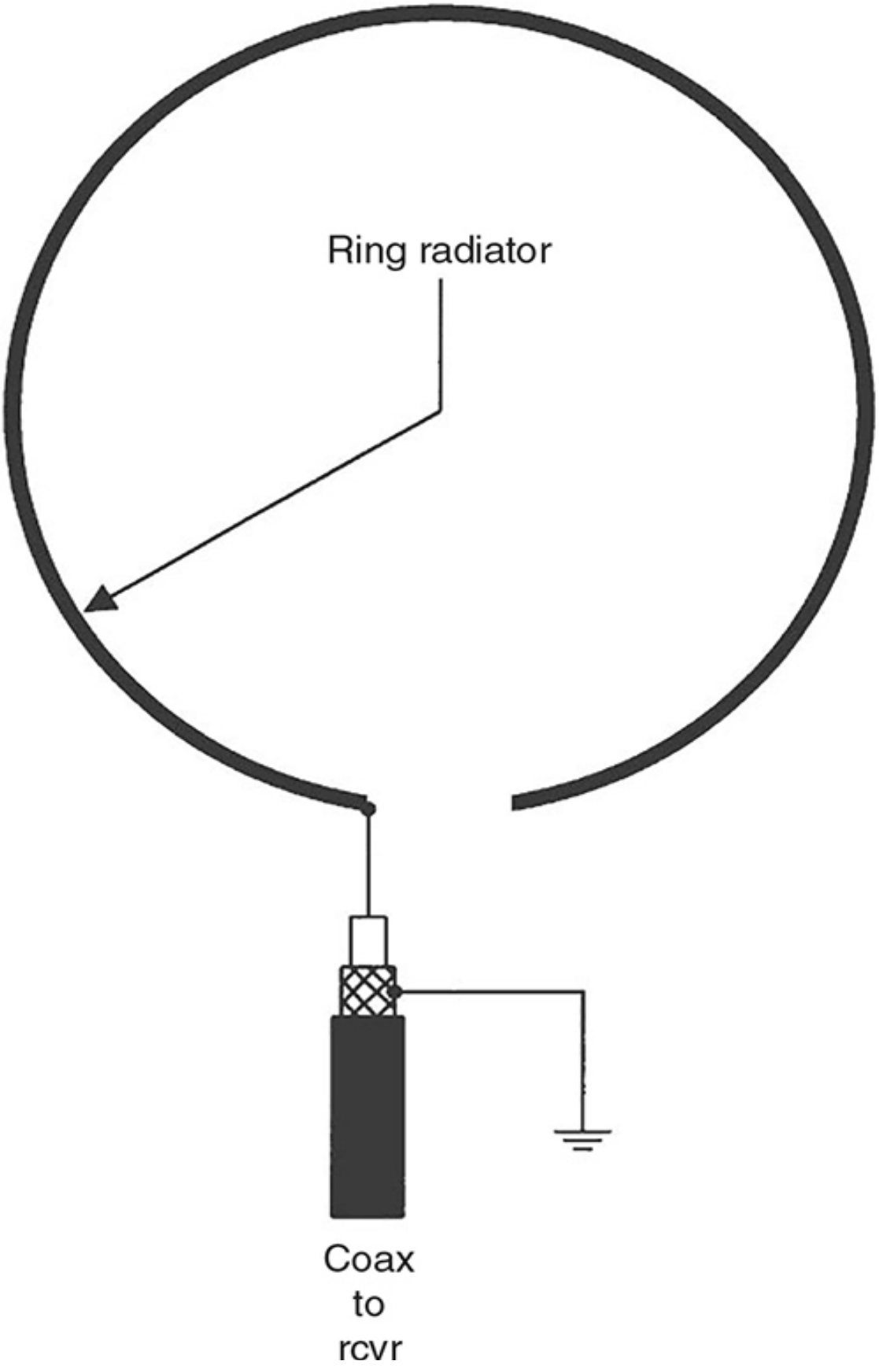

Ring radiator
Coax
to
rcvr

FIGURE 25.2 Ring radiator (single-ended).

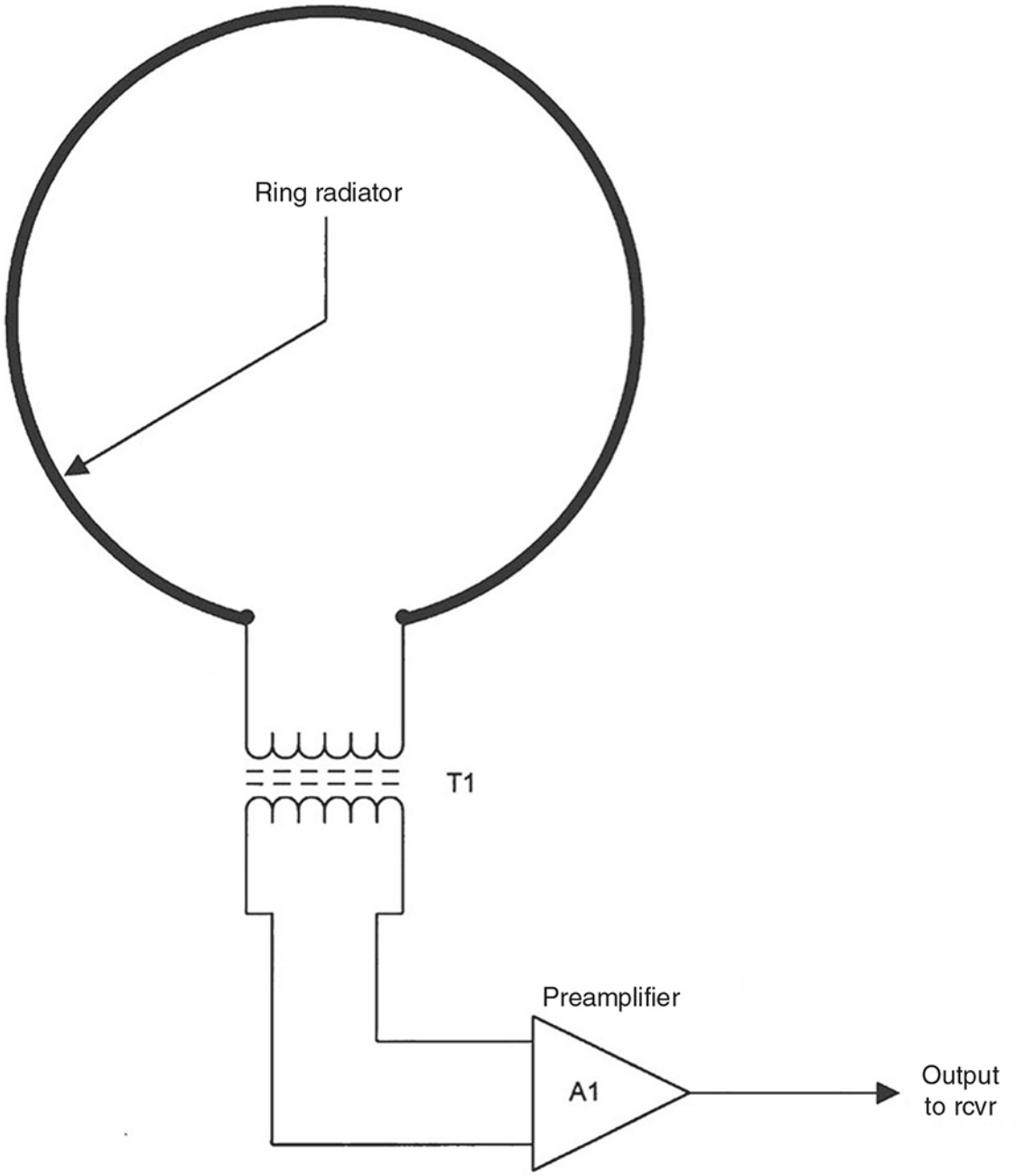
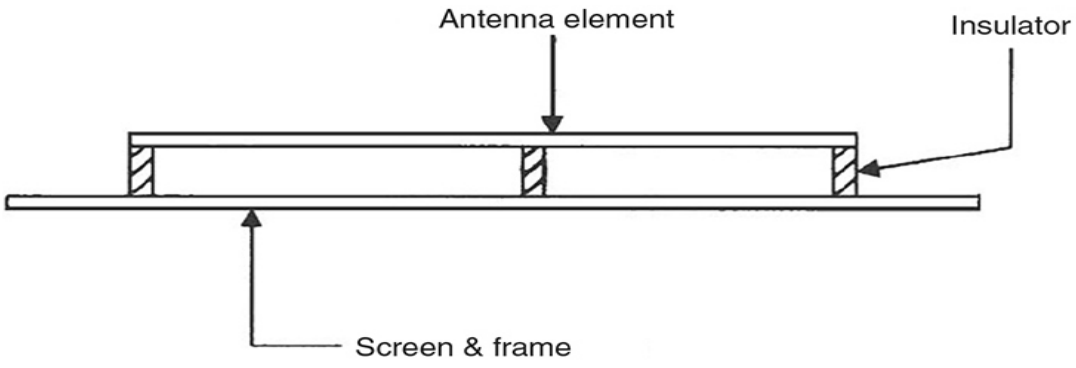


FIGURE 25.3 Ring radiator (balanced).

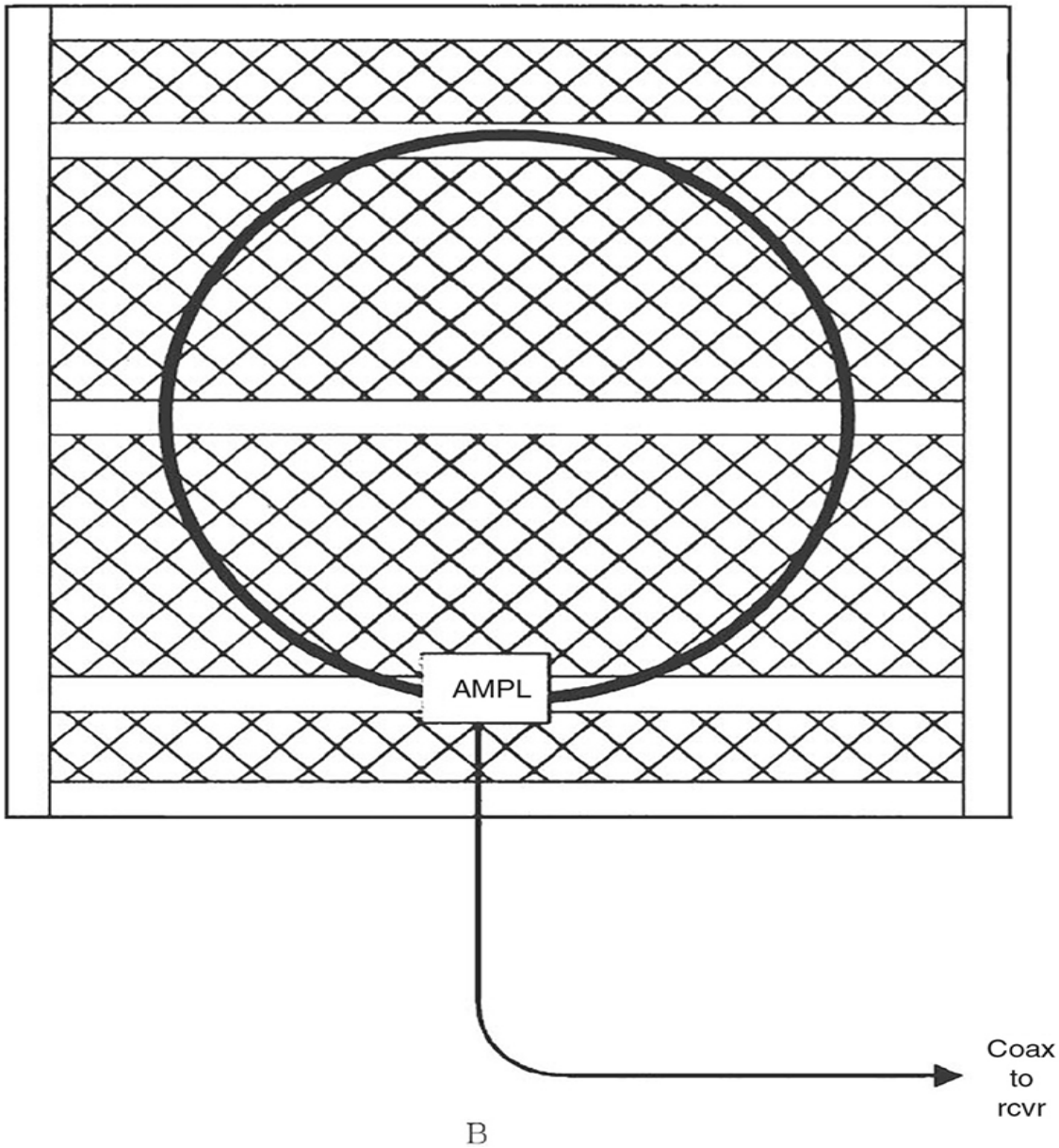
The ring radiator antenna should have a bandpass-filtered preamplifier as close as possible to the antenna. The filter minimizes the likelihood of strong terrestrial signals in the adjacent

bands interfering with the preamp or receiver. In the case of the single-ended amplifier, a single-ended preamplifier is used. But for the balanced version (Fig. 25.3), a differential preamplifier is appropriate. Either version of the loop or ring is mounted about 7 or 8 in above a ground plane made of chicken wire, metal window screen, copper sheeting, or copper foil.

Figure 25.4A is a mechanical side view of the ring radiator antenna, while Fig. 25.4B is a top view. The antenna is mounted above the screen with insulators. These can be made of wood, plastic, or any other insulating material. The frame is strengthened by interior crosspieces that also support the antenna. The larger outer perimeter is needed because the screen ground plane should extend beyond the diameter of the radiator element by at least 10 to 15 percent of that dimension.



A



B

FIGURE 25.4 Ring radiator mounted over a ground screen.

25.2.2 Directional Discontinuity Ring Radiator

The directional discontinuity ring radiator (DDRR) antenna is shown in Fig. 25.5. A side view showing the mounting scheme is shown in Fig. 25.6. The DDRR is typically mounted about 1 ft off the ground (
$H = 12 \text{ in}$
). It consists of two sections, one vertical and one horizontal. The short vertical section has a length equal to the height *H* of the antenna above the ground and the lower end of the vertical segment is grounded. The horizontal section is an open-circuited loop of diameter *D*, which varies with frequency. The optimum conductor diameter E for the DDRR is at least 0.5 in at 28 MHz, increasing to 4 in at 4 MHz. For monitoring Jupiter, a diameter of 1 in or slightly less is appropriate. Construction details of the DDRR are similar to those of the ring radiator and the same diagram can be used (see Fig. 25.4 again).

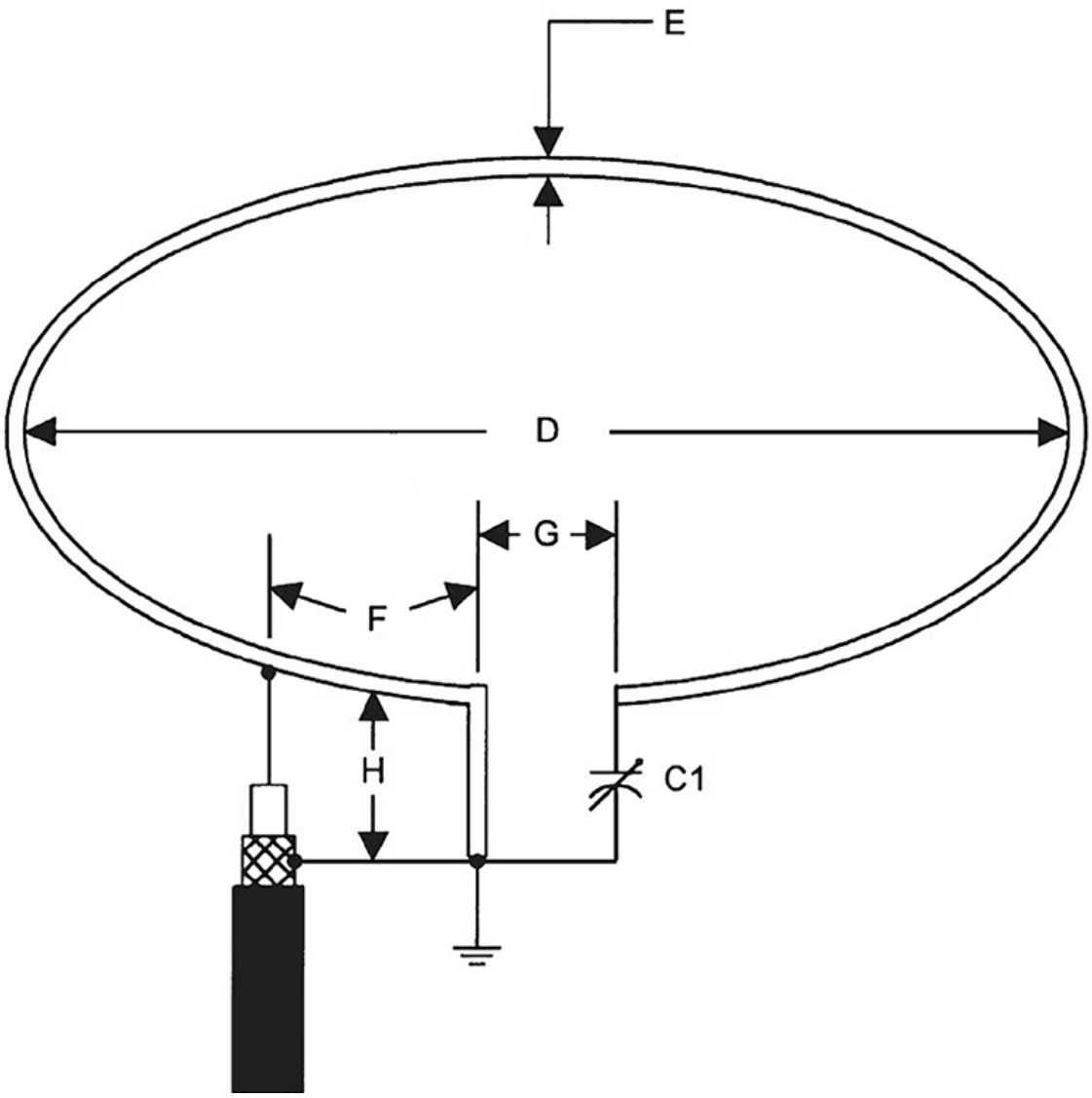


FIGURE 25.5 DDRR antenna.

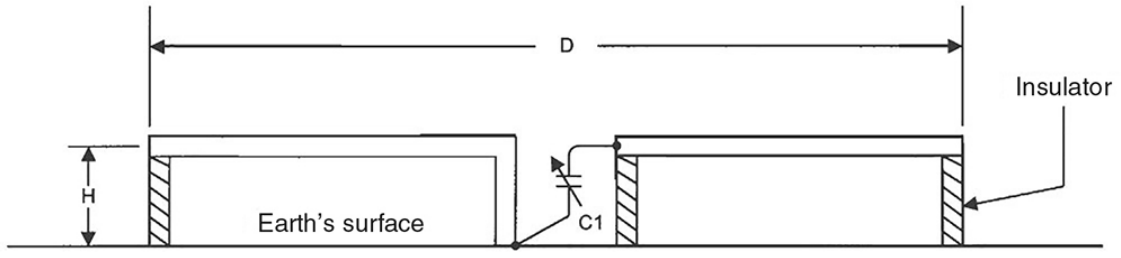


FIGURE 25.6 DDRR antenna: side view.

Because of the loop, some people call this the hula hoop antenna. The far end of the loop is connected to ground through a

small-value tuning capacitor C_1. The actual value of C_1, which is used to resonate the antenna at a particular operating frequency, is found experimentally.

The feedline of the DDRR antenna is coaxial cable; its shield is grounded at the bottom end of the vertical section. The center conductor is connected to the ring radiator at a distance of *F* from the vertical section in a circular variant of the autotransformer or gamma match. The length *F* is determined by the impedance that must be matched. The radiation resistance is approximated by

$$R_{\text{rad}} = \frac{2620H^2}{\lambda^2} \quad (25.1)$$

where *H* is the height of antenna above ground plane.

The normal attitude of the DDRR for communications is horizontal. However, for maximum signal when monitoring Jupiter, the entire antenna assembly, including the ground-plane screen, can be tilted to track Jupiter's position in the sky.

25.2.3 Helical Antennas

The helical antenna (Fig. 25.7) provides moderately wide bandwidth and circular polarization. The antenna (of diameter *D*) will have a circumference *C* of 0.75*λ* to 1.3*λ*. The pitch of the helix (*S*) is the axial length of a single turn, while the overall length $L = NS$ (where *N* is the number of turns). The ratio *S/C* should be 0.22*λ* to 0.28*λ*. At least three turns are needed to produce axial-mode main lobe maxima.

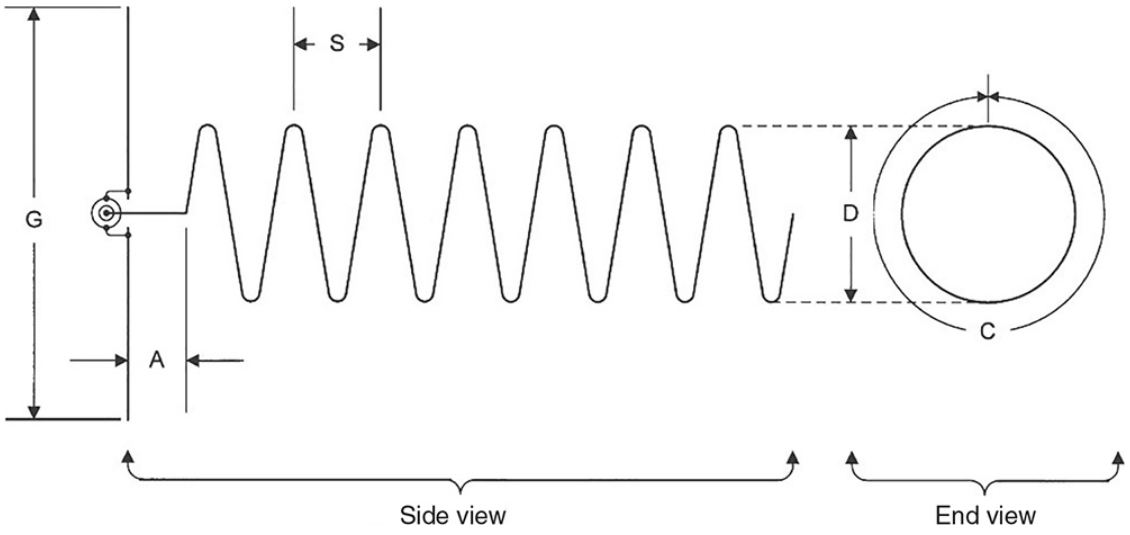


FIGURE 25.7 Helical antenna.

The diameter or edge of the ground plane *G* should be on the order of 0.8*λ* to 1.1*λ* if the conducting surface is circular or square, respectively. The offset between the ground plane and the first turn of the helix is 0.12*λ*. The short section between the helix and the ground plane is terminated in a coaxial connector, allowing the antenna to be fed from the rear of the ground plane. The feedpoint impedance is approximately 140 Ω.

The approximate gain of the helical antenna is found from

$$\text{Gain} = 11.8 + 10 \log \left(C^2 NS\right) \text{ dB} \quad (25.2)$$

The pitch angle ϕ and turn length γ for the helical antenna are given by

$$\phi = \tan^{-1}\left(\frac{\pi D}{S}\right) \quad (25.3)$$

and

$$\gamma = \sqrt{(\pi D)^2 + S^2} \quad (25.4)$$

The beamwidth of the helical antenna is

$$\theta = \frac{\mathrm{K}}{\sqrt{\mathrm{NS}}} \quad (25.5)$$

where *N* is the number of turns, *S* is the pitch in wavelengths, *C* is the circumference in wavelengths, and *K* is 52 for the 3-dB beamwidth and 115 for the beamwidth to the first null in the pattern.

25.2.4 Multiple Helical Antennas

Stacking helical antennas allows a radiation pattern that is much cleaner than the normal one-antenna radiation pattern. It also provides a good way to obtain high gain with only a few turns in each helix. If two helixes are stacked, then the gain will be the same as for an antenna that is twice the length of each element, while for four stacked antennas the gain is the same as for a single antenna four times as long. Figure 25.8 shows a side view of the stacked helixes.

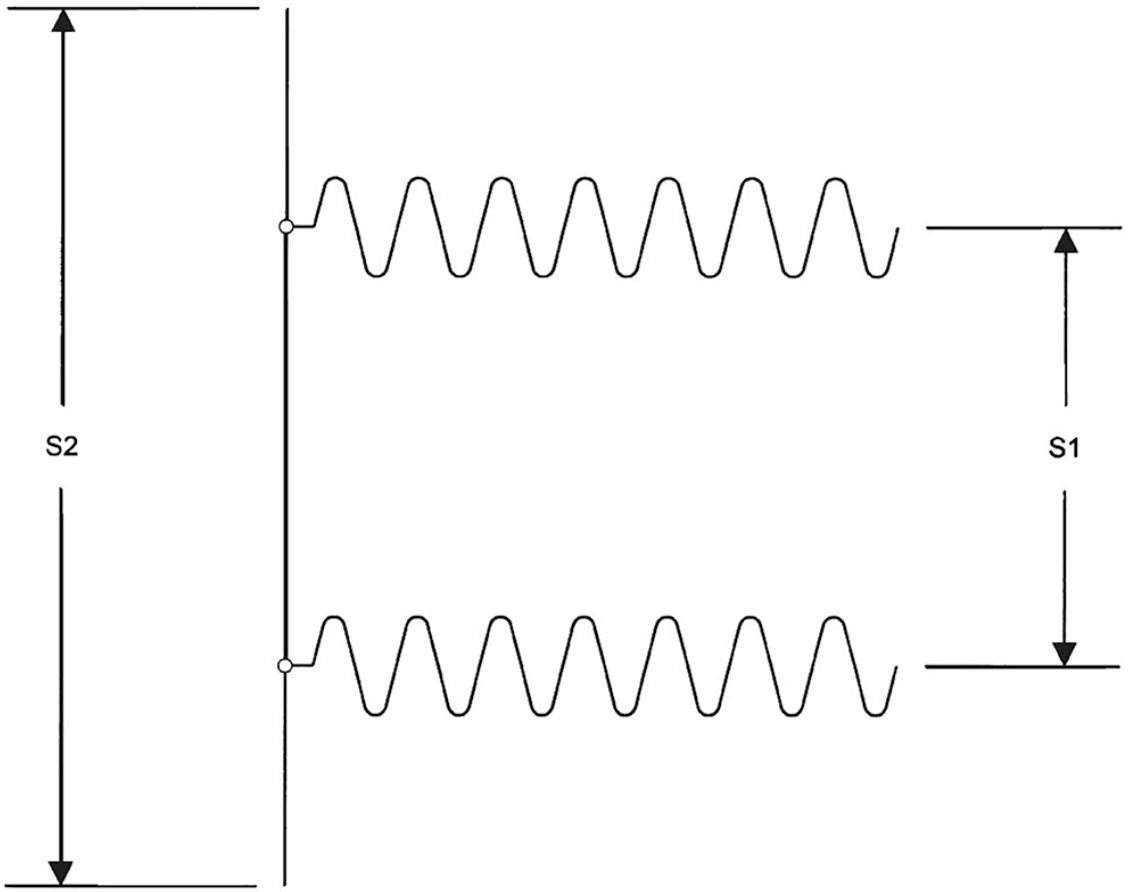
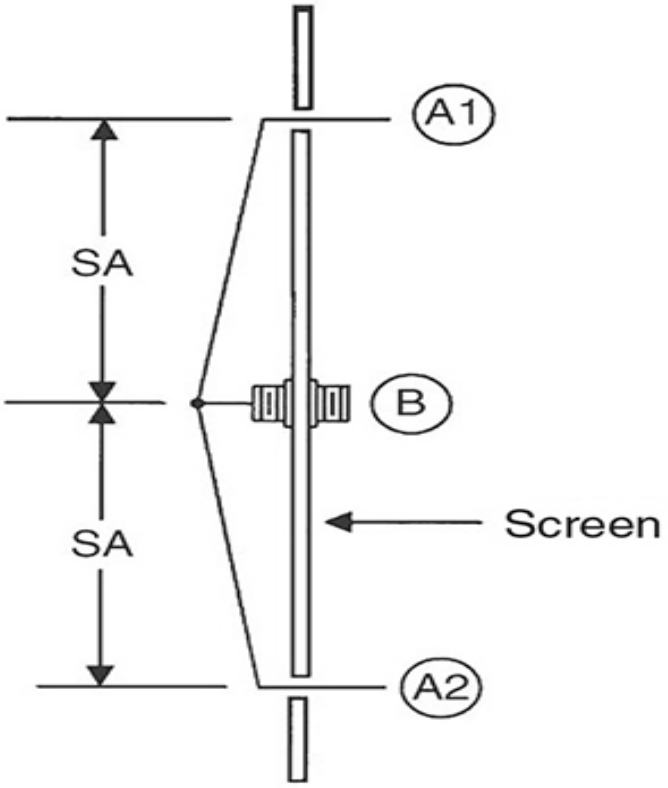
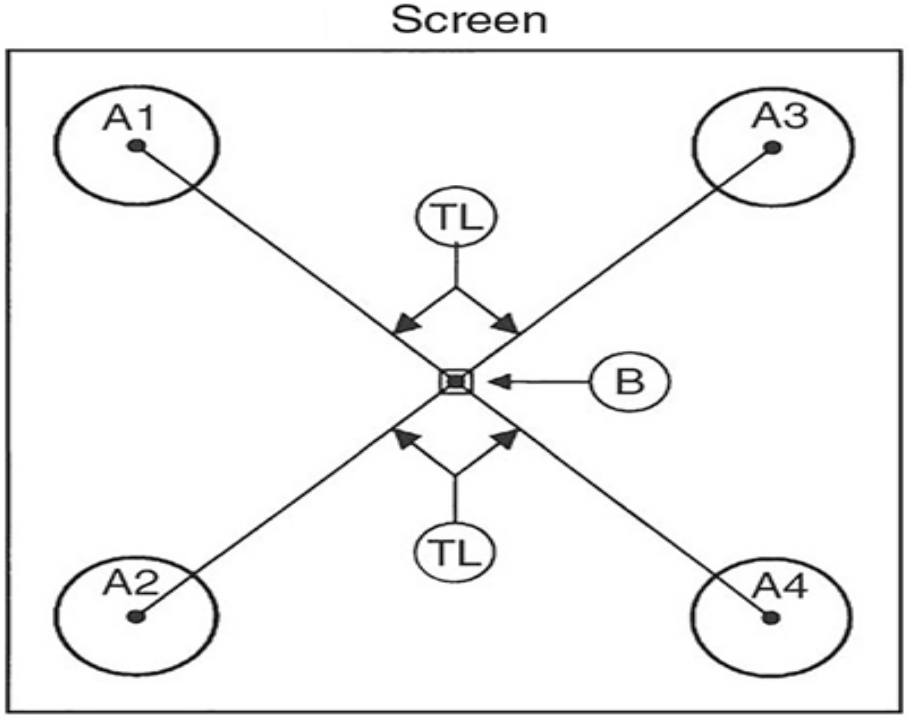


FIGURE 25.8 Dual/quad helical antenna.

The feed system for stacked helixes is more complicated than for a single helix. Figure 25.9A shows an end view of a set of four stacked helical antennas. Tapered lines (TL) are used to carry signals from each element and the coaxial connector (B). In this case, the coaxial connector is a feed-through barrel or back-to-back SO-239 device at the center of the ground plane (B). A side view of the tapered line system is shown in Fig. 25.9B. The length of the tapered lines is 1.06λ, while the center-to-center spacing between the helical elements is 1.5λ. The length of each side of the ground plane is 2.5λ. In the case of Fig. 25.9, the antenna is fed from the front of the ground plane.



B



A

FIGURE 25.9 Front view of quad helical antenna.

25.3 Interferometeric Arrays

The resolution of an antenna is a function of its dimensions relative to the wavelength λ of the received signal. Better resolution can be achieved by increasing the size of the antenna. For high resolution, when a single antenna becomes impractically large, multiple antennas can be used.

Figure 25.10 shows an interferometer array. Two antennas, each with aperture size a, are spaced S wavelengths apart. The radiation pattern is a fringe pattern (Fig. 25.11), consisting of a series of maxima and minima. The resolution angle ϕ to the first null is

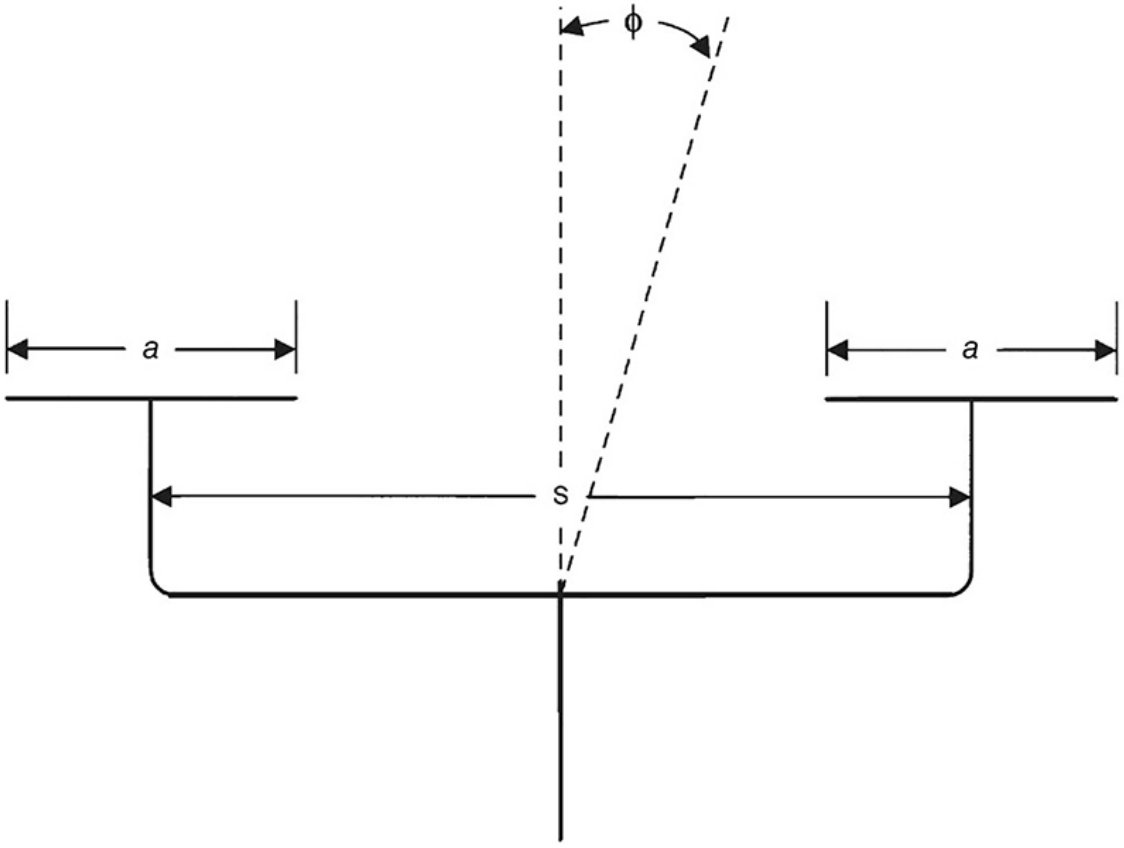


FIGURE 25.10 Interferometer array.

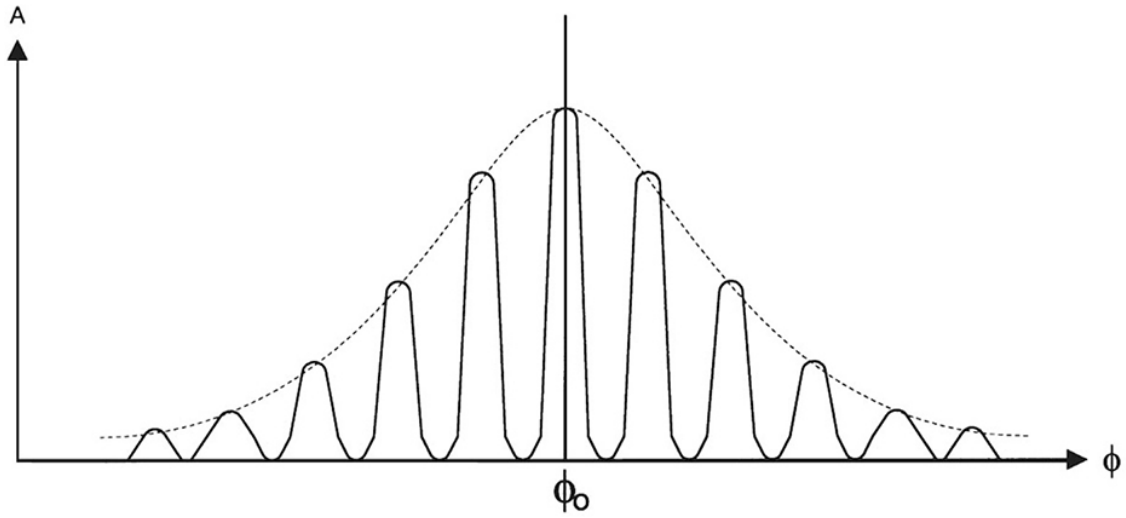


FIGURE 25.11 Interferometer pattern.

$$\phi = \frac{S + a}{57.3} \quad (25.6)$$

The interferometer can be improved by adding antennas to the array. The longest distance between antennas, or the longest baseline, determines the resolution of the interferometric array.

The elements of an interferometric array can be nearly any type of antenna. At frequencies below a few hundred MHz, the effective area of a dipole is large enough that dipoles can be used as array elements, such as in the European Low-frequency Array (LOFAR). At higher frequencies, dish antennas are used.

To achieve the highest possible resolution, signals can be combined from radio telescopes on different continents, as was done to create the first image of a black hole, or antennas can be placed in orbit on spaceborne platforms to further increase the resolution.

25.4 Dish Antennas

For radio astronomy applications in the hundreds of MHz through hundreds of GHz and even THz bands, parabolic reflector antennas are used. For amateur use, TV receive-only (TVRO) satellite dish

antennas have long been employed, and in some cases larger dishes can be obtained when they are surplused by government agencies. Software-defined radios (SDRs) and other commercially available hardware can be used for amateur radio astronomy.

Dish antennas can be used to observe radio reflections from meteor trails, create solar and planetary images, detect signals bounced from the moon, observe bright radio sources like Cygnus A, Taurus A, Cassiopeia A, Sagittarius A, and map emissions at 1.4 GHz from hydrogen in the galactic plane.

One of the major expenses of a dish-based radio telescope is the antenna pointing and tracking system. High gain means better signal reception from weak astronomical sources, but with high gain comes a narrow beam and the need to accurately aim the antenna at the source of interest. The larger and heavier the dish, the more costly the required positioner. Many observatories have outreach programs that allow the use of fairly large and sophisticated radio telescopes, and a wide array of information can be found on home-built receivers.

CHAPTER 26
Radio Direction-Finding (RDF) Antennas

Radio direction finding is the practice of using antennas to pinpoint the location of a distant source of RF energy. When the Federal Communications Commission (FCC) or other telecommunication authorities want to locate an illegal station that is transmitting, one of their primary tools is the use of radio direction finders to triangulate the bootleg station's position. In the same way that the global positioning system (GPS) receiver in your navigation system or smart phone determines where the device is located, monitors can determine the location of a transmitter from the intersection of bearing lines obtained at two or more receiving sites. In contrast to the accuracy available from GPS satellites, there is a fair degree of ambiguity in the bearings obtained for terrestrial targets, because multiple propagation modes, paths, and anomalies can obscure the true bearing of the signal source. To reduce uncertainty, radio direction finders may use several sites.

Direction finding has been used for aviation and maritime navigation in peacetime as well as during international conflicts. During the 1950s and early 1960s, AM radio dials in the United States were marked with two circled triangles at 640 and 1240 kHz. These were CONELRAD frequencies that people were encouraged to tune to in case of a nuclear attack. In each area of the country, selected radio stations would take turns transmitting in a carefully orchestrated and secret sequence. Local broadcast news and information services could be maintained at the same time the enemy was being prevented from using any AM broadcast signals for RDF because no one station was on the air long enough to allow a fix.

Radio direction-finding equipment utilizing the AM broadcast band looks much like Fig. 26.1. A receiver with an *S*-meter to measure signal strength is equipped with a rotatable ferrite loopstick antenna to form the RDF unit. A 360-degree scale around the perimeter of the antenna base can be rotated to line up with true north so that the compass bearing can be read off directly.

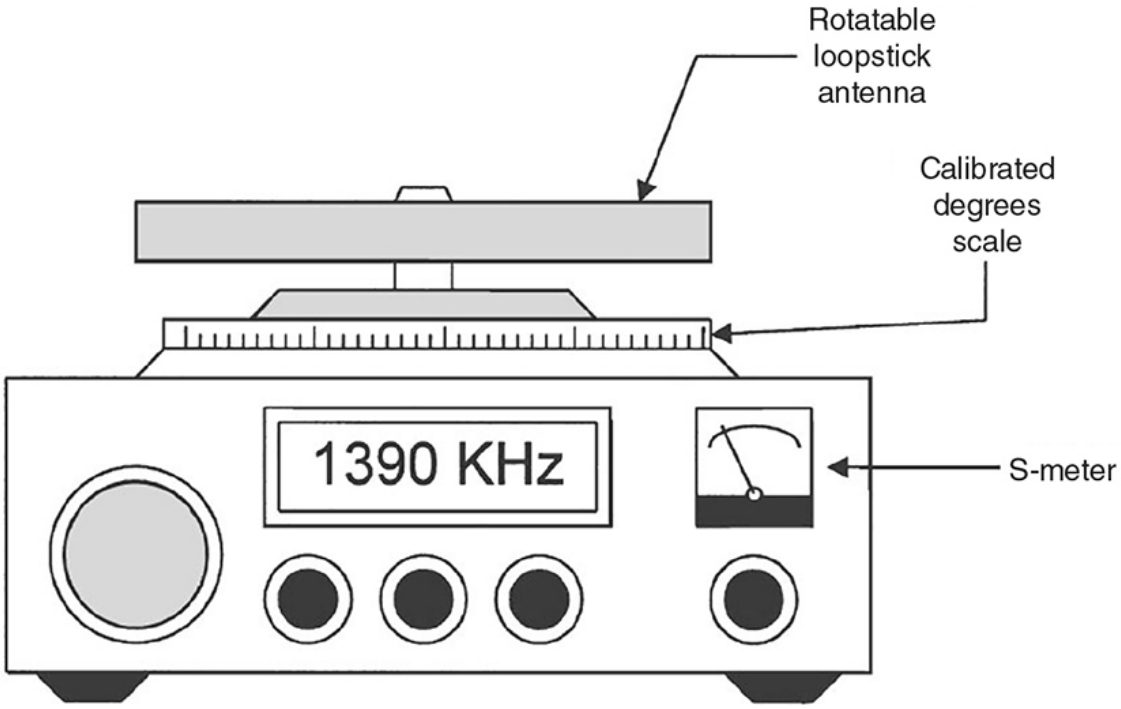


FIGURE 26.1 Radio direction-finding (RDF) receiver with loopstick antenna.

Loopstick antennas exhibit a figure eight reception pattern (Fig. 26.2) with the maxima (lobes) parallel to the loopstick rod and the minima (nulls) off the ends. When the antenna is broadside to the signal source, the received signal strength is at its maximum. The maxima are so broad that it is virtually impossible to find the true point on the compass dial where the signal peaks. The minima, on the other hand, are very sharp, and a good fix on the source can be obtained by nulling the signal. The null is found by rotating the antenna until the audio disappears or the S-meter dips to its minimum observed value.

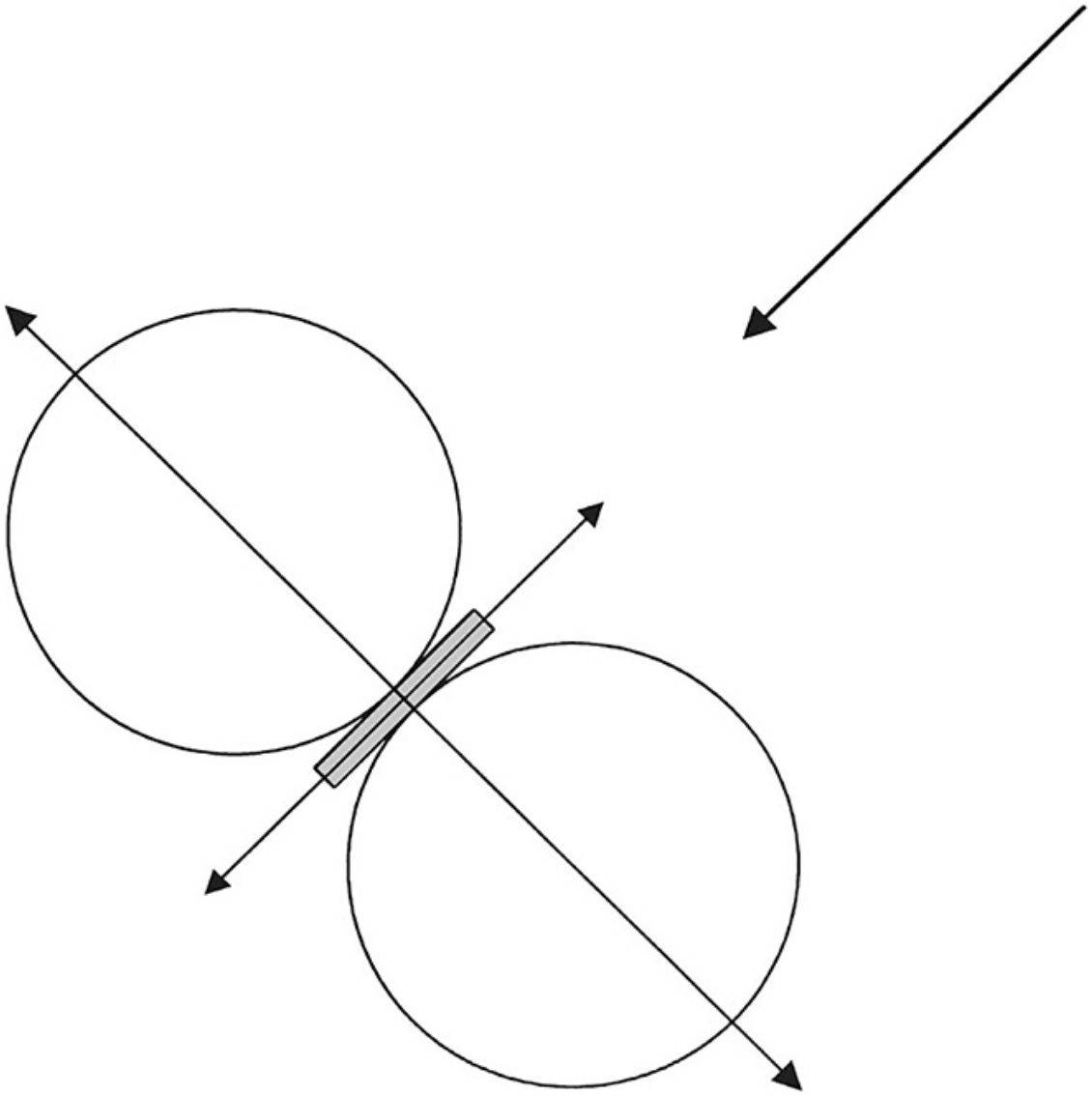


FIGURE 26.2 Pattern of loopstick antenna with the source arriving at a null.

Because the antenna is symmetric about its axis, there are two minima in the horizontal plane. The response is the same no matter which end of the stick is in the direction of the transmitter. The unassisted loopstick can only show you a line along which the transmitter is located, but can't tell you which half of the line to focus on.

One solution to the ambiguity problem is to add a sense antenna to the loopstick (Fig. 26.3). The sense antenna is an

omnidirectional vertical whip. Its signal is combined with that of the loopstick in an RC phasing circuit. When signals from the two antennas have roughly equal amplitudes, the resultant pattern will resemble Fig. 26.4. This pattern is called a cardioid because of its heart shape. Because it has only one null, it resolves the ambiguity of the loopstick pattern.

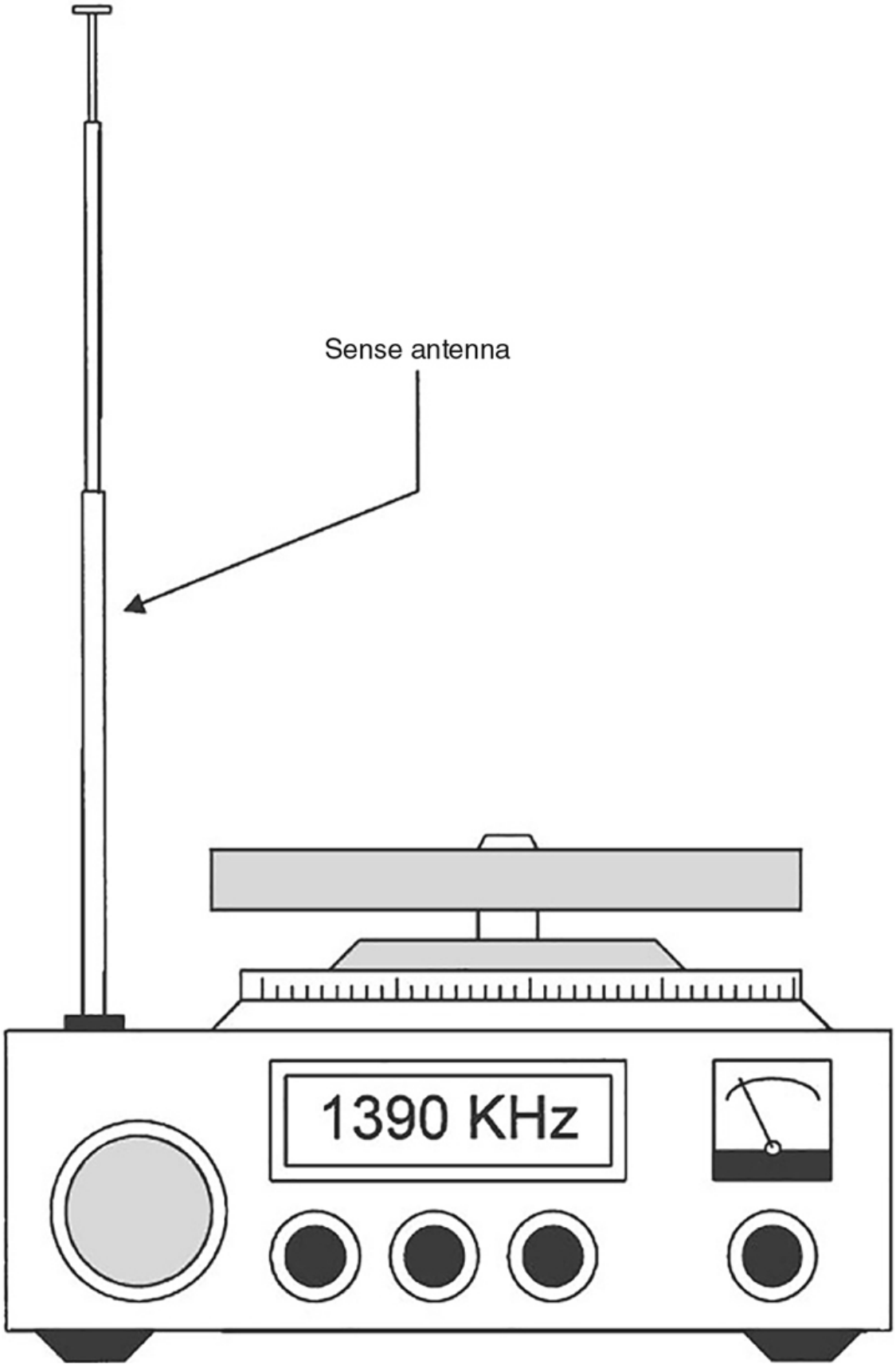

Sense antenna
1390 KHz

FIGURE 26.3 Addition of sense antenna overcomes directional ambiguity of the loopstick.

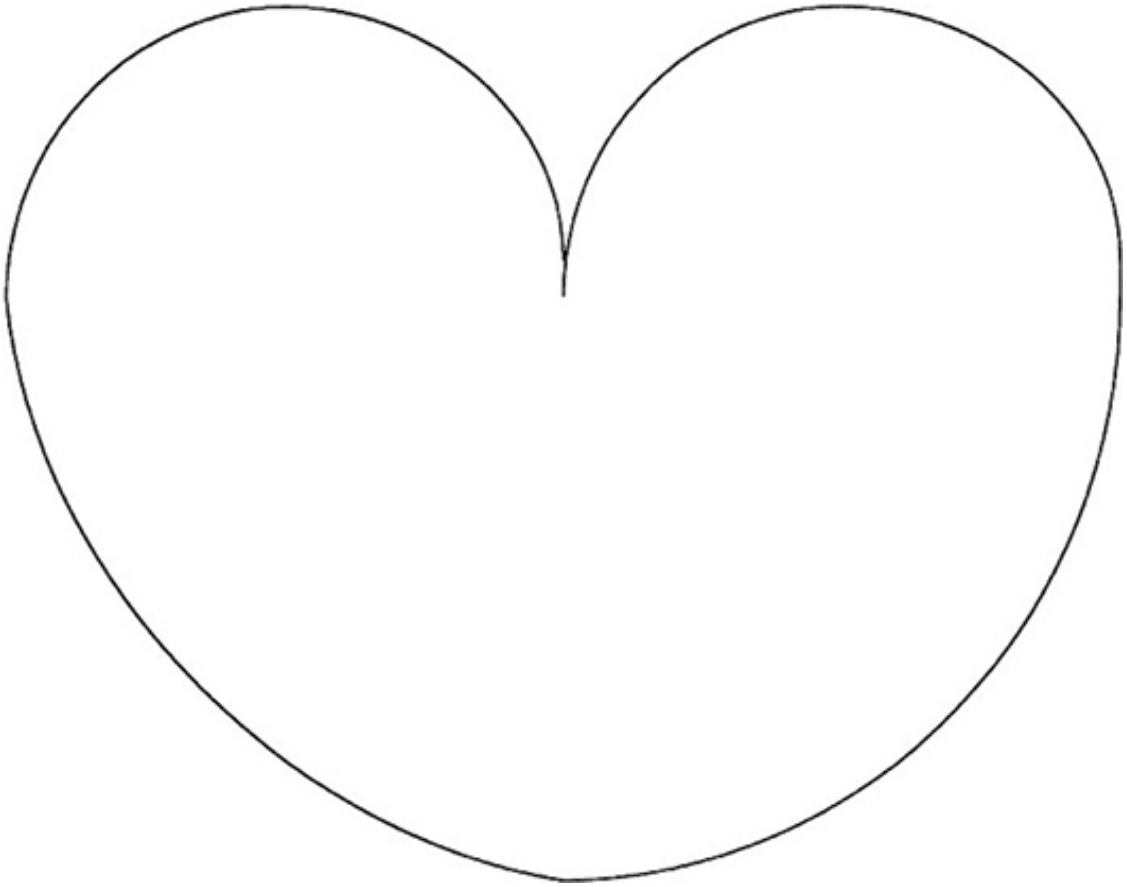


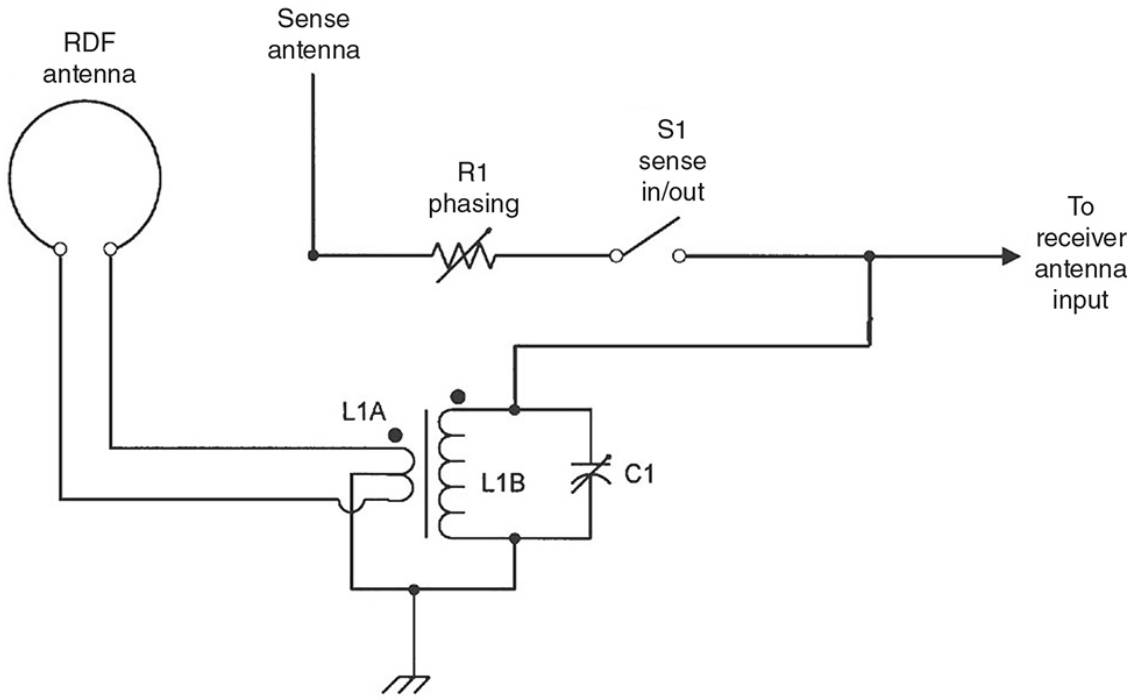
FIGURE 26.4 Cardioid pattern of sense-plus-loopstick antennas.

Even with the sense antenna, a single RDF receiver can provide only bearing information. Receivers at two or more sites and widely separated are necessary to obtain distance information and pinpoint the location of the target.

26.1 Sense Antenna Circuit

Figure 26.5 shows a method for summing together the signals from an RDF antenna and a sense antenna. The two terminals of the loop are connected to the primary winding of an RF transformer.

The primary (L1A) has a grounded center tap. The secondary of the transformer (L1B) is resonated by a variable capacitor C1.



F IGURE 26.5 Sense circuit for a loop/sense antenna RDF.

The top of L1B is connected to both the sense circuit and the receiver antenna input. Potentiometer R1 is a phasing control. The value of this pot is usually 10 to 100 kΩ, with 25 kΩ common. Switch S1 is used to disconnect the sense antenna from the circuit. Because the nulls of the loop or loopstick are typically deeper than the null on the cardioid pattern, the null is first located with the switch open. If the receiver S-meter reading drops when the switch is closed, the correct null has been selected. If not, reverse the direction of the antenna and repeat the process.

26.2 RDF at VHF and UHF

At lower frequencies, antennas for direction finding are relatively large, but at VHF and above, antennas with unidirectional patterns are smaller and more portable. A four-element 6-m Yagi is

reasonably easy to manipulate by hand, and the task gets simpler as one goes higher in frequency. At microwave frequencies, a parabolic reflector can be used.

26.3 Fox Hunting

An activity popular with ham radio operators is fox hunting. A mobile or portable transmitter is hidden. The "hunters" use RDF to locate the fox. The first to locate the antenna might become the fox in the next round. Fox hunting can be done at 10 m or 75 m as well as VHF and UHF and higher frequencies.

26.4 Tracking Down RFI

Direction finding can be used to track down sources of radio frequency interference (RFI). Radio frequency noise is emitted by appliances, computers, power lines, lamps, street-lights, and many other sources.

A good technique is to start hunting on the frequencies where you first notice the noise and then move upward in frequency as you get closer to the source. An initial search can be done while mobile by tuning an AM broadcast radio to an unused frequency near the top of the broadcast band (somewhere in the 1600- to 1700-kHz range). Keep your portable receiver in AM mode if at all possible.

At higher frequencies, smaller antennas can be used, and since RFI emissions are often weaker at higher frequencies, if the noise can be detected at all you may be close to the source. When it's time to switch to VHF antennas, amateur 2-m FM handhelds or mobile rigs that include coverage of the AM aircraft band between 108 and 130 MHz are especially useful for zeroing in on a specific RFI source.

Power line and power pole noise is particularly difficult because the noise can propagate for miles along the lines and go through many maxima and minima before tapering off. Some RFI is the

result of arcing when insulating surfaces get wet; others are the exact opposite. Some noises result from outdoor lighting powered by day/night sensors, and so on. To determine if the noise source is in your own home, turn off your house power at the main breaker box at a time when the noise is evident in your receiver. Continue to run your equipment on a 12-V car battery or an uninterruptible power supply (UPS).

If an S-meter is unavailable, the audio noise output of the speaker in a mobile VHF/UHF rig can be used with an audio signal level monitor smart phone app. For serious searches, two operators are typically required—one to maneuver the antenna and another to watch the meter and call out "Warmer!" or "Colder!" as the signal level goes up or down.

26.5 Shortwave and AM BCB "Skip" RDF

Radio direction finding is most accurate over relatively short distances, but with a little practice, you can RDF distant stations using skip propagation. There are some practical difficulties associated with skip RDFing. When we look at propagation drawings of skip in textbooks we usually see what's happening in a single plane at a time. Typically, a single imagined pencil beam radio signal is shown traveling at some angle up the ionosphere, where it is refracted back to earth. Unfortunately, the real world is not so neat and crisp.

An actual skywave signal is an amalgam of beams coming in from many directions and elevation angles. The strongest component is the path that has the least total loss from start to end. That path changes over time because the ionosphere is a continuously shifting medium. It's possible that the preferred path to your antenna for a signal source located to the northwest of you might be a so-called skew path, coming in from as far around the compass rose as the southwest.

Terrestrial reflections also cause problems, especially when RDFing a station in the high end of the HF band or the VHF/UHF

bands. Radio waves will reflect from geological features such as mountains and from man-made structures. If the reflection is strong enough, it might appear to be the real signal and cause a severe error in RDFing. Be wary of RDF results on HF when the "skip is in." Similarly, don't let the canyons of New York City or the mountains of upstate fool you on VHF and UHF.

26.6 Adcock Antennas

The Adcock antenna (Fig. 26.6) has been around since 1919, when it was patented by British Royal Engineer and Lieutenant Frank Adcock as a way of overcoming the tendency of RDF loops to lose directional accuracy at nighttime. Adcock surmised that the horizontally polarized component of incoming skywave signals induced loop output voltages indicating spurious directions when the output voltage due to the vertically polarized ground wave was near zero. To minimize this error, Adcock designed a number of different configurations based on the use of vertical elements, both monopole and dipole.

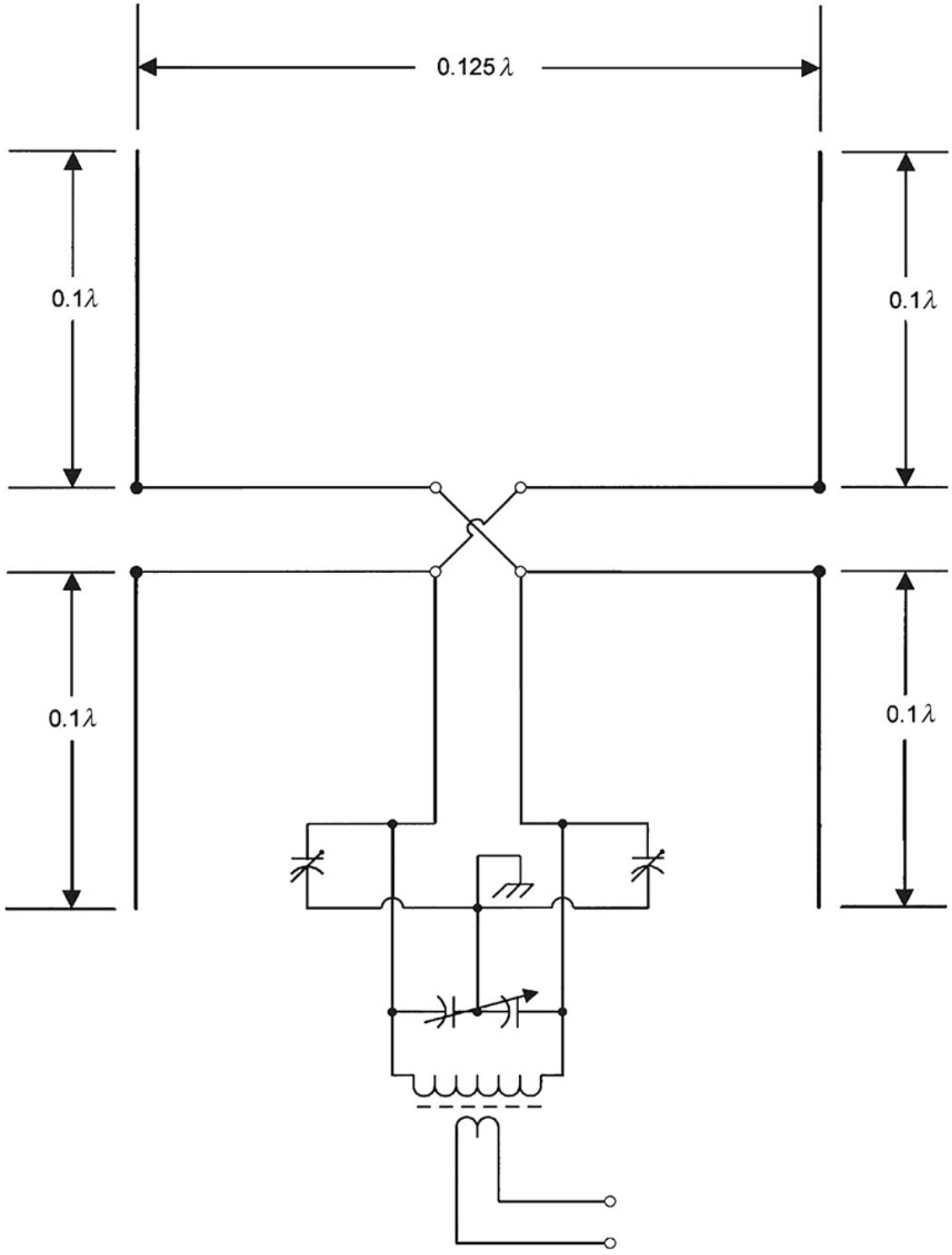


FIGURE 26.6 Adcock array RDF antenna.

The Adcock antenna shown here once found use in the United States by the FCC at monitoring sites. It consists of two center-fed,

nonresonant vertical radiators. Each side of each element is at least 0.1λ long, but need not be resonant (which means the antenna can be used over a wide band). The elements are spaced from 0.1λ to 0.75λ, although the example shown here is spaced 0.125λ. Spacing is ultimately limited by the need to avoid distorting the circular (omnidirectional) pattern of each element in the pair, so a given dipole configuration is suitable over a 3- or 4-to-1 frequency range. Pickup of horizontally polarized signals on the feedline to each element is minimized through the use of balancing or shielding techniques.

Other Adcock arrays are designed around two pairs of vertical dipoles or monopoles. All Adcock antennas are vertically polarized and respond only to vertically polarized waves, so they have earned an excellent reputation as being superior to loops for precise high-frequency shortwave RDFing.

A typical pattern for an Adcock antenna is shown in Fig. 26.7. The example antenna is a 10-MHz (30-m band) Adcock that uses 1.455-m elements (total 2.91 m on each side), spaced 4 m apart. The pattern is a traditional figure eight with deep nulls at 0 degrees and 180 degrees. The antenna can be rotated to find a null in the same manner as a loop.

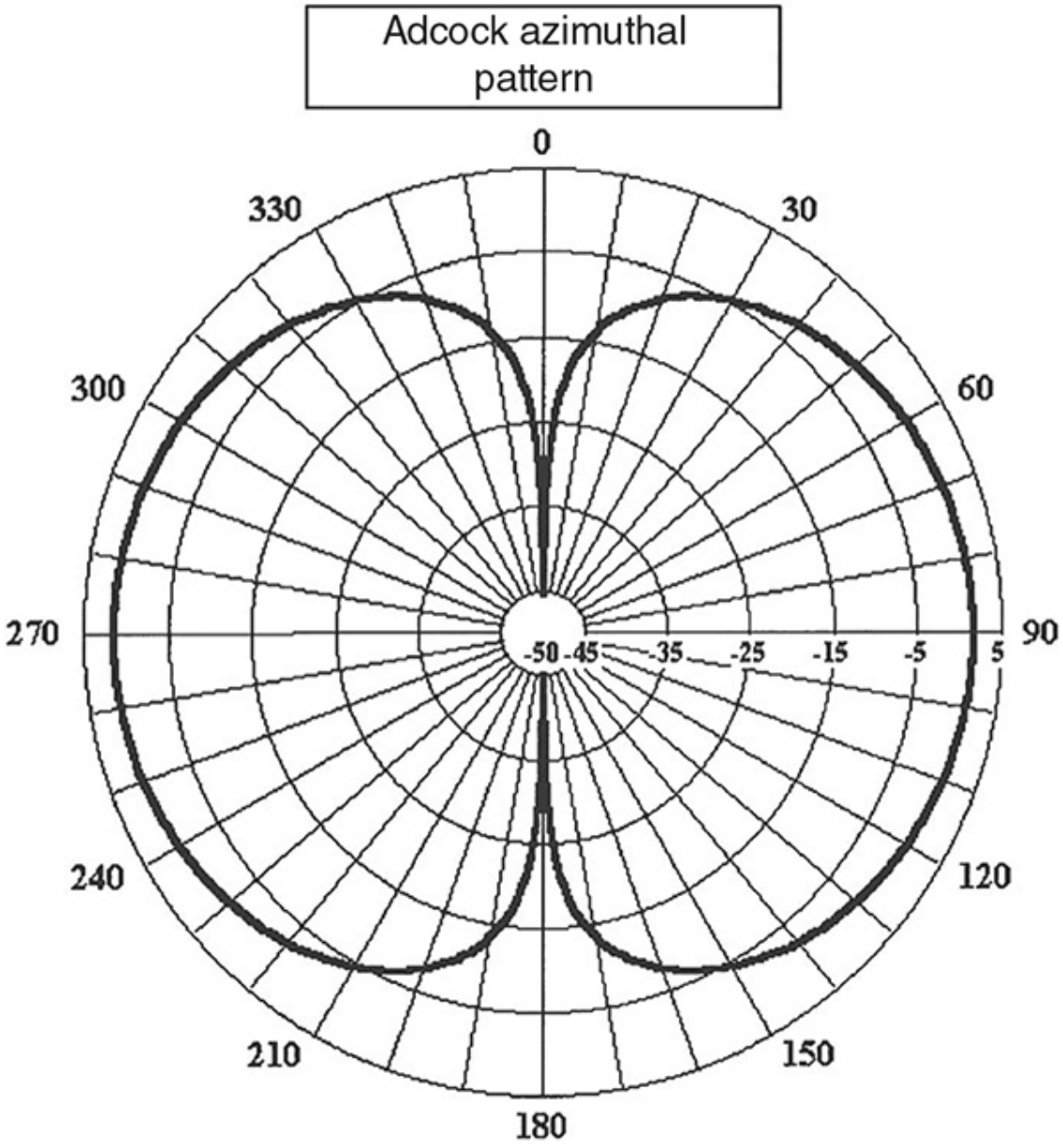


FIGURE 26.7 Adcock antenna pattern modeled using NEC.

26.7 Watson-Watt Adcock Array

Figure 26.8 shows the Watson-Watt Adcock RDF array, consisting of two Adcock pairs arranged orthogonally to each other. It is common practice to align one Adcock with an east-west line and the other with north-south. These are fed to identical receivers that are controlled by a common local oscillator (LO). The outputs of the

receivers are balanced and are used to drive the vertical and horizontal axes on an oscilloscope. Figure 26.9 shows the patterns that result when signals of various phases arrive at the Watson-Watt array. The patterns of Figs. 26.9 A and B are from signals 180 degrees out of phase, while the pattern of Fig. 26.9C is the result of a 90-degree phase difference.

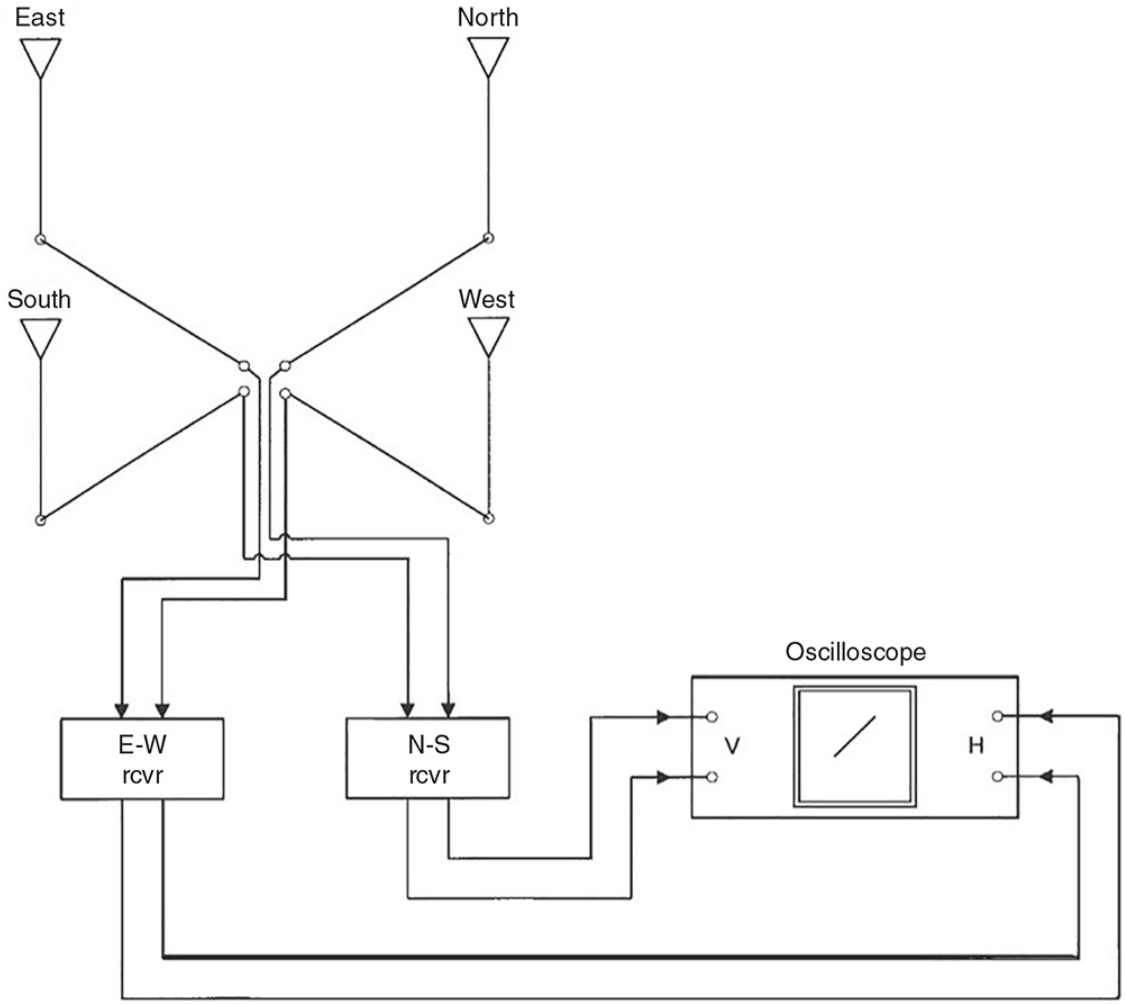
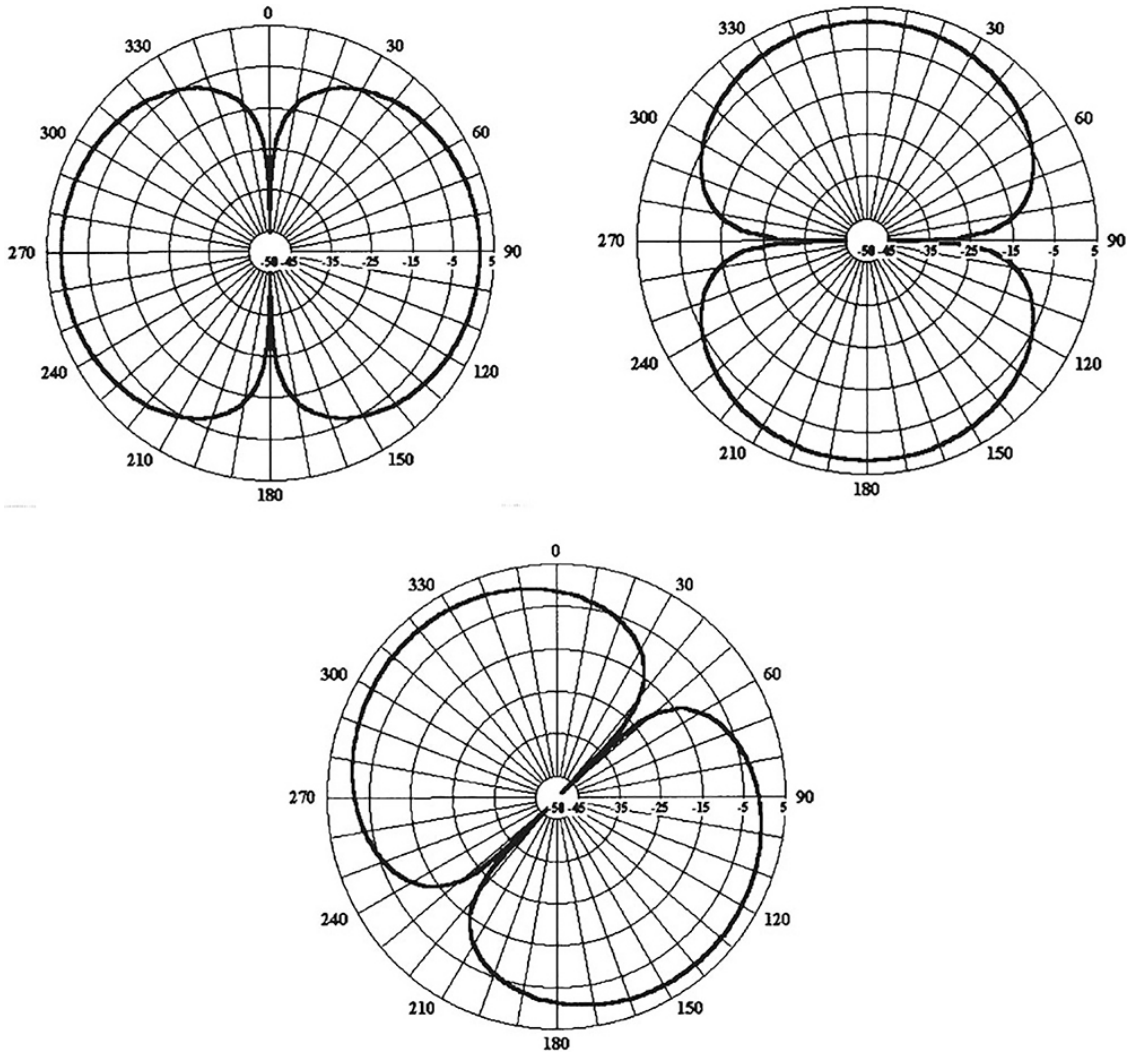


FIGURE 26.8 Watson-Watt Adcock array.



F IGURE 26.9 Patterns for different signals on the Watson-Watt array.

26.8 Doppler RDF Antennas

Figure 26.10 shows the basic concept of an RDF antenna based on the Doppler effect. In a radio system, when distance between a receiving antenna and a signal source is changing, a Doppler shift in the received signal frequency is detected. No Doppler shift occurs as a result of any change in motion at right angles to the line connecting the source and the receiver. When a hobbyist listens to the telemetry signals from a passing satellite or the international Morse code (CW) of a radio amateur making contacts through an

repeater satellite (see www.amsat.org), the received frequency continuously slides lower as the satellite comes into view, passes nearby, and disappears until its next orbit. At the point of the satellite's closest approach to the receiving station, the received frequency equals the actual transmitted frequency because neither antenna has any radial velocity relative to the other for that brief moment. Before then, the received frequency is higher; afterward, it is lower.

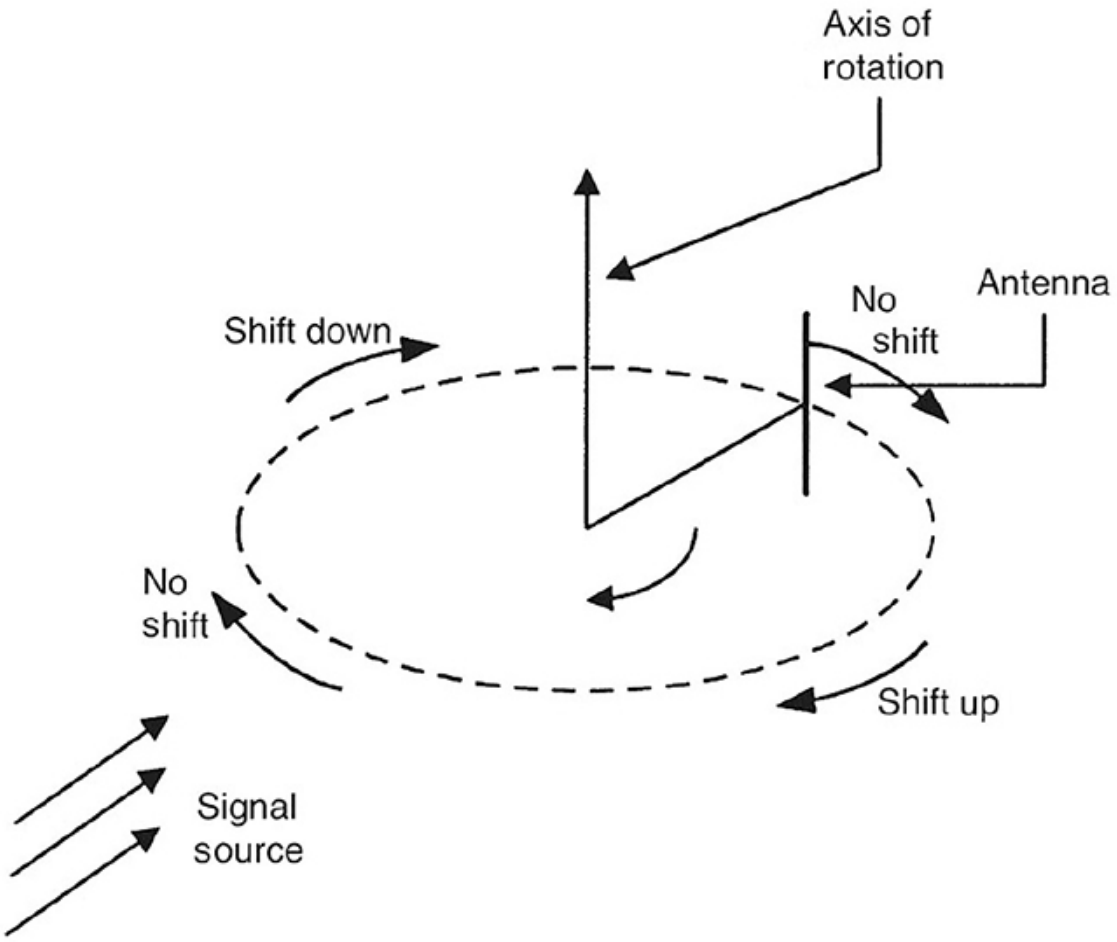


FIGURE 26.10 Doppler antenna.

The Doppler RDF antenna of Fig. 26.10 rotates at a constant angular velocity. The signal wavefront approaches from a single direction, so there will be a predictable Doppler shift at any point on the circular path of the antenna. The magnitude of the

frequency shift is maximum when the antenna is moving directly toward the source or away from it. When the antenna is closest to the source or farthest from it, there is very little Doppler shift. The maximum shift at the two sides can be found to be

$$S = \frac{R\omega f_c}{c} \quad (26.1)$$

where *R* is the radius of rotation, ω is the angular velocity of the antenna,
f_c
is the carrier frequency of the incoming signal, and *c* is the velocity of light.

In practice there are problems with the Doppler method, one of which is getting a Doppler shift large enough to measure. The mechanical rotational speed required of the antenna can be too high for practical use. The effect can be simulated by sequentially (i.e., electronically) scanning a number of antennas arranged in a circle. The result is a piecewise approximation of the effect seen when the antenna is rotated at high speed. In a typical electronically scanned system, verticals are switched to create a Doppler shift that allows pinpointing of the source.

26.9 Wullenweber Array

One of the problems associated with RDF antennas is that imperfections can cause pattern distortions. With a wide-aperture direction finder (WADF), one can average the signals from a large number of antenna elements distributed over a large-circumference circle. The Wullenweber array (Fig. 26.11) is such an antenna. It consists of a circle of many vertical elements. For HF the circle can be 500 to 2000 ft in diameter.

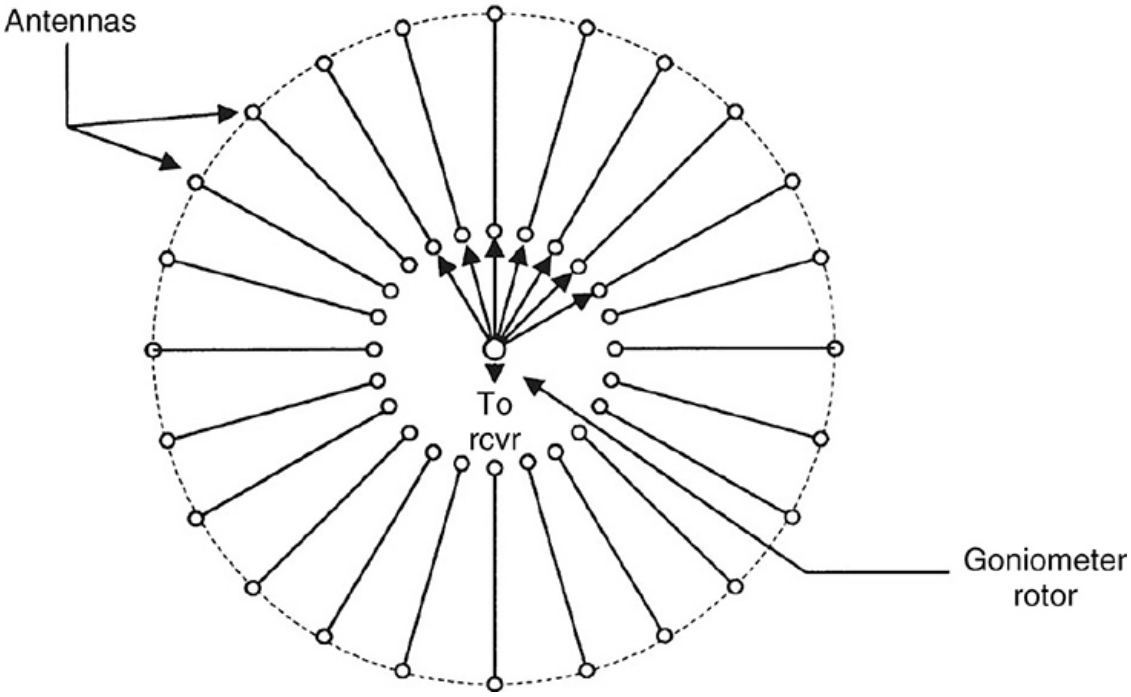


FIGURE 26.11 Wullenweber RDF array.

A goniometer rotor spins inside the ring to produce an output that will indicate the direction of arrival of the signal. The theoretical resolution of the Wullenweber array is on the order of 0.1 degree, although practical resolutions of less than 3 degrees are commonly seen.

26.10 Time Difference of Arrival (TDOA) Array

With two antennas, the direction of arrival of a signal can be estimated from the time-of-arrival difference. Figure 26.12 shows a wavefront advancing at an arbitrary angle to the line between two antennas. If the wavefront is parallel to the antenna axis, it will arrive at both antennas at the same time. If the signal arrives at an angle, it will arrive at one antenna some time before it reaches the second. From the difference in the arrival times at the two antennas we can calculate the arrival angle.

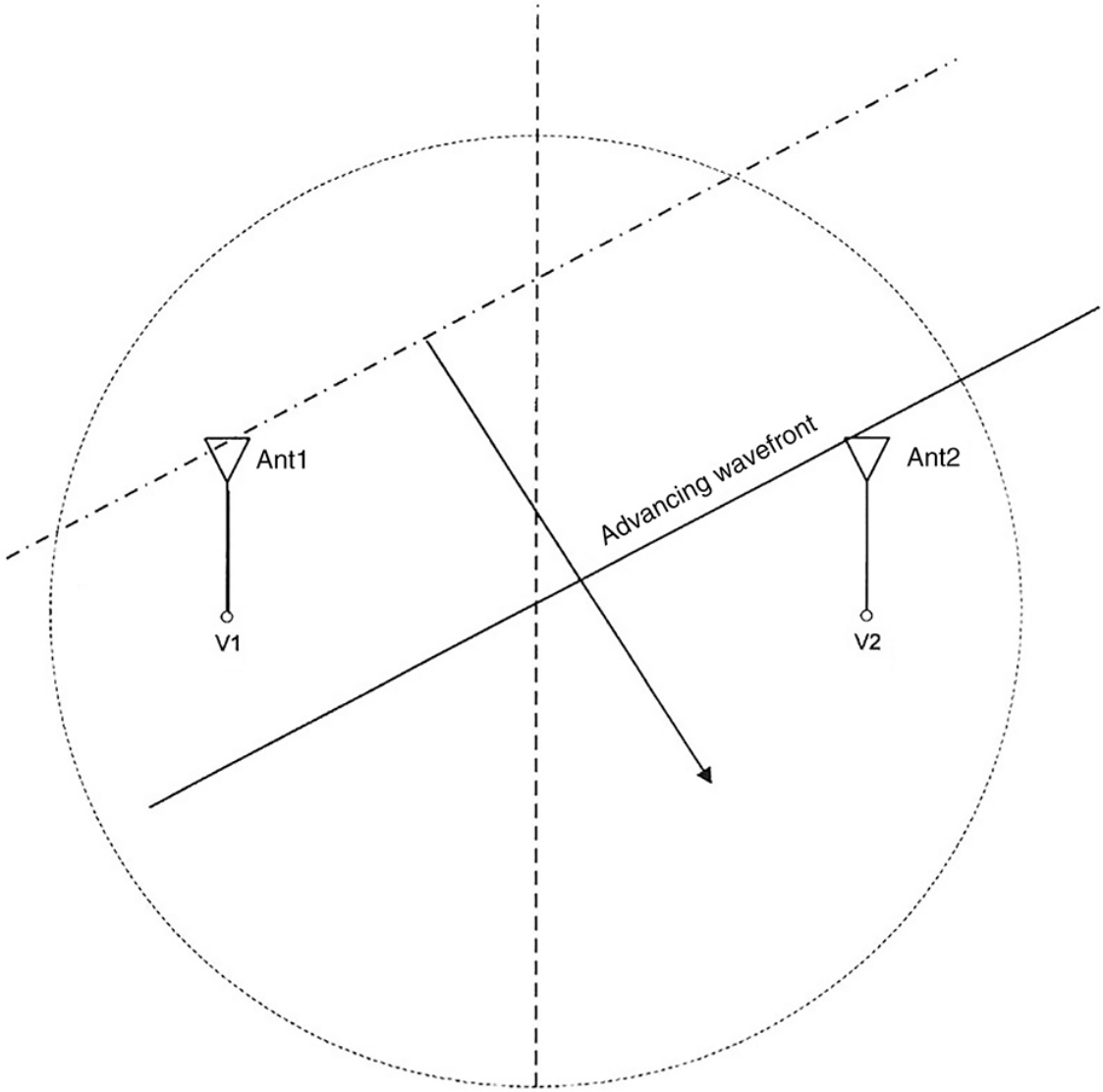


FIGURE 26.12 Time difference of arrival (TDOA) RDF array.

There is an ambiguity in the basic TDOA array in that the combined output will be the same for conjugate angles, i.e., at the same angle on opposite sides of a line bisecting the array. This problem can be resolved by adding the electronic system in Fig. 26.13. Signals from the antennas are detected by separate receivers and then thresholds detected to minimize the impact of noise. The outputs of the threshold detectors are used to trigger a sawtooth generator that controls the horizontal sweep on an oscilloscope. The two signals are delayed and one is inverted so that the operator can distinguish between their traces on a CRT

screen, thus eliminating the ambiguity. TDOA can also be implemented using a software-defined radio (SDR).

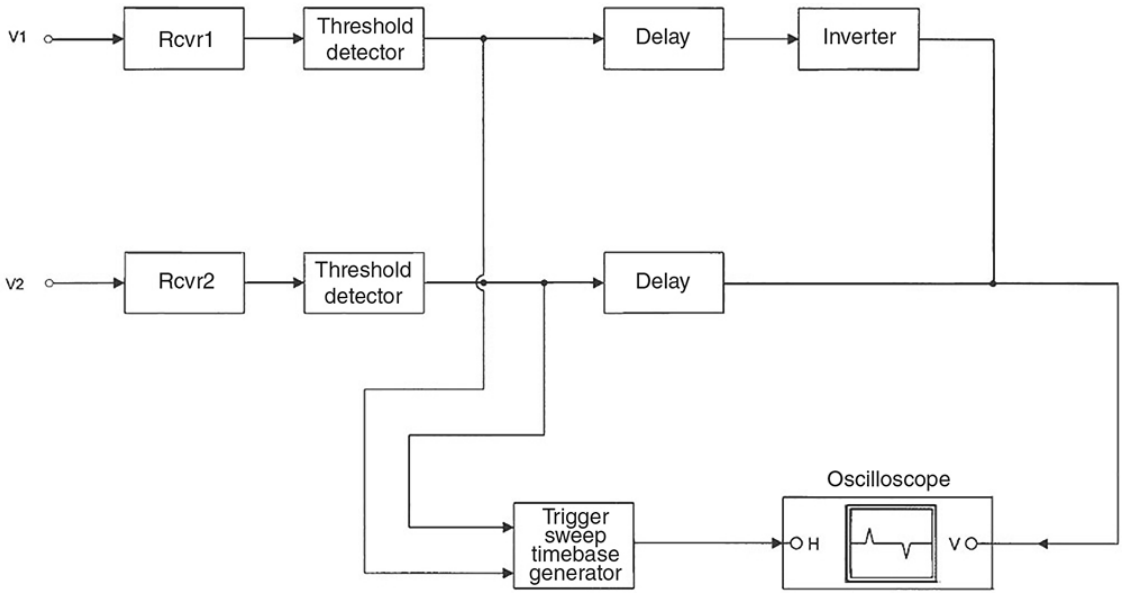


FIGURE 26.13 Block diagram of typical TDOA circuit.

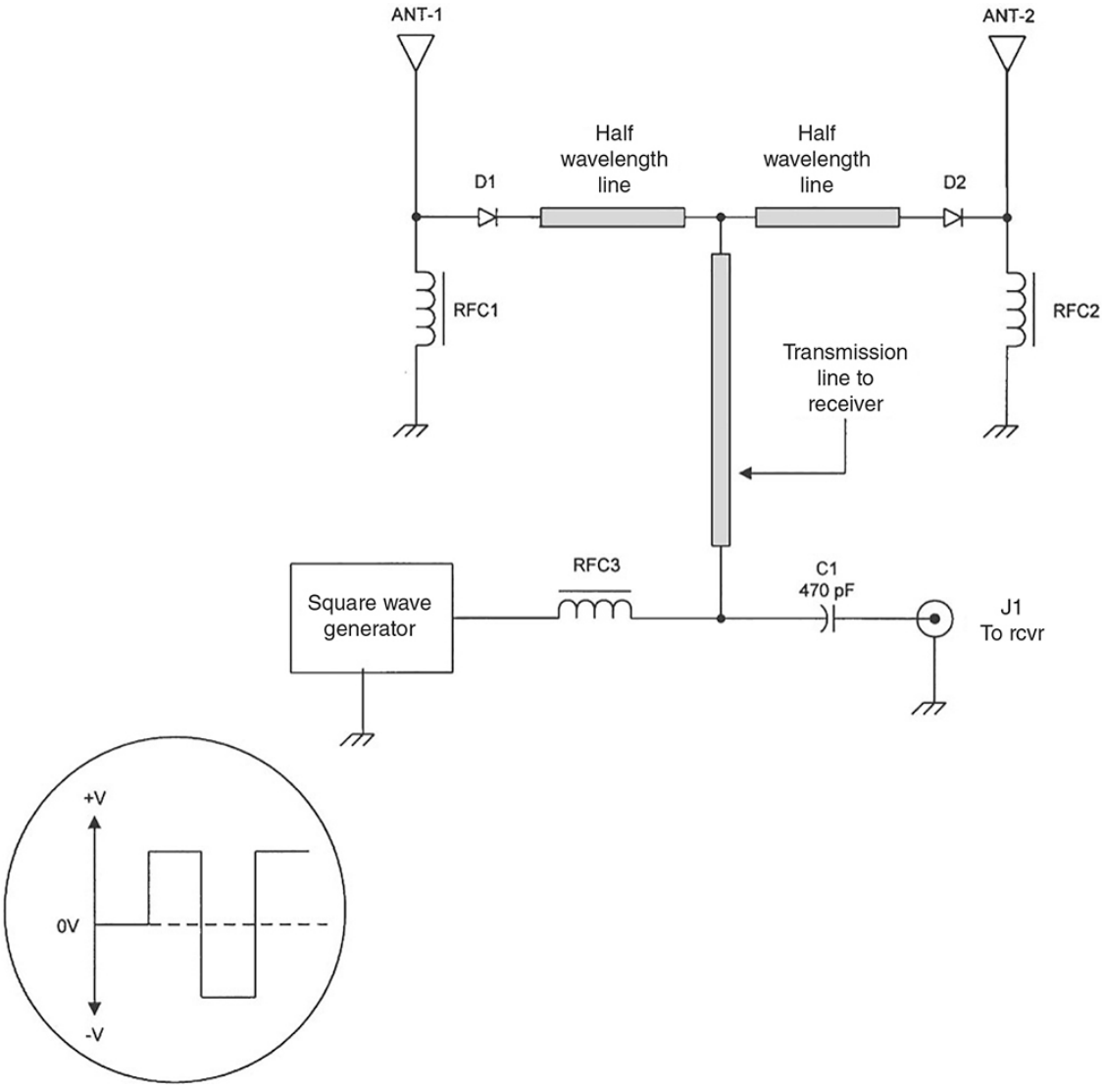
26.11 Switched-Pattern RDF Antennas

If an advancing wavefront arrives at two identical antennas from a direction perpendicular to the line connecting the antennas, the phase of the wavefront will be the same at both antennas. If signals from the two antennas are "chopped" at an audio rate with a square wave and alternately fed to a simple receiver, the resulting phase modulation will appear as an audio tone at the receiver output. Rotation of the antenna array until the tone is nulled out, which occurs when there is no phase difference between the signals at the two antennas, provides an accurate bidirectional fix on the source of the wavefront.

26.12 Double-Ducky Direction Finder (DDDF)

Figure 26.14 shows in simplified form a system originally presented to amateur radio operators by David T. Geiser, W5IXM, for use with

"rubber ducky" VHF antennas, the double-ducky direction finder (DDDF). The DDDF uses short, loaded, and rubber-damped antennas rather than full
$\lambda/4$
or 5⁄8λ whips because vibration and wind can cause enough relative motion between elements of an array built with full-length whips to make it difficult to get a clean null in normal outdoor use.



F IGURE 26.14 Double-ducky switched-pattern RDF system.

Spacing of the two antennas is not critical; 0.25λ to 1λ apart over a good ground plane (such as the roof of a car or truck)

should work well. As with all vertical monopoles, if no ground plane exists, a sheet metal ground plane should be provided. In Fig. 26.14, the antennas are fed from a common transmission line to the receiver. To keep them electrically the same distance apart, identical half-wavelength sections of transmission line are used to connect the antennas to the shared section of line.

Switching is accomplished by using a bipolar square wave and PIN diodes. The bipolar square wave (see inset to Fig. 26.14) has equal positive and negative peak voltages on opposite halves of the cycle. The PIN diodes are connected in opposition such that diode D1 conducts on negative excursions of the square wave and D2 conducts on positive excursions. In any given half-cycle of the audio frequency square wave, one antenna is connected to the receiver when its diode is conducting and the other diode is back-biased, creating a high impedance between its antenna and the shared line. Thus, the active antenna alternates rapidly between ANT1 and ANT2.

The combined signal from the two antennas is coupled to the receiver through a small-value capacitor that blocks the baseband audio frequency square wave from appearing at the receiver input. This allows use of the shared transmission line segment for both the RF signal and the switching signal. An RF choke keeps RF from entering the square-wave generator.

The DDDF antenna phase-modulates the incoming signal at the frequency of the square wave, and the resulting tone can be heard in the receiver output. When the signal's direction of arrival is perpendicular to the line between the two antennas, the phase difference is zero and the audio tone disappears.

The pattern of the DDDF antenna is bidirectional, so the ambiguity problem found with loop antennas exists here as well. The ambiguity can be resolved by placing a reflector $\lambda/4$ behind each antenna. This is an attractive solution, but it tends to distort the antenna pattern a little bit. Alternately, rotate the antenna through 90 degrees or walk an L-shaped path.

26.13 Summary

This chapter has touched on the basics of radio direction finding. Much more depth can be found on the Internet and among other amateurs specializing in RDF technology and techniques. RDF can be very useful for locating RF sources, both good and bad; as such, it is an important tool for locating illegal stations or undesired noise sources. RDF is a great way to build club spirit and teach a wide variety of technical concepts through fox-hunting training exercises and group construction projects. Try it…you'll like it!

CHAPTER 27
Antennas for Wireless Devices

While the focus of this book is on antennas for traditional amateur radio applications, it's worth discussing some of the antennas available for short-range communications and wireless devices. The range of hardware available for short-range wireless communications: internet-of-things (IoT), connected appliances, handheld devices, and vehicle-to-vehicle links has expanded dramatically in recent years. Chipsets, modules, antennas, interfaces, protocols, and software-defined radios are available for a wide variety of amateur radio communications projects. Most of the newer hardware operates in the microwave region at GHz frequencies and above, so antennas tend to be small and packaged with the transmitter and receiver, and other hardware.

The antennas considered in this chapter will help round out our treatment of the basic radiation modes used by antennas. Antennas can be broadly categorized into those that radiate like a dipole, a loop, a resonant patch antenna, or an aperture type antenna fed by one of these basic antennas. Nearly all other antennas can be viewed as modifications of these radiating structures. Over the last few decades, as more antennas have been crammed into small wireless devices, uncounted modifications of these basic antenna types have been developed. Despite differences in appearance, nearly any antenna can be traced back to some notional basic antenna type.

Frequency allocations for wireless devices are rather complicated. Some wireless devices operate in allocated frequency bands for which the operator has paid tens of millions of dollars or more in license fees. Other devices like Wi-Fi nodes operate in bands that are not exclusively allocated to one provider, but the devices

require FCC certification and must operate under transmitted power constraints and other limitations. Certain bands are allocated for industrial, scientific, and medical (ISM) applications and can be used without a license as long as power constraints and other regulatory requirements are respected. ISM bands are used for low-power wireless links like Bluetooth, garage door openers, baby monitors, IoT devices, connected homes, and many other short-range communications applications.

27.1 Design Considerations for Small, Low-Gain Antennas

Antennas for wireless devices must be small. A smart phone can have over a dozen antennas. For the primary voice and data communication links, there may be 4 to 5 or more multiband antennas covering LTE, 5G, and other protocols. The primary communication antennas are often antenna arrays with two or more elements per group of frequency bands. Whereas single antennas are susceptible to fading caused by multipath, wireless devices use antenna arrays to exploit multipath to actually improve channel capacity using multiple input multiple output (MIMO) communication techniques. In addition to the main communication bands, the device will also have antennas for positioning, Wi-Fi networks, short-range links like Bluetooth, ultrawideband antennas for fine positioning, and near-field communications (NFC), which is in the traditional HAM radio HF band and used for point-of-sale transactions. Packing all of these antennas into a small device is a tour de force of engineering.

Wireless devices typically require low-gain antennas. High gain would mean a larger antenna than we typically want, and the device would need to be oriented to aim the beam at a base station or another network node. For this reason, antennas for wireless devices are nearly always low gain and electrically small. When higher gain is needed, an array of low-gain elements can be used with the beam electronically steered as the device moves and

changes orientation to maintain pointing at a base station, as is done in MIMO communication systems. In certain specialized applications, like mobile satellite phones, on occasion an external antenna with moderate gain is used, but this is normally a last resort for handheld devices.

Since the antenna radiation pattern is relatively unimportant for wireless device applications, the antenna only needs to resonate at the proper frequency or frequencies with reasonably low loss. The RF environment in commonly used microwave bands is crowded with interference, so radiation efficiency is less important than it would be for satellite communications and other applications where the ambient noise environment is quieter. An efficient antenna receives more of both the signal of interest and the interference than a lossy antenna, and thermal noise from the antenna is small in relation to the interference, so the overall SNR does not improve much with lower loss and increasing antenna efficiency.

Another issue for antennas in wireless devices is the tradeoff between size and bandwidth discussed in Sec. 5.14. For compact wireless devices, we want a small antenna, but based on fundamental physics that means a narrow bandwidth. Loss in the antenna is normally undesirable, but it allows for a small antenna to have wider bandwidth. Since radiation efficiency is relatively unimportant for wireless devices (efficiencies of around 70% are common), antenna losses are tolerable from an SNR perspective. Dielectric and ohmic losses in the antenna broaden the frequency response, lower the quality factor, and make it easier to design an antenna that is both electrically small and at least moderately wideband.

In the world of wireless devices, one way to look at antenna design is that we simply need a configuration of metal and in many cases also dielectric material that has a low VSWR at a desired frequency. The actual shape of the antenna is otherwise quite arbitrary. Since the antenna is typically packaged into a device, it is near the case and other structures that make up the device. Nearby parts of the device can perturb the resonant frequency of the antenna, so we need the antenna to be tuned to the proper

frequency in its embedded environment. The device itself including its case and other components is part of the antenna, and this often needs to be accounted for by modeling and designing the antenna together with the device. The larger embedded environment is relevant as well, since the antenna must be designed to avoid detuning by the hand and body of the user.

27.2 Planar Antennas

At microwave frequencies, it is convenient and practical to integrate antennas into a device. This avoids having a large whip protruding from the device or an externally connected antenna. Since the electronics in a wireless device is assembled on a printed circuit board (PCB), it stands to reason that antennas can also be integrated on the PCB as well.

The broad class of antennas that can be constructed on a PCB or PCB-like stackup are refered to as planar antennas. A planar antenna is created by patterning the conductor (usually copper) on the top side, and in some cases the top and bottom layer, of a PCB in such a way that the structure resonates at the desired frequency and radiates with reasonable efficiency.

A wide variety of planar antennas are available. These include the patch antenna (see Sec. 27.3), printed monopole, bowtie, slots of various types including the tapered slot antenna (TSA) and Vivaldi antenna, leaky wave antennas, printed Yagis, loops and rings, and many others. Planar antennas are inherently low profile and compact. Planar antennas can be modified to produce conformal antennas that wrap around cylinders, fuselages, and other curved surfaces.

27.3 Patch Antennas

The most basic PCB antenna is the microstrip patch antenna, or simply patch antenna. The patch is a square or other shape of conductor patterned on the top surface of a PCB separated by

dielectric from a ground plane below the patch. The patch is fed by a micostrip trace or a conductive via to a trace on a lower board layer. The space between the patch and the ground plane forms a resonant cavity. Energy in the cavity leaks out of the gaps on the sides of the patch and radiates to the far field.

When designing a patch antenna, the starting point is a square patch with dimensions equal to a half wavelength at the operating frequency in the dielectric substrate that separates the patch and ground plane. EM modeling software can be used to optimize the design, but it normally is not worth overdesigning at the modeling stage. There is enough process variability in the substrate dielectric constant that the fabricated antenna tuning is different from the design. A few cycles of build-test-adjust can be used to tune the patch resonance to the desired frequency.

To estimate the gain of a patch antenna, we can approximate it as a pair of dipoles along two of the sides of the patch. With the ground plane below the patch creating images of the dipoles, the resulting gain is

$$G = 2 \times 2 \times 1.5 = 6$$

or about 8 dB.

There are many variants of the patch antenna. The basic patch has a fairly narrow bandwidth, but the shape can be adjusted or multiple resonant modes can be used to enhance the bandwidth. The basic patch antenna is linearly polarized, but circular polarization patch antennas are available. There are patches with air dielectric, slotted patches, patch arrays, and any of hundreds of other modifications as well. Patch antennas are normally used at frequencies above a few GHz, but occasionally larger patches are used for sub-GHz applications.

27.4 Planar Inverted F Antenna (PIFA) and Derivatives

In the quest to fit more antennas into small wireless devices, many variants of the patch antenna as well as monopole and dipole

antennas have been developed. One of these is the planar inverted F antenna (PIFA).

If we start with a patch antenna and cut the size in half, the antenna is no longer resonant at the same frequency, but adding a short at one edge of the patch makes it appear electrically to be a half wave in size and restores the resonance in a similar way to a monopole being half a dipole over a ground plane. Together with a probe feed, the shorting pin and the horizontal patch has the shape of an inverted F. The PIFA antenna is related to the inverted L in Sec. 20.8 and can be viewed as having a gamma match as used in Sec. 14.7.2 for lower frequency antennas.

The PIFA is low-profile, relatively easy to integrate into a compact device, and the location of the shorting pin and feed provide degrees of freedom that can be used to tune the antenna. PIFAs are often located along the sides of a wireless device.

27.5 Chip Antennas

To reduce the size of an antenna, it can be embedded in a dielectric material. This shrinks the wavelength of the fields in the material and the antenna size drops proportionately. A chip antenna consists of a metallic radiator embedded inside a block of ceramic or other type of dielectric with high permittivity and reasonably low loss. Chip antennas can be assembled onto a PCB in the same manufacturing process as other components.

27.6 Dielectric Resonator Antenna

A dielectric resonator antenna can be considered to be a type of chip antenna. The geometry and composition of the dielectric is designed to make the dielectric a resonant structure at the antenna design frequency. Rather than having an antenna that radiates inside a dielectric block, the dielectric itself becomes the radiating element. The size of a DRA is roughly the free space wavelength divided by the square root of the relative dielectric constant. Like a

chip antenna, the DRA can be fed with a microstrip transmission line or other coupling structure on a PCB. By adjusting the size and shape or exciting multiple resonant modes, the DRA antenna parameters can be adjusted or the bandwidth increases.

27.7 Modules and Integrated Antennas

At microwave frequencies, the small size and integratability of chip antennas and planar antennas has enabled radio communications transceivers to be modularized. For many short-range wireless applications, one can simply purchase a module with power and data connections, and the antenna is integrated into the module.

Deploying small, modular transceivers with embedded antennas is quite different from the large outdoor tower with ground structure needed for amateur radio in the HF band. The convenience of modular antenna systems comes at a cost, as the high propagation loss and inability to use ionosphere bounce means that the range of small wireless devices operating at microwave frequencies is generally limited to tens of meters or a few kilometers.

PART VI

Tuning, Troubleshooting, and Design Aids

CHAPTER 28
Antenna Tuners (ATUs)
CHAPTER 29
Antenna Modeling Software
CHAPTER 30
Testing and Troubleshooting

CHAPTER 28
Antenna Tuners (ATUs)

The task of an impedance-matching or antenna tuning unit (ATU) is to present an impedance that is a close match to the output impedance of the transmitter. This results in maximum power transfer from the power amplifier stage of the transmitter to the feedline and antenna. A properly designed ATU provides multiple benefits:

- It minimizes the standing wave ratio (SWR) that the transmitter or power amplifier sees. RF power amplifiers are not tolerant of high SWR and if operated anywhere near full power will almost certainly be damaged with a significantly mismatched load.
- It minimizes the occurrence of reduced output power triggered by power amplifier "fold-back" protection circuitry when a high VSWR is sensed.
- It facilitates operation of the transmitter final amplifier stage at maximum efficiency.
- It provides attenuation of out-of-band emissions, including harmonics generated in or amplified by the amplifier.
- If left in the line to the receiver during "key up" periods, it helps protect the receiver from overload caused by strong out-of-band signals.

If the ATU is located at the antenna end of the transmission line from a transmitter to the antenna, there are additional benefits:

- It minimizes the occurrence and amplitude of high-voltage standing waves along the line that can destroy

the transmission line and any associated switches, relays, or other components when high power is employed.

- It minimizes additional mismatch signal loss on long transmission lines relative to the loss that is present with line SWRs greater than 1.0:1.
- It minimizes radiating currents along the transmission line which produce fields that can distort the antenna radiation pattern and reduce directivity and gain.
- It can include additional protection for station electronics (receivers, transmitters, accessories, etc.) from lightning-induced events, serving as a first line of defense in preventing voltage or current surges from entering the radio room.

As the second list makes clear, it can be preferable to locate the ATU at the antenna. At low frequencies, low power levels, or for relatively short distances between the transmitter output and the antenna, an ATU at the transmitter can provide acceptable results. When an transmitter-end tuner is used, however, operating a poorly matched antenna and transmission line at full power is a bad practice even if the tuner achieves a low SWR.

Many transceivers include an internal automatic ATU. With short feedlines (such as mobile or portable installations) on HF, this configuration is compact and can be as efficient as a separate ATU. Third-party manufacturers offer external transmitter-end auto-tuners that will interface with the built-in tuner control circuitry of most transceiver models. These reduce transmit power temporarily while searching for lowered SWR by switching in different tuner network component values in sequence.

In some cases, matching devices can be included at both ends of the system transmission line. Fixed-tuned impedance-matching networks at the antenna feedpoint get the SWR on the system transmission line down to a manageable level, and a tunable ATU back at the transmitter and receiver ends of the main transmission

line completes the task of presenting the transmitter with a matched load. The impedance matching network at the antenna may be a simple balun or a distributed network, such as the transmission line transformers described in Chap. 6.

Occasionally, fully tunable ATUs are located at or near the antenna end of the system transmission line. These are often capable of being tuned remotely, or the intended operating range is over such a small percentage bandwidth that adjustment of the ATU is required infrequently or never. An example of such a configuration is found in Fig. 10.4. An open-wire line (which is the best choice when the possibility of high SWR on the line exists) connects the center of a multiband dipole to a tunable ATU directly beneath it. The ATU is adjusted to provide minimum SWR on the coaxial transmission line that comes from the station equipment. It may also incorporate RF chokes to drain static buildup from the antenna and OWL, or remotely actuated mechanical contactors to directly ground the OWL and dipole when not in use. An even more compelling case for a remote ATU is the bobtail curtain of Chap. 12. Since the feedpoint for this array is at a high impedance point, the SWR on any common coaxial cable back to the transmitter would put the cable at risk of destruction when substantial power levels are used. An attempt to run open-wire line over that distance could materially degrade the antenna pattern. Locating a tunable ATU at the base of the antenna, as shown in Fig. 12.11, is a better approach.

28.1 ATU Circuit Configurations

At any given operating frequency, the antenna feedpoint impedance,
Z_{ant}
, can be represented as

$$Z_{\text{ant}} = R_{\text{ant}} + jX_{\text{ant}} \quad (28.1)$$

where

X_{ant}
can be either positive or negative. If we could count on
R_{ant}
always equaling the system transmission line impedance, we could use a single variable capacitor in series to tune out a positive
X_{ant}
and a single variable inductor in series to tune out a negative
X_{ant}
. As the operating frequency changes, however, the real resistance
R_{ant}
also varies.
R_{ant}
and
X_{ant}
are fairly complicated functions of frequency. Thus, the task of the successful antenna matching unit or ATU is to be flexible enough to present the system transmission line with an apparent antenna feedpoint impedance of
$Z_{\text{ant}} = Z_0$
over a reasonable range of operating frequencies and a practical range of complex antenna impedances. In the process, the ATU should dissipate as little of the transmitter RF power output in internal losses as possible. There are several network topologies that can do this with varying degrees of success.

28.1.1 L-Section Network

The L-section network is one of the most used antenna matching networks in existence. One form of L-section network is shown in Fig. 28.1A. Values for *L* and *C* can be found that will allow this circuit to match all possible load impedances having
R_1
, the resistive part of the source impedance, less than
R_2
, the resistive part of the load impedance. The circuit can also match some load impedances for the range

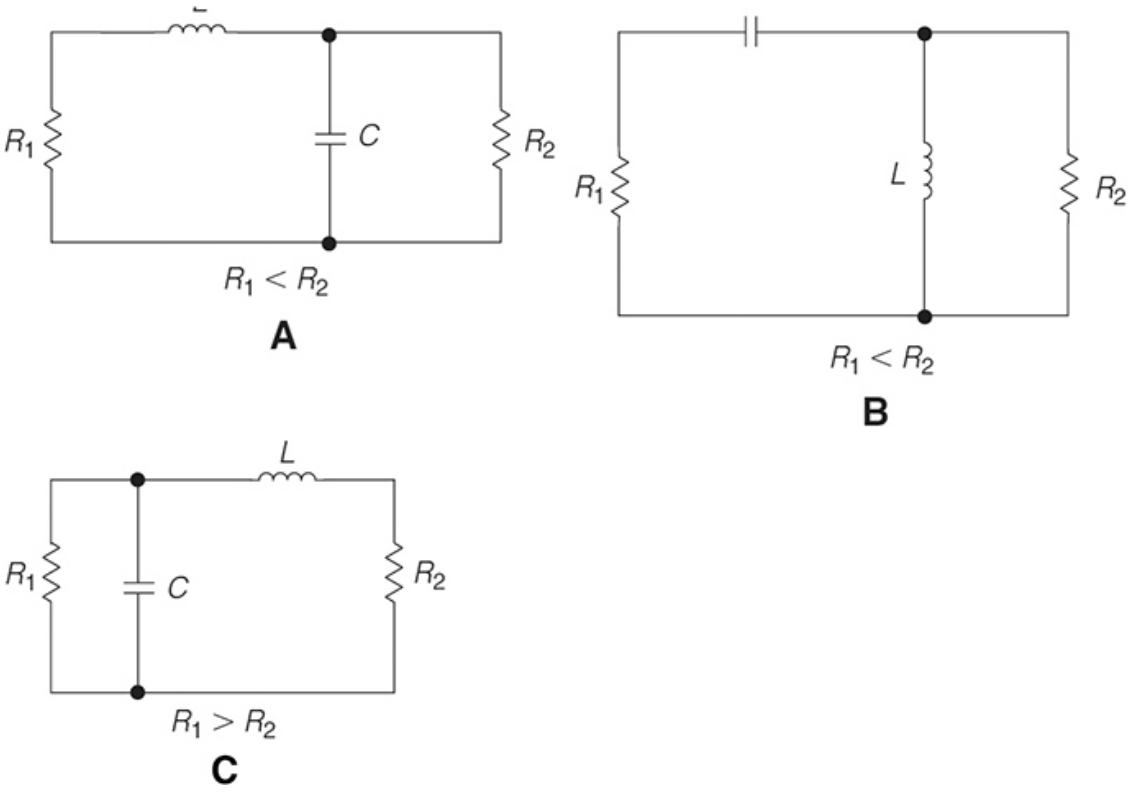
$R_1 > R_2$
, depending on the magnitude of
X_2
, the reactive part of the series-representation load impedance. For the special case when both the source and the load impedances are purely resistive, the design equations are:

$$Q = \sqrt{\frac{R_2}{R_1} - 1} \quad (28.2)$$

$$X_L = QR_1 \quad (28.3)$$

$$X_C = -\frac{R_2}{Q} \quad (28.4)$$

where
$X_C = -1/(2\pi fC)$
and
$X_L = 2\pi fL$
.



F IGURE 28.1 (A) L-section network. (B) Reverse L section. (C) Inverted L-section network.

The L-section network of Fig. 28.1B differs from the previous circuit in that the coil and capacitor locations are swapped. As before, values for *L* and *C* can be found that will allow this circuit to match all possible load impedances having $R_1 < R_2$, and some load impedances for the range $R_1 > R_2$, depending on the magnitude of X_2. For the special case when both the source and the load impedances are purely resistive, the design equations are:

$$Q = \sqrt{\frac{R_2}{R_1} - 1} \quad (28.5)$$

$$X_L = \frac{R_2}{Q} \quad (28.6)$$

$$X_C = -R_1 Q \quad (28.7)$$

Figure 28.1C is similar to Fig. 28.1A with the exception that the capacitor is at the input rather than the output of the network. Values for *L* and *C* can be found that will allow this topology to match all possible load impedances having $R_1 > R_2$, and some load impedances for the range $R_1 < R_2$. The equations governing this network are:

$$Q = \sqrt{\frac{R_1}{R_2} - 1} \quad (28.8)$$

$$X_L = R_2 Q \quad (28.9)$$

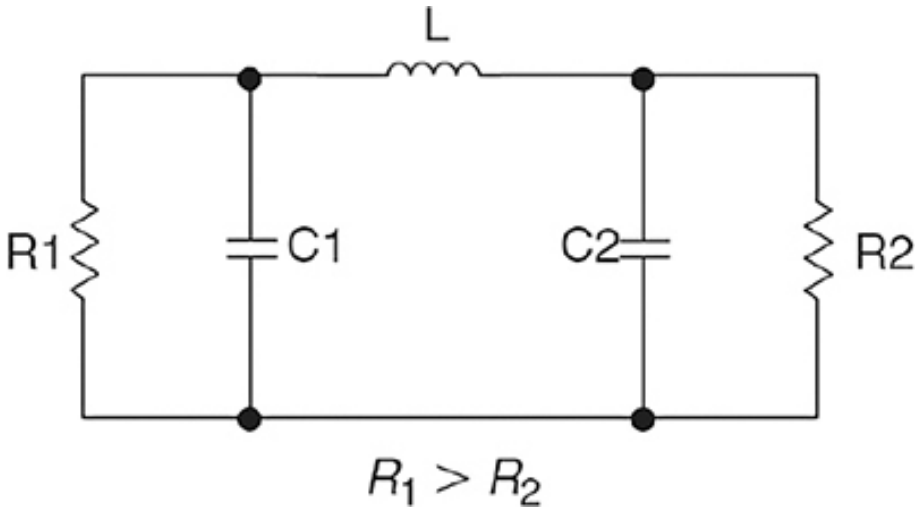
$$X_C = -\frac{R_1}{Q} \quad (28.10)$$

Because of the limitations on the range of load impedances it can match when either the source or the load has a reactive component, an L network located at the transmitter end of a feedline should be configurable to accommodate the different topologies shown in Fig. 28.1. Alternatively, an L network located at the antenna feedpoint can be a fixed topology once the nature of the feedpoint impedance of the antenna is known.

28.1.2 Π Networks

The pi network (or pi-section network) shown in Fig. 28.2 is used to match a high source impedance to a low load impedance. These circuits are typically used in RF power amplifiers with high output impedances of a few thousand ohms when matching to much lower system transmission line impedances. The name of the circuit

comes from its resemblance to the Greek letter pi (π). The equations for the pi network are:



F IGURE 28.2 Pi network.

$$R_1 > R_2 \text{ and } 5 < Q < 15$$

$$Q > \sqrt{\frac{R_1}{R_2} - 1} \quad (28.11)$$

$$X_{C2} = \frac{R_2}{\sqrt{\left(\frac{R_2}{R_1}\right)(1+Q^2) - 1}} \quad (28.12)$$

$$X_{C1} = \frac{R_1}{Q} \quad (28.13)$$

$$X_L = \frac{R_1\{(Q + (R_2/X_{C2})\}}{Q^2 + 1} \quad (28.14)$$

Of course, if $R_1 < R_2$, the pi network can be flipped left to right; that is, R_1

can become the load and
R_2
the source. Because the pi network typically has a higher *Q* than an L-section network, its bandwidth will generally be narrower. If bandwidth is an important issue, a pair of cascaded L sections may be a better solution to a particular matching task.

28.1.3 Split-Capacitor Network

The split-capacitor network shown in Fig. 28.3 is used to transform a source impedance that is less than the load impedance. In addition to matching antennas, this circuit is also used for interstage impedance matching inside communications equipment. The equations for design are:

$$R_1 < R_2$$

$$Q > \sqrt{\frac{R_2}{R_1} - 1} \quad (28.15)$$

$$X_L = \frac{R_2}{Q} \quad (28.16)$$

$$X_{C\,\mathrm{IB}} = \sqrt{\frac{R_1 (Q_2 + 1)^2}{R_2} - 1} \quad (28.17)$$

$$X_{C1\mathrm{A}} = \frac{R_2 Q}{Q^2 + 1} \left(1 - \frac{R_1}{Q X_{C1\mathrm{B}}} \right) \quad (28.18)$$

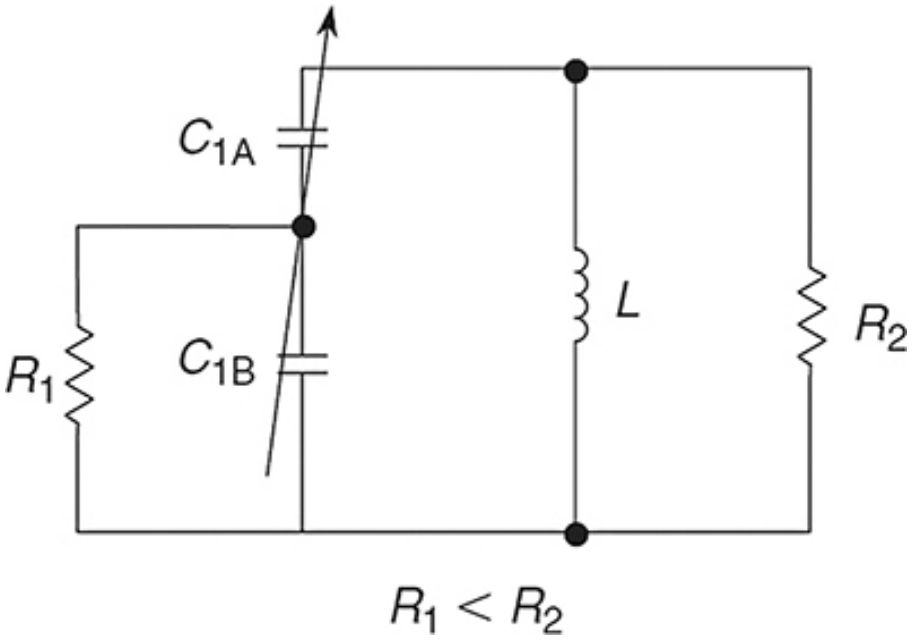


FIGURE 28.3 Split-capacitor network.

28.1.4 Transmatch Circuit

One version of the transmatch is shown in Fig. 28.4. This circuit is a combination of the split-capacitor network and an output tuning capacitor (C_2). For the HF bands, the capacitors are on the order of 150 pF per section for C_1, and 250 pF for C_2. The tapped or roller inductor should have a maximum value of 28 μH. A popular use of the transmatch is as a coax-to-coax impedance matcher.

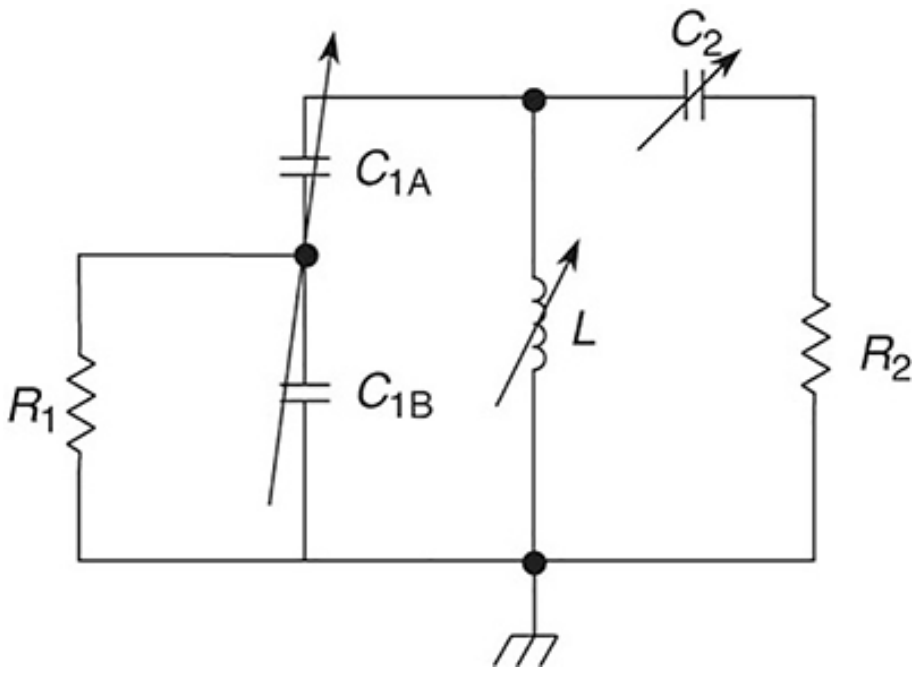


FIGURE 28.4 Split-capacitor transmatch network.

The long-popular E. F. Johnson *Matchbox* is a balanced variant of this circuit, designed primarily for use with balanced feedlines such as open-wire line and twin-lead. Coupling to the transmitter or receiver is accomplished via an inductive link (an isolated secondary winding on *L*), thus eliminating the connection of R_1 across half of C_1. C_2 becomes a split capacitor connected across the full length of the tapped inductor *L*, and the junctions of the two halves of C_1 and C_2 are tied together and grounded to the chassis. Each half of the new C_2 is split again, and the two sides of the balanced transmission line are connected at the midpoint

C_2
on each side of the ground. The center conductor of an SO-239 chassis-mounted coaxial receptacle for feeding unbalanced transmission lines is hard-wired to the midpoint of one side of
C_2
.

A common form of transmatch circuit is the tee network shown in Fig. 28.5. Like the reverse L-network of Fig. 28.1B, it is a high-pass filter and provides no transmitter harmonic attenuation.

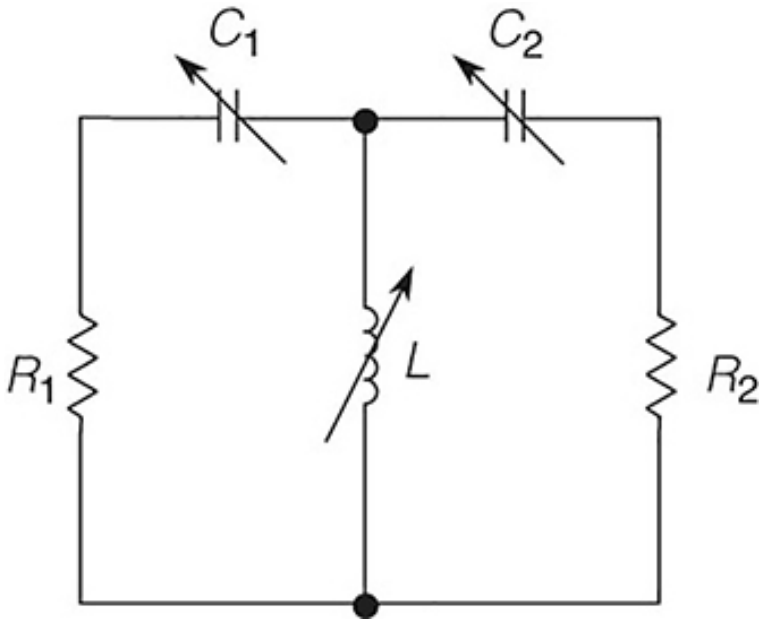


FIGURE 28.5 Tee-network transmatch.

An alternative network, called the SPC transmatch, is shown in Fig. 28.6. This version of the circuit offers some harmonic attenuation.

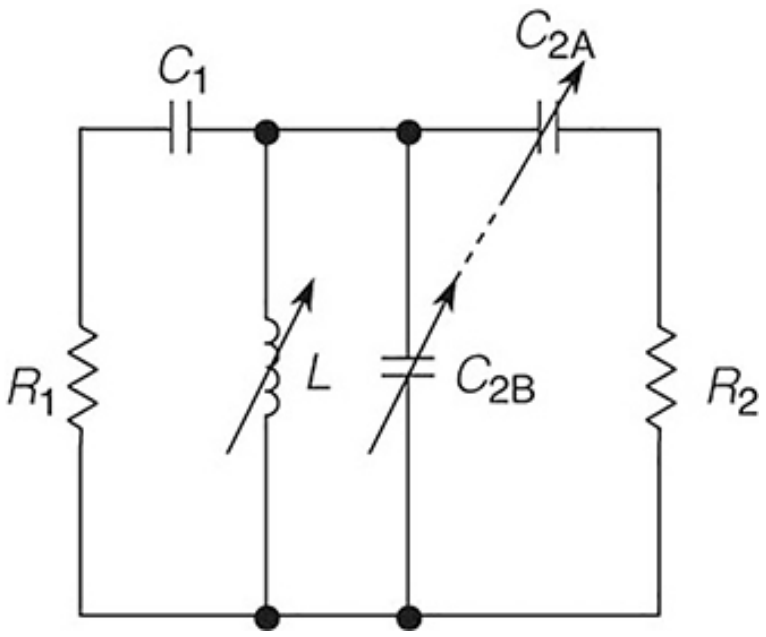


FIGURE 28.6 Improved transmatch offers harmonic attenuation.

Lumped-component matching networks (or ATUs) for MF and HF are relatively easy to build. Inductors and variable capacitors capable of withstanding the voltages and currents associated with transmitter power levels can be expensive. When building your own ATU, to avoid reductions in efficiency, leave space between a metallic enclosure and the internal components.

Commercially produced transmitters, transceivers, and amplifiers are shielded and filtered against the unwanted emission of harmonics and other spurious products of the signal generation and amplification process. Similarly, shielding of the ATU circuitry is a best practice.

28.2 Baluns

A balun is a transformer that matches a balanced load (such as a horizontal dipole or Yagi antenna) to an unbalanced source (such as a transmitter output or a coaxial feedline). The baluns discussed in this section are lumped-circuit implementations typically wound on

ferrite cores. Baluns can also be formed from sections of a transmission line properly interconnected (see Chap. 6).

28.2.1 Toroid Impedance-Matching Transformers

The toroidal transformer provides a broadband match between an antenna and transmission line or between a transmission line and transmitter or receiver. Other matching methods can be frequency-sensitive and must be readjusted when the operating frequency is changed by even a small percentage. Although this problem is of no great concern to fixed-frequency radio stations, it is an issue for stations that operate on a variety of frequencies.

Figure 28.7 shows a trifilar transformer that provides a 1:1 impedance ratio and transforms an unbalanced transmission line (e.g., coaxial cable) into a balanced signal required to feed a dipole antenna. This helps to balance the feed currents in the two halves of the antenna and makes it possible to obtain a more symmetrical figure eight dipole radiation pattern in the horizontal plane. Many station owners make it standard practice to use a balun at the antenna feedpoint when using a coaxial (unbalanced) transmission line, even if the feedpoint impedance is close to the

Z_0

of the cable.

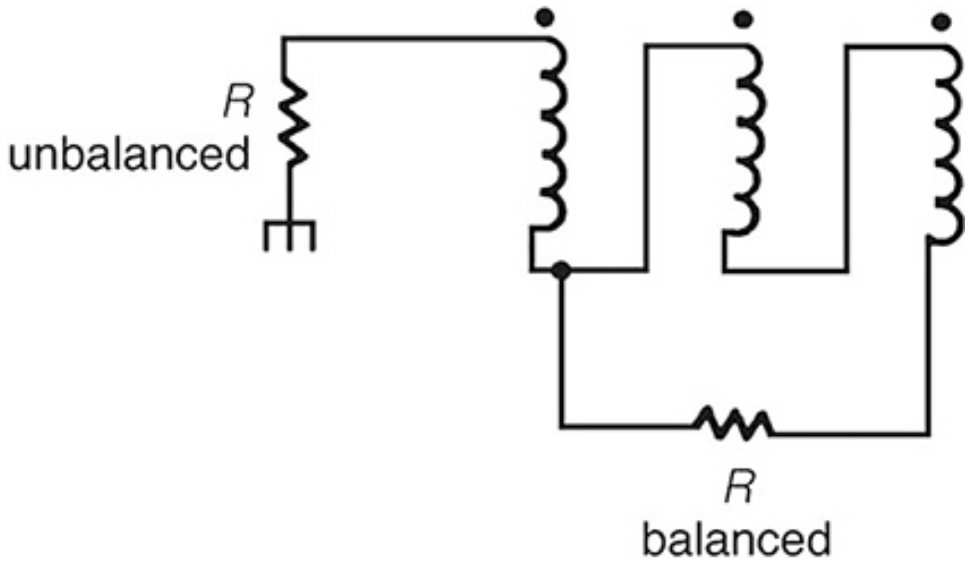


FIGURE 28.7 Balun 1:1.

The balun shown in Fig. 28.8 is designed to provide an unbalanced to balanced transformation in conjunction with a 4:1 impedance transformation.

L_1

and

L_2

have equal turns. A 300-Ω folded dipole feedpoint impedance will be transformed to 75-Ω unbalanced. This type of balun is often included inside ATUs.

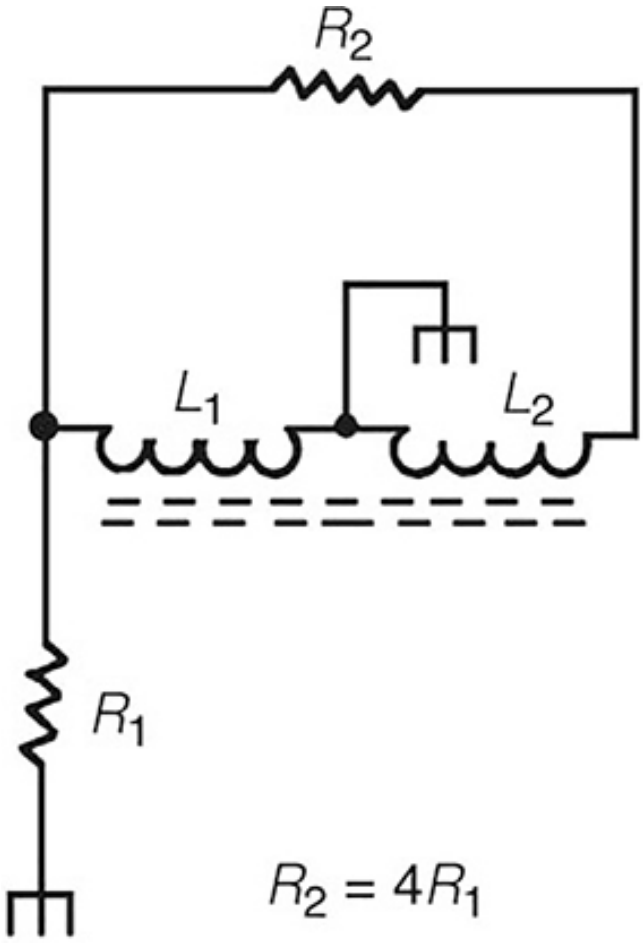


FIGURE 28.8 Balun 1:4.

Another multiple-impedance transformer is shown in Fig. 28.9. The operator can select impedance transformation ratios of 1.5:1, 4:1, 9:1 or 16:1.

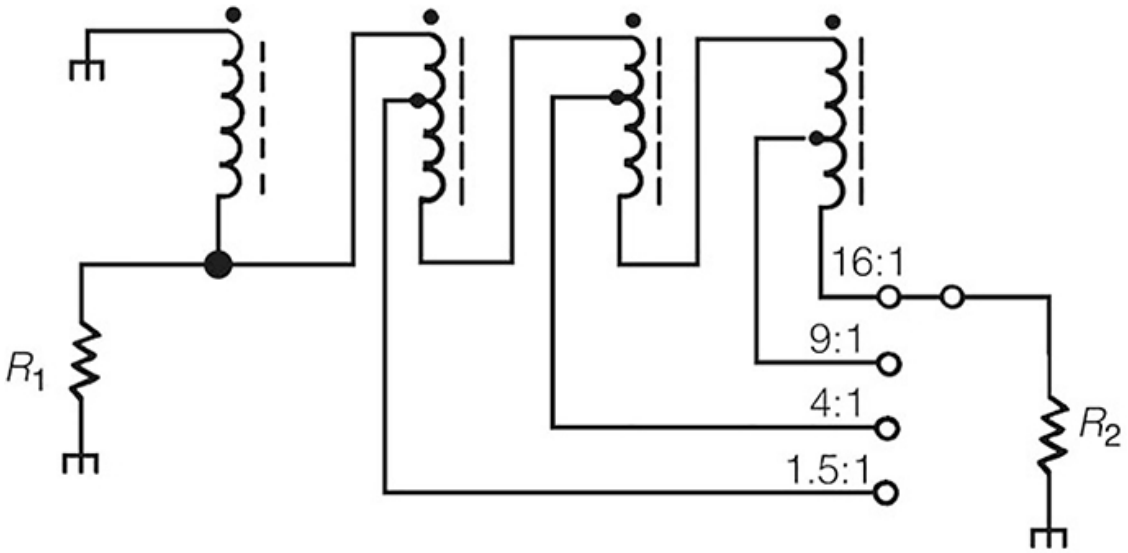


FIGURE 28.9 Tapped multiple-impedance balun.

Figures 28.8 and 28.9 are examples of what can be accomplished using a toroidal core of the right material and power-handling capability in conjunction with an autotransformer winding, where both input and output loads are referenced to a common ground. If ground isolation is important, separate windings can be used in a configuration analogous to a link-coupled coil.

28.3 Ferrite Core Inductors

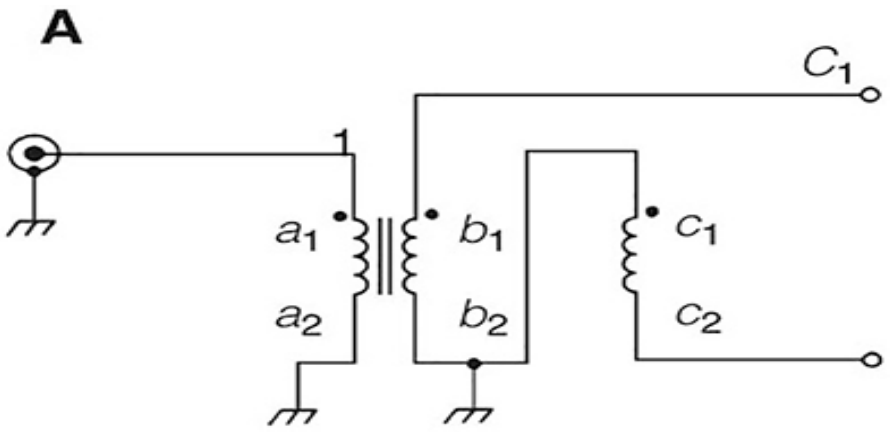
Ferrite refers to any of several examples of a class of materials that behave similarly to iron compounds and are used in radio equipment as the cores for inductors and transformers. Ferrites are available with relative permeability up to 5000 or more. As the frequency becomes higher, loss in a ferrite core increases, so at microwave frequencies high quality materials are used or the ferrite can be omitted and an air core inductor used.

28.3.1 Toroidal Transformers

Windings on a toroid are generally wound simultaneously in a multifilar manner. Multiple wires of equal length are slightly twisted together first, then wrapped around and through the donut form as

if they were a single winding. The preferred number of wires in parallel is a function of the required step-up or step-down ratio and whether an autotransformer or isolated primary and secondary windings are required.

Figure 28.10A is the schematic of a 4:1 balun transformer with separate primary and secondary windings. The dot at the top of each winding indicates the polarity or phase of the winding. For proper operation as an autoformer, all three wires must wrap around the donut in the same direction. This particular circuit is said to have a trifilar wound core.


A
C_1
1
a_1
b_1
c_1
a_2
b_2
c_2

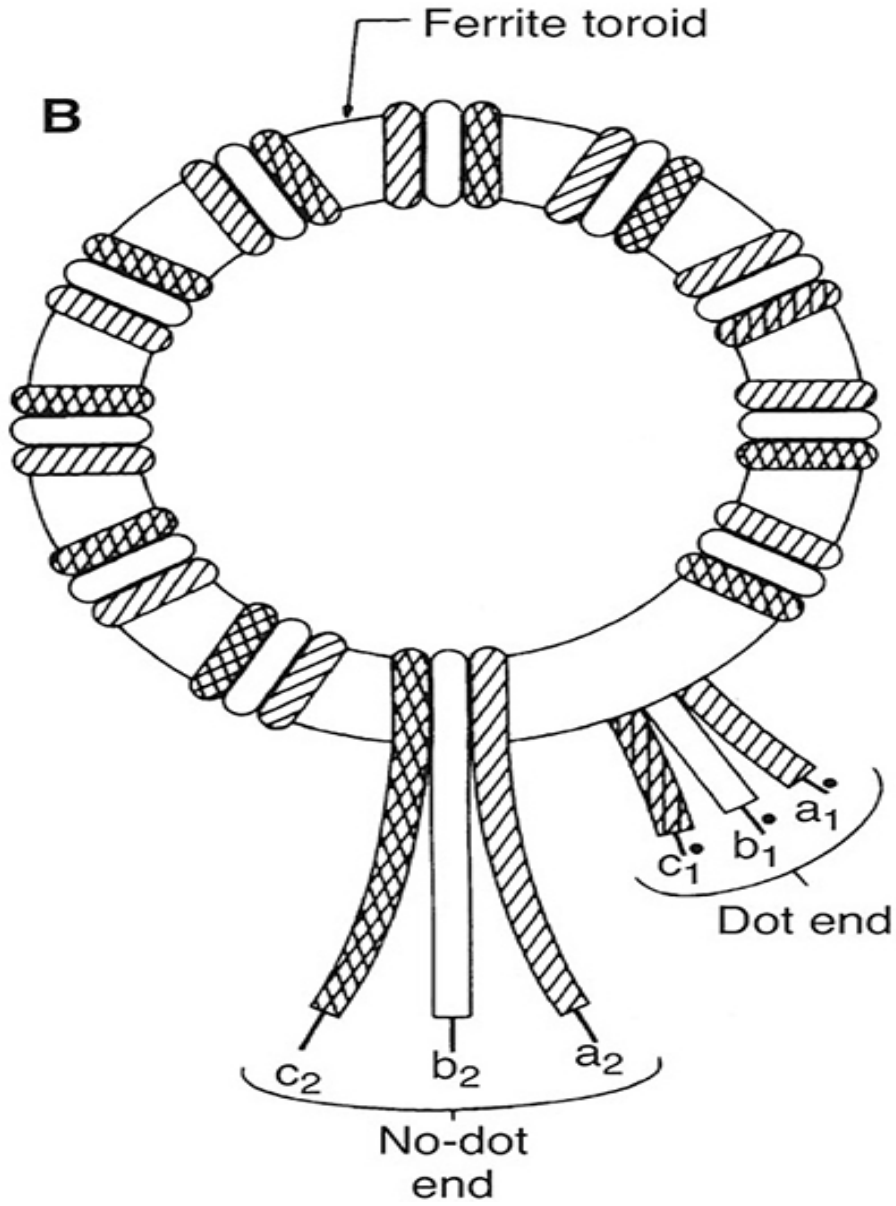

Ferrite toroid
B
a_1
b_1
c_1
Dot end
c_2
b_2
a_2
No-dot
end

FIGURE 28.10 Broadband RF transformer.

Figure 28.10B is a detail of the trifilar winding. Most small construction projects (e.g., Beverage antenna transformers) use #26, #28, or #30 enameled wire to wind coils. For higher-power transmitting antenna baluns, use #16, #14, #12, or #10 wire. Enamel and Teflon are common insulation layers. To help identify wires of these diameters, label their ends with adhesive labels.

Figure 28.11 shows two accepted methods for winding a multifilar coil on a toroidal core. Figure 28.11A is a toroid wound with the method shown in Fig. 28.10B. The wires are laid down parallel to each other, as shown previously. The method of Fig. 28.11B uses twisted wires. Prior to winding the toroid, the three wires are "chucked up" in a drill and twisted together before being wound on the core. With one end of the three wires secured in the drill chuck, anchor the far ends using a vise. Run the drill at slow speed and allow the wires to twist together until the desired pitch is achieved.

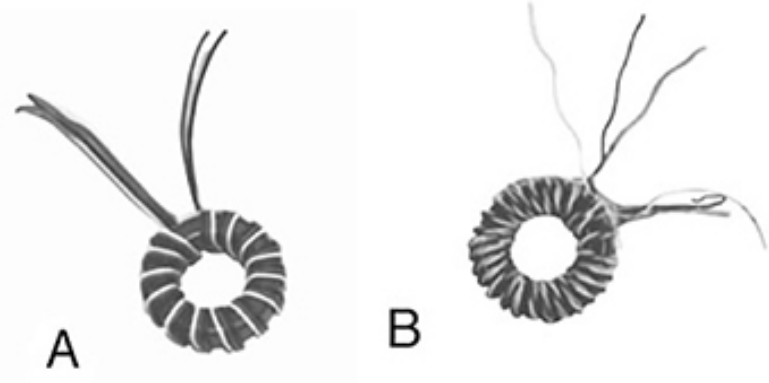


FIGURE 28.11 Winding a toroidal transformer. (A) Parallel wound. (B) Twist wound.

To minimize the chance of a chipped core breaking through the wire insulation, wrap the core with a slightly overlapped layer of fiberglass packing tape before starting to wind. To wind the bundle of wires on the toroid, pass the bundle through the donut hole until

the toroid is about in the middle of the length of wire. Loop the wire over and around the outside surface of the toroid and pass it through the hole again. Repeat this process until the correct number of turns is wound onto the core. Press the wire against the toroid form and keep it taut as you wind the coils. To prevent the coil from unwinding, secure the ends of the wires with an adhesive.

28.3.2 Mounting Toroid Cores

To mount a properly wound toroidal inductor or transformer, if the wire is heavy enough, use the wire connections to the circuit board or terminal strip to support the component, especially if the connections are directly above or below the toroid. Smaller toroids can be laid flat on the circuit board and cemented in place with silicone seal or rubber cement. Drill a hole in the wiring board and use a screw and nut to secure the toroid. Do not use metallic hardware for mounting the toroid! Metallic fasteners will alter the inductance of the component and possibly render it unusable. Use nylon hardware for mounting the toroidal component.

28.3.3 How Many Turns?

Three factors must be taken into consideration when making toroid transformers or inductors: toroid size, core material, and number of turns of wire. Toroid size is a function of power-handling capability and handling/installation convenience. The ore material depends on the frequency range of the circuit and the application (balun, common mode choke, etc.).

A wide tolerance exists on amateur-grade ferrite cores. Accuracy is relatively unimportant for baluns but can be a concern when making fixed inductors for a tuned circuit.

Equation (28.19) provides a rough rule of thumb for the turns count. Based on the availability of a parameter called the A_L factor, the number of turns needed to obtain a specified inductance is given by

$$N = 100\sqrt{\frac{L}{A_L}} \quad (28.19)$$

where *L* is the inductance in microhenrys and
A_L
is the core factor, in microhenrys per 100 turns.

Example 28.1 Calculate the number of turns required to make a 5-*μ*H inductor on a T-50-6 core. The
A_L
factor for the core is 40.

Solution

$$\begin{aligned} N &= 100\sqrt{\frac{A}{A_L}} \\ &= 35 \end{aligned}$$

28.3.4 Ferrite Rods

Another form of ferrite core is the rod, shown in Fig. 28.12. Primary and secondary windings are wound in a bifilar manner over the ferrite rod.

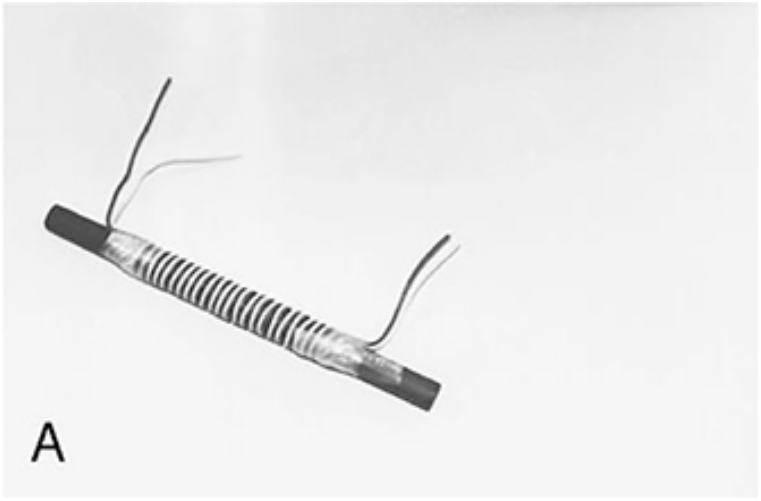


FIGURE 28.12 Ferrite rod inductor construction.

Ferrite rods are used in receiving antennas when high inductance in a small space is required. The ferrite rod antenna (or loopstick) is popular in small receivers for MF and below, and in portable radio direction-finding equipment. Some amateurs use sharp nulls of loopstick receiving antennas to null out interfering signals on the crowded HF bands. You would normally not want to use a small loopstick rod in a transmitting application.

Ferrite rods are light enough to be suspended from their own wires or to be glued or cemented to a panel or printed circuit board inside the receiver. In place of the simple nylon screws that hold toroids in place, use insulating cable clamps to secure the ends of the rod to the board.

CHAPTER 29
Antenna Modeling Software

Many types of antennas can be designed using formulas or trial and error, but more complex antennas and arrays are designed using numerical algorithms. Available tools range from online calculators and free or low-cost software to expensive and powerful software packages. It is much easier to fine-tune a design by computer than to make repeated antenna measurements in a test range or climb a tower to make an adjustment.

29.1 Antenna Analysis and Synthesis

The basic task of a numerical model is to start with a given antenna design, in some cases, including mounts and other structures around the antenna, and find the electromagnetic fields radiated by the antenna. From the radiated fields, useful parameters are computed in post-processing, including directivity, gain, radiation efficiency, radiation pattern, beamwidth, and other parameters. This is referred to as antenna analysis.

The reverse problem, going from desired antenna characteristics to an antenna design, is called antenna synthesis. Because there are many designs that correspond to a limited set of desired antenna characteristics, antenna synthesis is a more open-ended problem than antenna analysis, and depending on the complexity of the design, it can take a lot of computation time.

Antenna synthesis is typically done by starting with a parametric model for an antenna and adjusting the geometry until the antenna characteristics meet the desired performance requirements. An optimization algorithm is used to create one or more sets of geometrical parameters for the antenna. The candidate designs are analyzed using the numerical model, and a cost function is

computed based on how well the design performs. This process is repeated until the cost function is maximized or minimized and the antenna design converges to one that achieves the performance criteria.

In simple cases where a conductor length, diameter, or one or two geometrical parameters of the antenna are adjusted, this process can be done using hand tuning. For an antenna with a complex shape, more sophisticated optimization techniques are needed. These kinds of algorithms are typically included in expensive antenna modeling software packages.

29.2 Types of Numerical Methods

Many different numerical methods are used for antenna analysis and design, but the most common are the method of moments (MoM), the finite different time domain (FDTD) method, and the finite element method (FEM). For antennas made only of good conductors, the method of moments has the advantage that only the conductor surfaces need to be represented as a grid or mesh. The other algorithms generally require more sample points over a volume around the antenna, which increases the memory requirement and run time of the algorithm.

29.3 Finite Difference and Finite Element Methods

The physical laws that govern electromagnetic fields can be expressed using a set of differential equations called Maxwell's equations. The derivatives in these equations can be approximated as difference equations. The difference equations can be solved by iterating in time or by inverting a matrix to find the fields radiated by an antenna.

There are two basic frameworks for creating the necessary difference equations. One is based on a rectangular grid and a mathematical construction called a Yee cell. The Yee cell connects

samples of the electric and magnetic fields at different points in space and time. Using equations from the Yee cell, the electric and magnetic fields around an antenna can be found at a future time step from the fields at past time steps. By iterating the process over time, we can model the fields as they propagate away from an antenna. This is the finite difference time domain (FDTD) method.

The other major framework is the finite element method (FEM). In the finite element method, the antenna is divided into tetrahedra or other 3D geometrical shapes called mesh elements, and the electric and magnetic fields on each element are represented using samples at the corners of the element. The electromagnetic fields are found by solving a matrix equation that minimizes the stored energy of the fields. This is based on the physics concept of the variational principle, in which the behavior of a physical system in nature tends to minimize some quantity, such as stored energy or action.

Because the finite element method is based on tetrahedral shapes, it is more inherently geometrically flexible than the basic FDTD method on a rectangular grid. Mesh elements can be conformal to a curved antenna surface, and the mesh can be refined with smaller elements near sharp or small features of the antenna, which can make numerical results more accurate. Most commonly, the FEM method is used to find the fields radiated by an antenna for a time-harmonic source with a single frequency. To find the response over frequency, the FEM algorithm must be rerun multiple times with different source frequencies. The FDTD method, on the other hand, can find the response of an antenna with a single pulsed source containing many frequency components. Variants of the FDTD method are available that allow for conformal grids and mesh refinement as well.

In addition to FDTD and FEM, there are many other numerical methods that have been developed for use in electromagnetic modeling and antenna design. These include the finite integration technique and the transmission line matrix method. While there are advantages and disadvantages under the hood, from the perspective of the user the end result is essentially the same—a

solution for the electric and magnetic fields around an antenna based on the shape and geometrical structure of the antenna and the frequency or time dependence of the source that is used to excite the antenna.

Hundreds of software packages ranging from freeware to commercial design environments based on FDTD and FEM and similar types of numerical methods are available. For these packages, the learning curve required to generate a 3D antenna model and simulate it can be fairly steep. Early users often follow tutorials to develop a working antenna model for a simple case and then make changes until it meets their needs. Some software packages include a library of predefined antenna types that can be used as building blocks or tuned to meet a specific design requirement.

29.4 Method of Moments

The method of moments is based on an integral relationship between an electrical current and the field it radiates. The relationship uses what is called a Green's function. To transform the integral relationship into discrete equations that can be solved numerically, an antenna is broken into many short segments. By solving a matrix equation, the current in each segment can be found that satisfies the electromagnetic boundary conditions at the antenna's conducting surface. The fields radiated by the antenna can be calculated by summing the contributions from the currents on all segments of the antenna. From the radiated fields, the gain, radiation pattern, beamwidth, and other antenna parameters can be calculated. If needed we can also find the input impedance, feedpoint voltage and current, voltage standing wave ratio (VSWR), the currents on the antenna, or near fields around the antenna.

The earliest method of moments codes were based on a thin-wire formulation. In the thin-wire formulation, the current on a wire is approximated as constant over a cylinder or concentrated on a thin filament. The volumetric current flow through the wire is

represented by one-dimensional samples of the total current through each segment of the wire. The thin-wire formulation minimizes the number of current samples used to represent an antenna and is fast and convenient. The thin-wire formulation has a numerical instability which must be dealt with by careful choices for the aspect ratio of the segments of a conducting wire. To represent two-dimensional and three-dimensional structures, a wire mesh must be used.

Surface and volume formulations of the method of moments are also available. In the surface method of moments, the antenna is approximated as a perfect conductor and the current is confined to an infinitely thin, two-dimensional skin on the surface of the antenna. The computational mesh is formed from triangles or in some cases quadrilaterals or other shapes that tile the surface of the antenna. If the antenna is modeled as an imperfect conductor or dielectric materials are included, the volume method of moments formulation can be used.

29.5 Numerical Electromagnetics Computation (NEC)

One of the oldest method of moments codes is the Numerical Electromagnetic Computation (NEC) package and its many derivatives (NEC-2, NEC-4, miniNEC, EZ-NEC, NEC-Win Plus, cocoaNEC, 4nec2, etc.). Some versions of NEC are freely available, whereas others require a paid license. Early versions of NEC lacked a graphical user interface and the antenna model was created using text scripts, but more recent versions are available with a GUI.

NEC-based programs model wire antennas using the standard three-dimensional cartesian coordinate system, as shown in Fig. 29.1. The z axis represents the vertical direction and Φ and θ are the azimuth and elevation angles, respectively. The antenna to be examined is placed in this space by identifying the x, y, z coordinates of both ends of each straight wire element that is part of the antenna, along with similar representation of any nearby

conducting surfaces likely to affect the performance of the antenna being modeled.

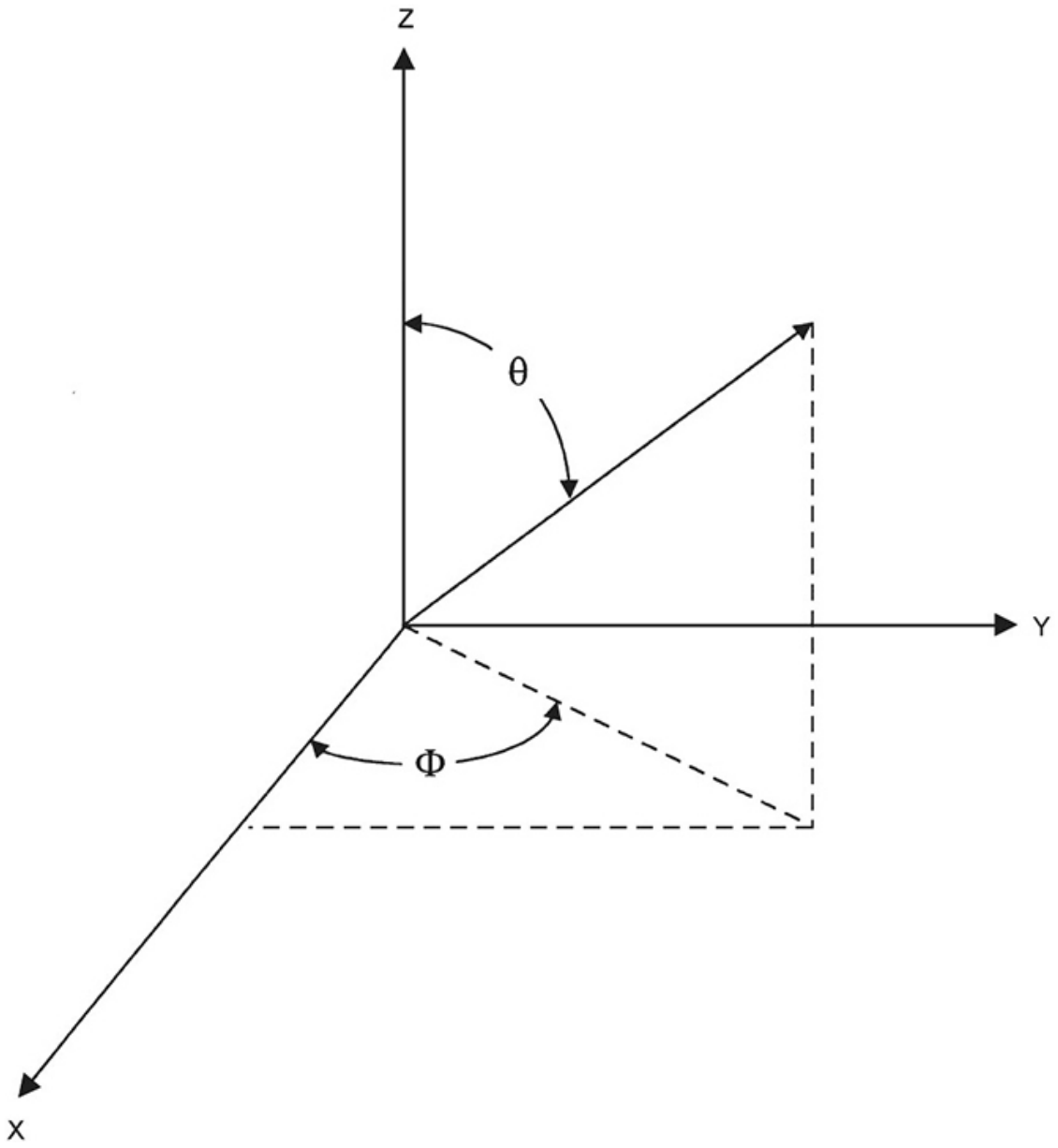


FIGURE 29.1 *x*, *y*, *z* coordinate system for *NecWin* modeling software.

If the antenna is located in free space, the orientation of the antenna with respect to the *x*, *y*, *z* coordinate system is not particularly important, but if the antenna is near a conducting or dielectric medium (such as ground, seawater, etc.) having

characteristics different from those of free space, the orientation of the antenna properly with respect to the z-axis becomes important since some antenna modeling packages assume a flat surface (earth, a car top, the ocean, etc.) spanning the xy-plane that corresponds to

$$z = 0$$

.

Figure 29.2 depicts a single-wire antenna in free space. It is centered on the origin (0,0,0) and extends from

$$x = x_1$$

(expressed as a positive number) to

$$x = x_2$$

(expressed as a negative number). If the antenna were above a flat earth or ground surface, we would put it at a positive *z* value instead of at

$$z = 0$$

.

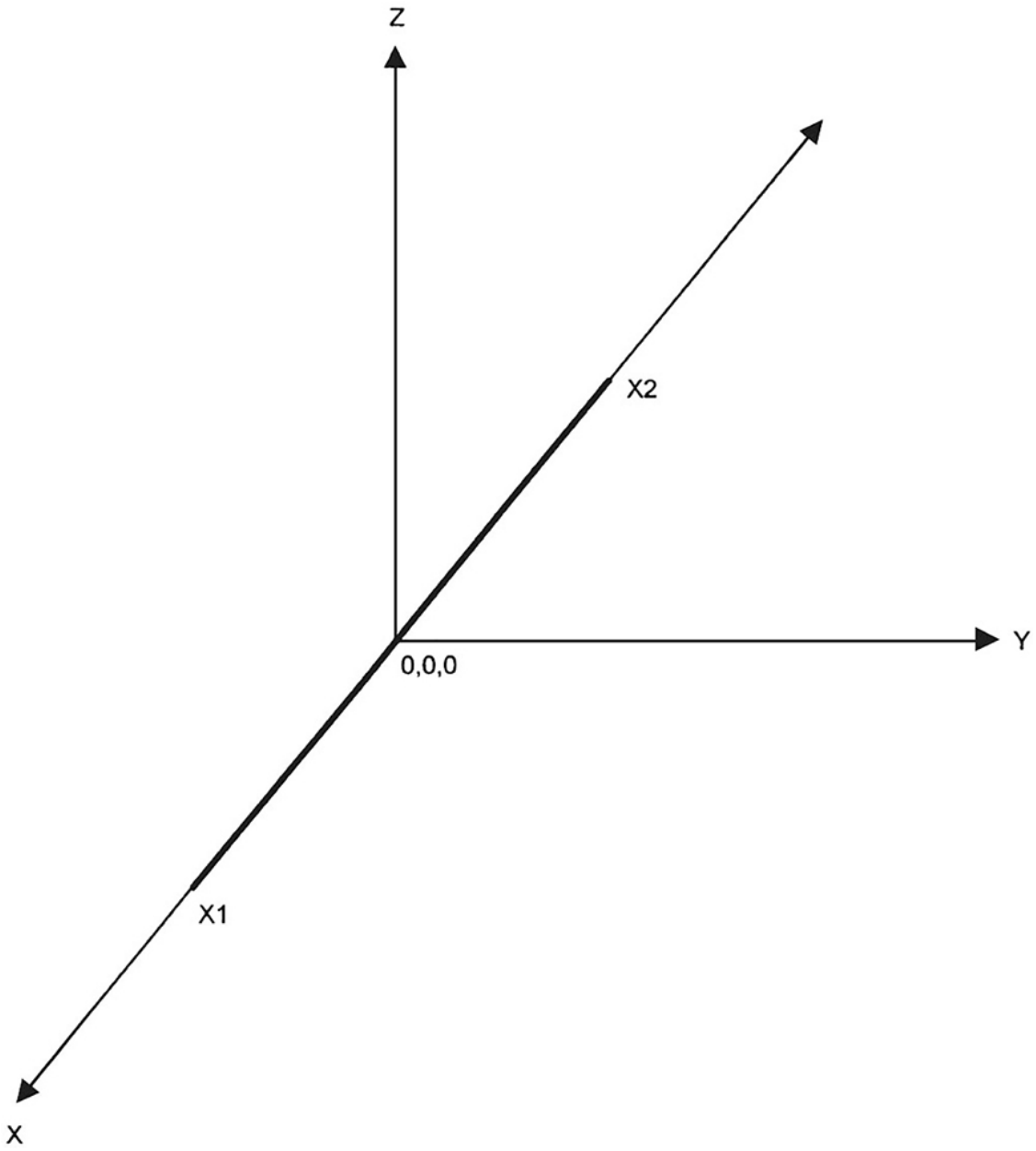


FIGURE 29.2 Example of a single wire laid out on *x*, *y*, *z* coordinate system.

We need to specify where the source of RF power that excites the antenna is located. This can be in the middle of the antenna for a center-fed dipole or at the end in the case of an end-fed dipole or long wire antenna. In some MoM implementations, the user needs to specify the number of segments that are used to model the antenna. With practice, the user can learn how to quickly specify

wires and their positions in model space, select an appropriate number of segments for each wire, and position sources within segments. When modeling antenna elements, such as dipoles, that are fed in the center, use an odd number of segments for the wire. This allows placement of the source or feedpoint in the exact center of the middle segment.

29.5.1 A Typical NEC-2 GUI

Figure 29.3 shows the opening screen in EZNEC+ ver 5.0 for a typical antenna project (in this case, an 80-m bent dipole). Visible in this window are file name, frequency, free-space wavelength (in meters), number of wires and segments, number of sources, number of loads, transmission lines, ground type, wire loss, the units employed (meters, feet, etc.), plot type (azimuth, elevation), elevation angle, step size, a reference level, and an alternate Z_0 for the SWR plot (to conveniently handle both 50- and 75-Ω system impedances).

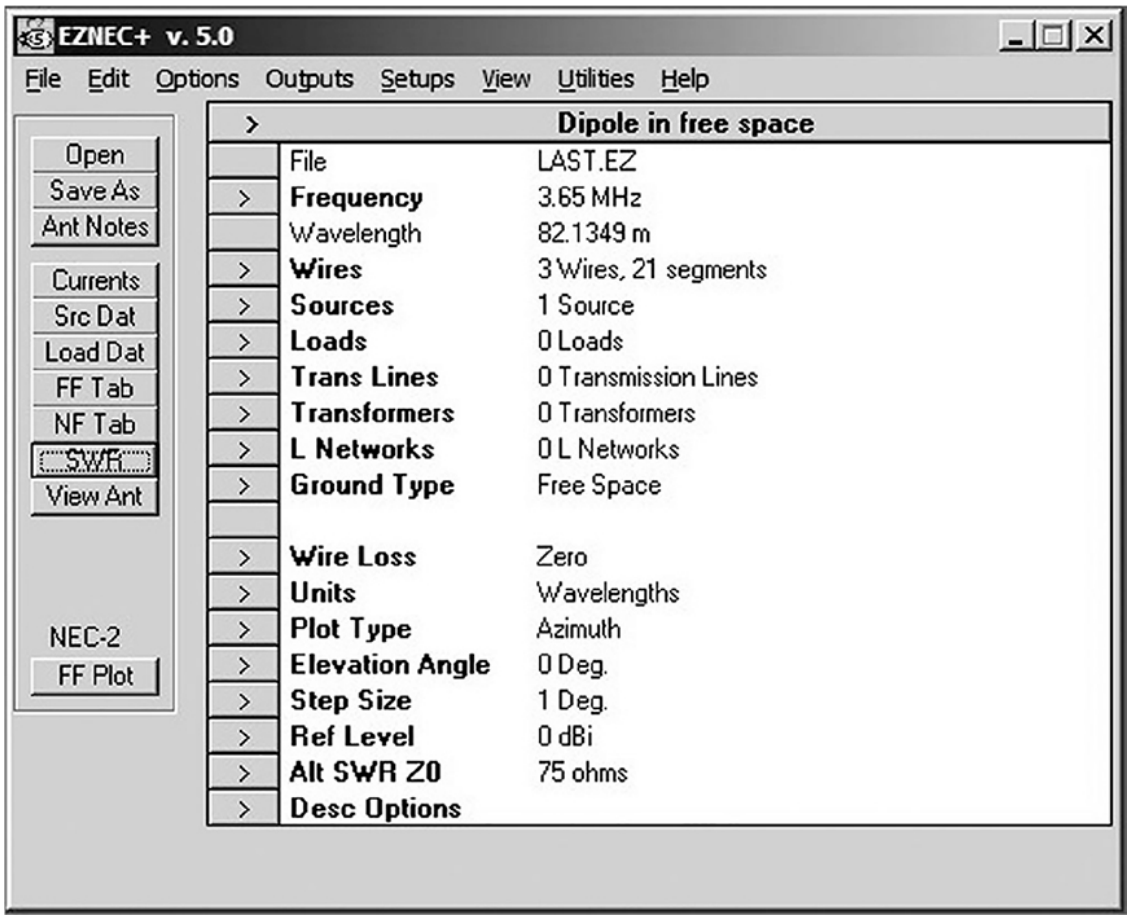


FIGURE 29.3 *EZNEC+ ver.* 5 opening screen.

Command buttons include Open, Save As, Ant[enna] Notes, Currents, Src [Source] Dat[a], Load Dat[a], F[ar]Field] Tab[le], N[ear]F[ield] Tab[le], SWR, View Ant[enna], and F[ar]F[ield] Plot. The "FF Plot" button graphs the far field of the antenna as a function of azimuth (compass heading) for a given radiation (elevation) angle, or as a function of radiation angle for a user-selected azimuth (heading). Alternatively, the user can select a 3D view with all elevations and azimuths in a single plot. Clicking on the "SWR" button brings up a new window for the user to define a frequency range and sample frequency intervals prior to running an SWR plot.

The wire and source tables used to create this antenna are shown in Fig. 29.4. The number of segments chosen for each of the three wires that make up the antenna keep the segment lengths as

similar as possible. The horizontal wire in this example is divided into an odd number of segments to allow placement of a source at the exact midpoint of the center segment and, hence, at the midpoint of the antenna, in case it is to be a center-fed dipole. As can be seen from the rightmost two columns of the wire table, this antenna is being modeled with insulated wire.

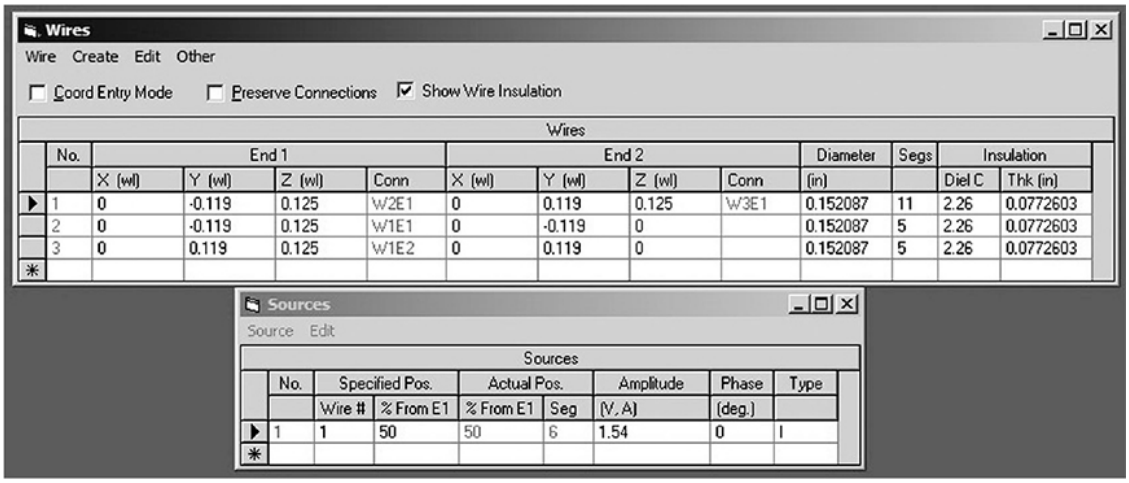


FIGURE 29.4 Bent dipole wire and source tables.

The "View Antenna" window (Fig. 29.5) displays the *x*, *y*, *z* coordinate system and places all the wires specified in the Wires table according to the coordinates entered there for each wire end. This window includes its own set of tools, including scroll bars to the left of the main window for zooming and repositioning the image within the window. Other options available from the drop-down menus include printer selection and setup, and a list of useful related objects (legends, currents, etc.) that can be displayed or not, as the user chooses. Once a set of calculations has been run, this window graphically shows (in a separate color) the magnitude of the current at all points on the antenna. A separate tabular listing of the current and its phase in each segment is also available for viewing, printing, or saving to a stand-alone file.

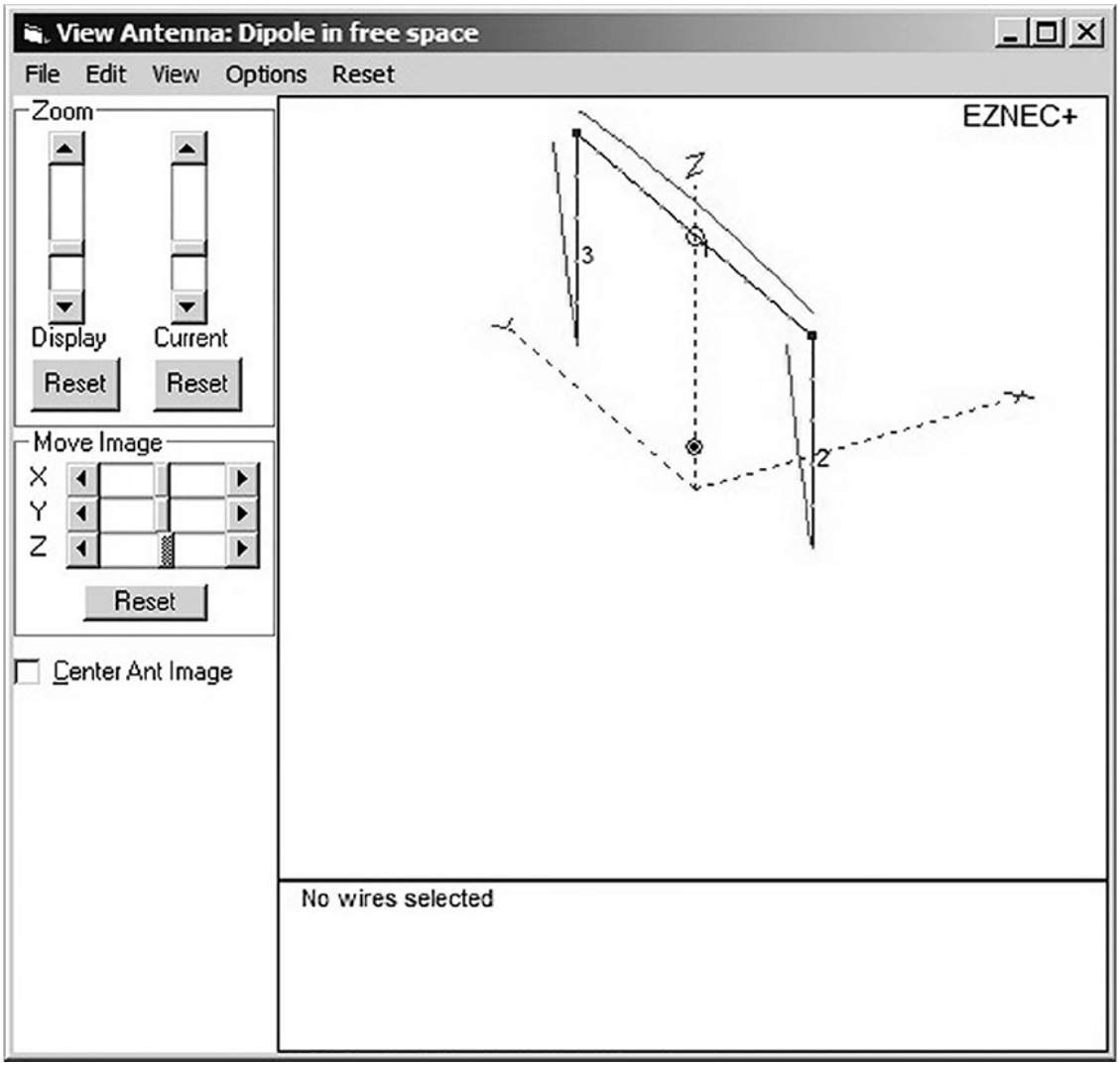


FIGURE 29.5 Bent dipole antenna view.

Once the antenna details have been entered, pressing the "FF Plot" button in the main window opens a new window (Fig. 29.6) displaying the antenna's far-field radiation pattern. The plot shown here is the radiated field versus azimuth for a bent dipole in free space. The plot also shows the elevation angle used for the calculation, the gain of the strongest lobe (wherever the pattern touches the 0-dB outer circle), front-to-back (F/B) ratio, beamwidth, and the azimuthal angle (listed as "Cursor Az") for the strongest lobe. The user can run an Elevation plot for a specific heading by making those selections and clicking on the "FF Plot" button again.

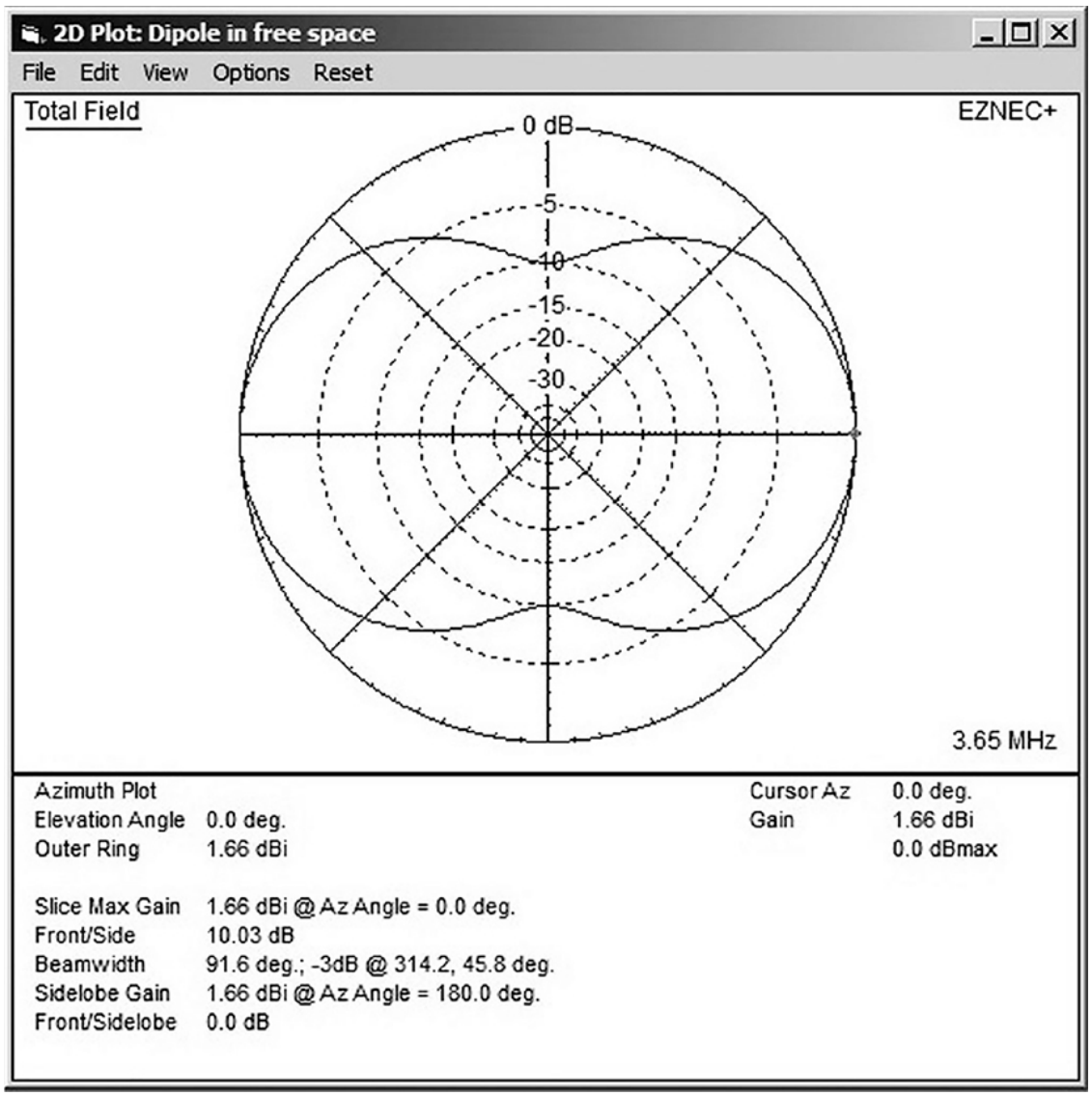


FIGURE 29.6 Bent dipole far field pattern.

For a horizontally polarized antenna over a perfectly conducting ground, the fields at a 0-degree elevation angle are zero. To find the azimuthal pattern of a horizontal dipole or Yagi over ground, a small but nonzero elevation angle of around 10 or 20 degrees should be used. Cubical quads and loops fed in the center of their top or bottom sides should be handled the same way.

The resistive component of a nonideal ground with finite conductivity will sap power from an advancing wavefront in contact with it, causing the zero elevation angle component of the wave to

become nonexistent long before reaching a distant receiver. When modeling an antenna over one of NEC's real ground options, choose a nonzero elevation angle to view the low-angle azimuthal radiation pattern of the antenna.

Some versions of EZNEC include SWR plotting capability. Figure 29.7 shows the 50-Ω plot for the dipole under consideration. The displayed data include SWR, complex feedpoint impedance (*Z*), reflection coefficient, and return loss.

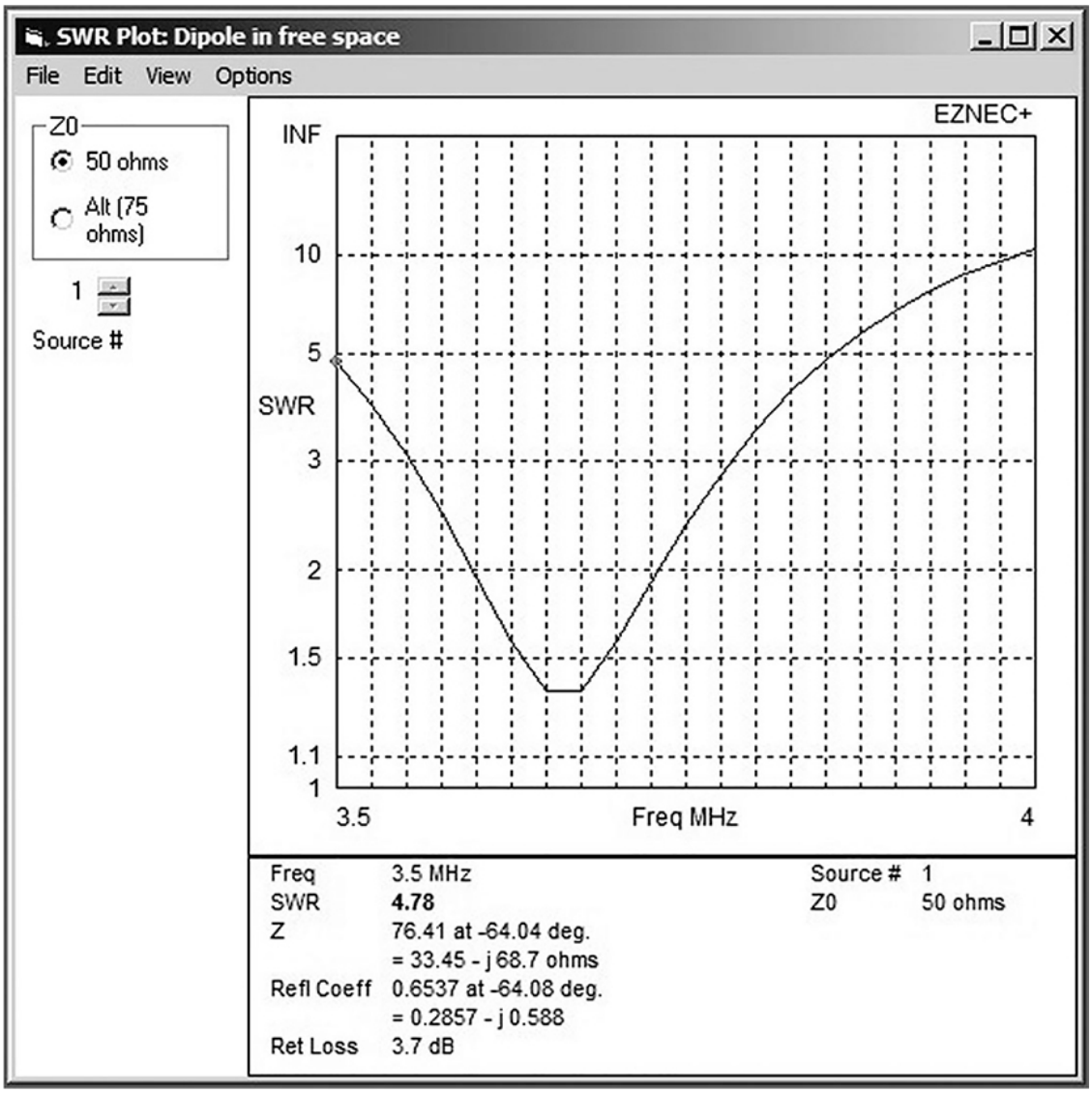


FIGURE 29.7 Bent dipole SWR plot.

29.5.2 Advanced Capabilities

Even early versions of NEC were quite useful and could model a variety of antennas, but newer software usually includes far more powerful features. Commercial antenna design packages include full 3D CAD capability that allows conformal and curved antennas to be modeled and optimized. NEC has been updated many times to include some of these features, including the ability to taper or step the diameter of thin wire antennas.

An interesting feature of EZNEC 5 is the ability to model adjacent grounds of differing characteristics. To access this capability, either of the real grounds ("High Accuracy" or "MiniNEC") must be selected. The user can choose to have a single ground extending infinitely far in all directions or split the surface between two grounds in either of two ways: by a straight line dividing the grounds at a user-specified x coordinate or by a circle having a user-specified radius "R" where the ground characteristics change. The first option requires the user to rotate model space around the x and y axes until the model's wires are oriented correctly with respect to the boundary between the two grounds, since this boundary is always parallel to the y axis. This feature is useful, for example, when the antenna is near a body of water.

Most software packages include a way to model transmission lines and other components attached to an antenna. Standard transmission line models are often included and can be incorporated quickly into an otherwise custom antenna model. In general, the user must specify where on the modeled antenna each conductor at the antenna end of the transmission line connects, the line length and characteristic impedance (Z_0), its velocity factor, and the loss per 100 ft at a user- or catalog-specified frequency.

The most powerful packages can in theory model the antenna, supporting structures, ground and buildings around the antenna, connecting transmission lines, and even the circuitry in the transmitter or receiver, but this is usually overkill for amateur

antenna designers. One of the most important principles in antenna design is that we can almost always design the antenna for a specific input impedance, and it will work with a standard transmitter. In systems with sophisticated performance requirements, the antenna and transmitter or receiver circuits can be jointly optimized or codesigned.

29.5.3 Accuracy and Computation Time

For numerical simulations, there is generally a tradeoff between the accuracy of the results and the required computation time. For the method of moments, this is controlled primarily by the number of wire segments used to represent the antenna and neighboring conductors. If the model takes too long to run, the number of segments can be reduced. A rule of thumb is 10 or more segments for each half-wavelength of wire length plus a few added segments wherever two wires meet at an angle or where two wires of different diameter are joined.

Although a model of an antenna could take into account all conducting structures in the vicinity of the target antenna, some common sense is required. For horizontal HF antennas, there is likely to be little effect if we leave out vertical masts and lattice towers having horizontal dimensions less than

$\lambda/2$

0 at the modeling frequencies of interest. On the other hand, modeling an 80-m grounded vertical or a 40-m elevated ground-plane antenna within a wavelength of a 15-m or taller guyed tower supporting a triband Yagi might not yield accurate predictions unless the tower, Yagi, and metallic guy wires are added to the model.

One tricky detail with NEC is issues with wires in contact with or close to the ground. The software offers multiple ways to model the ground and an assortment of representative parameters for multiple soil types, freshwater lakes and ponds, seawater, and custom permeability and dielectric constants defined by the user. But if a wire that is part of an antenna is too close to ground, the

modeled results can be inaccurate. Grounded MF and HF verticals with extensive radial systems can be approximated by placing the radials a few inches above ground. Model results for antennas based on the use of counterpoises (low wires running in close proximity to ground directly underneath the antenna wire) are similarly suspect, and the user should proceed with caution.

Once you have put together a model for your particular antenna, you can test the model for convergence. This is done by increasing the number of segments used to model conductors and ensuring that the results remain fairly stable. A good model will yield results for parameters like maximum forward gain or feedpoint impedance that vary by 1 percent or less with small changes in segment counts.

29.6 Terrain and Propagation Modeling Software

When modeling antennas, the overall goal is usually to maximize the ability to communicate via radio waves. While development of an optimized antenna is an important step, putting that antenna in the right location and at the right height can be just as important as the antenna itself.

Various categories of modeling software are available for antenna siting. These include:

- Terrain analysis codes that use ray tracing or similar high-frequency approximations to model the propagation of radio waves over topographic features. Terrain analysis helps in determining the height or location of antennas and their supports on a specific parcel of land and to compare the relative merits of candidate sites for an antenna facility. The terrain profile can be entered by hand, or topographical databases can often be downloaded for automatic use by the software.

- Ionospheric propagation programs provide statistical predictions of optimum working frequencies, the probabilities and times for communicating between two points at any given time of the day, the propagation modes providing the most efficient path between those points, and can predict received signal levels at distant receiving locations anywhere on the globe for a given transmitter power level.

See Chap. 3 for additional information regarding these software tools.

29.7 Summary

Many capable and inexpensive modeling tools as well as powerful commercial packages are available to antenna experimenters. Modeling can be used to avoid hours or even days that might previously have been spent trudging through swamps, prickly brush, or deep snowdrifts to erect a configuration of wires and test it in the field without knowing beforehand how it will perform.

NEC and other modeling software packages have many details under the hood that can take a long time to learn. Textbooks, online tutorials, and courses are available from ARRL and other organizations and publishers.

CHAPTER 30
Testing and Troubleshooting

This chapter examines techniques for testing antenna systems. We will explore some of the instruments that help evaluate the performance of the system, including the antenna and the RF path to and from the transceiver. Because the feedpoints of many antennas, especially on the MF, HF, and VHF bands, are high up and far from reach, many of the devices described here are intended to give us some idea of conditions topside on the antenna while we are at ground level.

When measuring, monitoring, or troubleshooting an antenna system, we are typically interested in certain key attributes of the system components:

- Signal strength at a distance
- Impedance matching throughout the system
- RF attenuation (loss) throughout the system
- Useable bandwidth of the system

Testing can be done at various junctions in a system. We can measure the RF signal amplitude immediately at the output of a transmitter or before or after a separate power amplifier if one is used as well as before or after a transmission line to the antenna feedpoint. The SWR can be measured before the transmission line or at the antenna feedpoint to determine how well the system is matched. Some transceivers include a wattmeter with a directional coupler that is capable of indicating forward (incident) power and reflected power or SWR. If no external power amplifier is present, the transceiver wattmeter is adequate for measurements of antenna and transmission line performance.

Older test equipment based on clever circuits like the Wheatstone bridge or the grid dip meter or dip oscillator have largely been replaced by antenna analyzers and similar highly sophisticated, low cost, and easily portable instruments.

30.1 A Review of SWR

Many measurement techniques for antenna systems either measure directly or are used to determine the voltage standing wave ratio on a transmission line in the system or the impedance of the antenna which can cause a poor SWR. The standing wave radio (SWR) is covered in Sec. 6.8.1. If the impedance looking into an antenna is different from the characteristic impedance of the transmission line connecting it to the transceiver, there will be both a forward and reflected wave on the line. Interference between the forward and reflected waves produces a standing wave pattern on the line.

As discussed earlier, we generally want no reflected wave and only forward power, which corresponds to an SWR of 1. Higher values mean that there is reflected power on the line which is normally wasted and undesirable. Reflections are caused by an antenna impedance with a real part different from the characteristic impedance of the transmission line (usually 50 Ω) or a nonzero reactance. If the antenna impedance has a nonzero imaginary part, the antenna is reactive which also causes reflections on the transmission line and leads to an SWR greater than 1.

The frequency range over which the SWR is acceptably low is the SWR bandwidth. The bandwidth can be too narrow, particularly if the antenna is small compared to the wavelength at its operating frequency. An antenna with resistive or dissipative losses comparable to the radiation resistance of the antenna itself can have a wide SWR bandwidth but may not perform well due to attenuation of the signal. High SWR on a transmission line causes greater losses along the line than when the line is matched to the load impedance as current maxima along the line increase the

I^2R
power dissipation.

30.2 Antenna Analyzers

An antenna analyzer or SWR analyzer uses a low-power RF signal generator to measure the VSWR on a system transmission line connected to an antenna or other complex load impedance. Because these units employ wideband small-signal detectors to keep costs down, they are susceptible to errors from ambient RF fields within their frequency ranges, such as might be experienced when using the device in the presence of active transmitters or in the vicinity of broadcast stations. Filters can be added to suppress interference, but the filter can alter the impedances reported by the instrument.

Antenna analyzers can usually measure other parts of the RF path in a typical installation. They can measure the velocity factor of a transmission line, help in tuning or adjusting matching stubs or matching networks, measure capacitance or inductance, and determine the resonant frequency of LC networks.

30.3 Spectrum Analyzers

One of the difficulties with RF power measurements is dealing with crowded propagation environments with many strong signals, wanted and unwanted, and wideband signals with complex modulations. An RF power meter lumps all signals within its measurement bandwidth into one power level. More information about a complex signal environment can be obtained from a spectrum analyzer. A well-calibrated commercial grade spectrum analyzer may cost several tens of thousands of dollars, but for routine use, handheld, battery-powered analyzers are readily available for a few hundred dollars or less.

A spectrum analyzer converts an input signal into the frequency domain and gives the power (typically in dBm per bin) within many

small frequency bins. The frequency range of displayed power results can be adjusted by changing the spectrum analyzer span. Spectrum analyzers have a marker function that allows power at a specific frequency to be read off. For noise or fluctuating signals, averaging can be used to reduce power estimation error and smooth out variations in time.

To measure the total power for a wideband modulated signal, the power in each displayed bin must be added, or the spectrum analyzer's resolution bandwidth set to be large enough that all the power is included in one displayed bin. Most spectrum analyzers have a power spectral density mode that gives power levels for wideband noise in units of dBm/Hz.

30.4 Vector Network Analyzers

A recent addition to the stable of affordable analyzers for antenna systems is the vector network analyzer. Precision network analyzers have and still remain large and expensive laboratory instruments, but small, low-cost VNAs are now available that operate well into the GHz range for around $100. Kits are available that allow VNAs to be built by hand.

A single-port VNA provides resistance, reactance, and the sign of the reactance. A multiport VNA can provide S-parameters between two or more ports. This allows simultaneous measurement of the input match or return loss, insertion loss, and output match for a two-port device like a length of transmission line or a filter. Measurements are normally shown as the magnitude of S-parameters in dB as a function of frequency, but results can also be displayed on a Smith chart or as impedance (resistance and reactance) curves.

A VNA can be used to measure the attenuation and electrical length of a transmission line. VNAs have a procedure for calibrating out the effects of a transmission line between the antenna and the VNA. A low-cost VNA may not be well calibrated enough to measure loss to better than a dB or so, but by following the recommended

instrument calibration procedure or by making a careful A-B comparison with a well-characterized short, low-loss through line, better measurement accuracy can be obtained.

In addition to swept plots of VSWR, feedpoint impedance and reactance, and Smith chart displays of antenna performance across the entire band, VNAs can function as time domain reflectometers (TDRs), providing a detailed graphical look at pulsed waveforms sent down a transmission line. At each point along the line that an outbound pulse encounters an impedance other than the Z_0 of the line, a reflected pulse is created that travels back to the instrument, where it is detected and plotted. With this capability, the instrument can locate connectors, splices, transitions in the feedlines, and potential faults (short circuit, open circuit, or other unexpected discontinuity) between the radio room and an antennas hundreds of feet away. Figure 30.1 shows an example VNA trace of the entire RF path from a radio room to an antenna. Reflections from cable transitions, in-line components, and the antenna feedpoint can be seen.

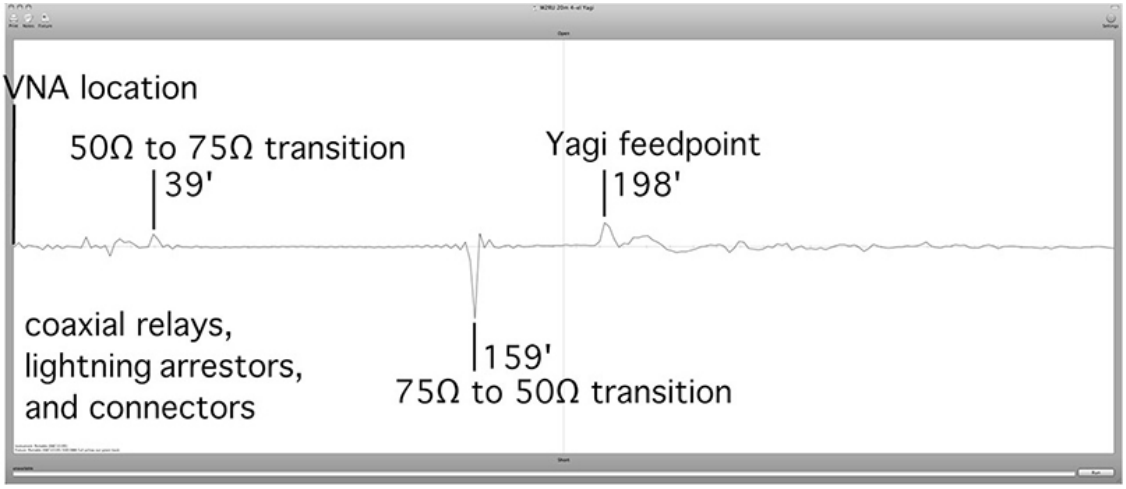


FIGURE 30.1 VNA TDR trace example.

30.5 Field Strength Meters

A field strength meter (FSM) measures the radiated field from an antenna directly. Commercial grade instruments are calibrated in watts per square centimeter or volts per meter and are used for precise measurements. For adjusting antennas, a simpler instrument is sufficient. This section describes two simple passive (no DC power is required) field strength meters useable for adjusting HF radio antennas.

Two circuits are shown in Figs. 30.2 and 30.3. Figure 30.2 shows a simple untuned FSM. In this circuit, a small whip antenna used for signal pickup is connected to one end of a grounded RF choke (RFC). The RF voltage developed across the RFC by the signal being measured is applied to a germanium diode detector; either a 1N34 or an 1N60 can be used. Silicon diodes are normally preferred, but in this case we need the lower contact potential of germanium diodes in order to improve sensitivity (V_g is 0.2 to 0.3 V for Ge and 0.6 to 0.7 V for Si). A potentiometer is used both as the load for the diode and as a sensitivity control to set the meter reading to a convenient level.

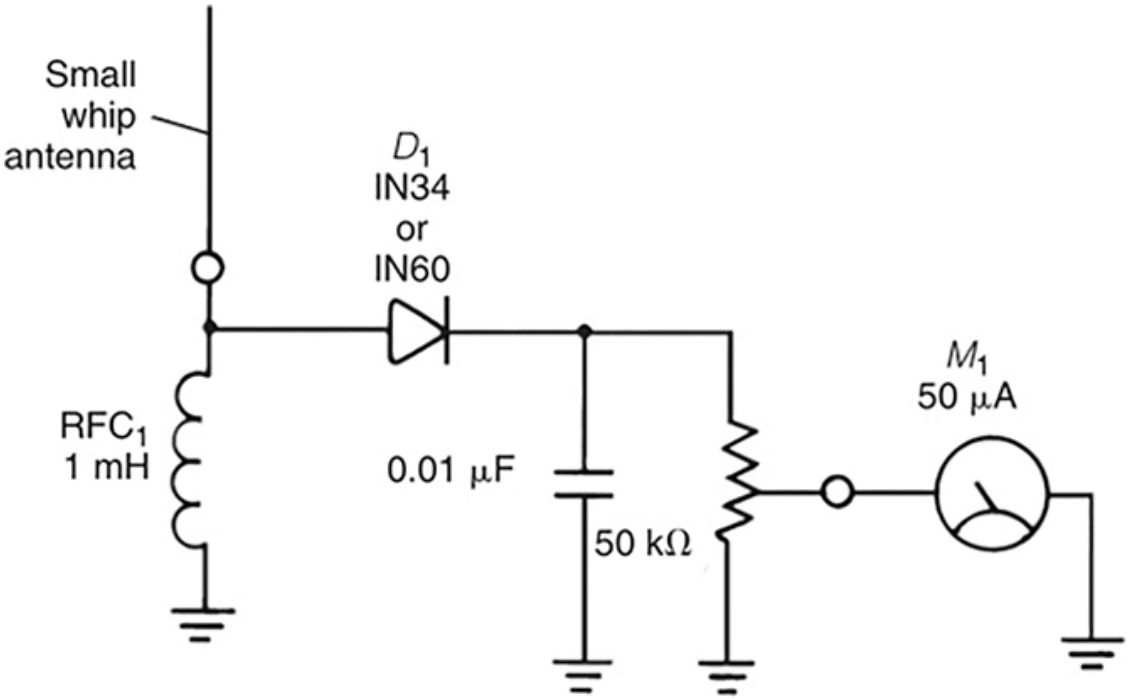


FIGURE 30.2 Simple field strength meter.

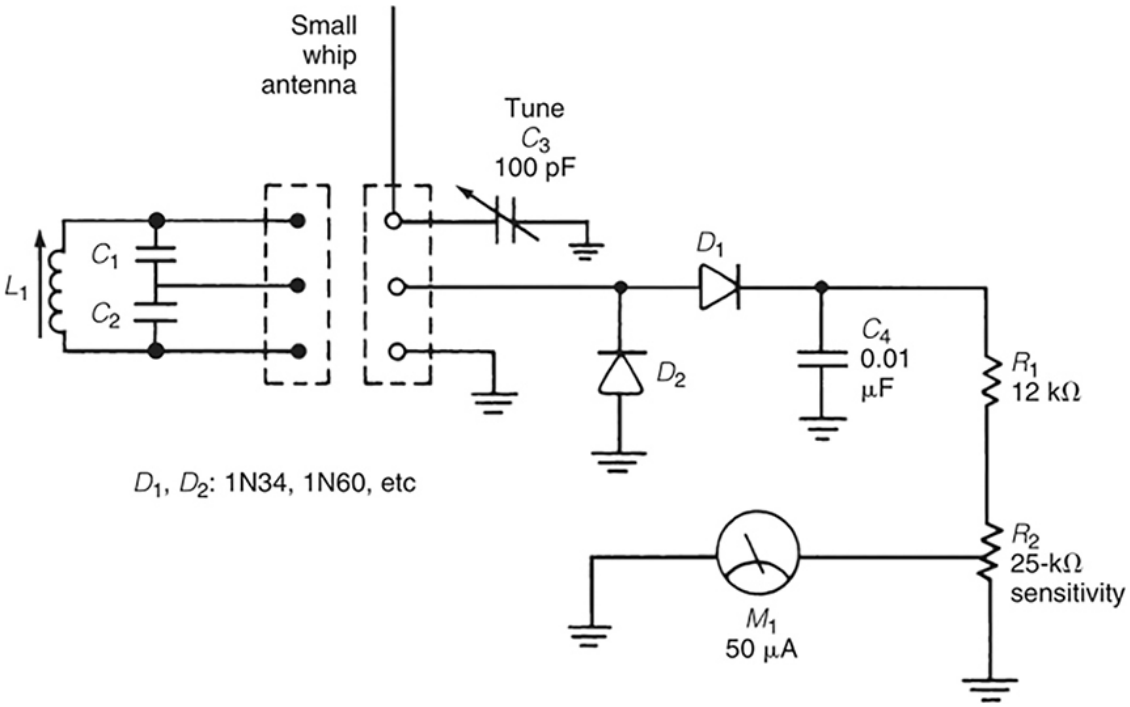


FIGURE 30.3 Tuned field strength meter.

The untuned version of the FSM suffers from a lack of sensitivity. It can be improved by adding a tuned circuit and an impedance-matching scheme, as shown in Fig. 30.3. A variable capacitor C_3 is used as a tuning control in parallel resonance with inductor L_1. A tapped capacitive voltage divider (C_1/C_2) is used to provide impedance matching to the diode. Figure 30.4 gives values of capacitance and inductance for various bands.

Band (meters)	L_1 (μH)	C_1 (pF)	C_2 (pF)
75/80	27	25	100
40	10	15	68
30	4.4	15	68
20	2.25	10	68
17	1.7	10	56
15	1.3	10	47
13	0.8	10	39
10	0.5	10	39

FIGURE 30.4 Component values for field strength meters.

Operation of the tunable FSM is straightforward. Set the sensitivity control to approximately half-scale and key the transmitter. Adjust the tuning control (C_3) for maximum deflection of the meter pointer (readjustment of the sensitivity control may be needed). After these adjustments are made, the tunable FSM works just like any other FSM. The FSM is

best utilized at least a few wavelengths from the transmitting antenna, outside of the reactive near field.

30.6 RF Wattmeters

Measuring RF power is straightforward for a single RF tone, but for complex waveforms or a crowded environment with many signals it can be challenging. We can measure power on unmodulated sinusoids or equivalently a continuous wave (CW) signal by measuring either the RMS or peak values of voltage or current and assuming that a constant resistance load is present. For a CW tone, power dissipated in a resistive load is related to the peak and RMS voltages by

$$P = \frac{(V_{\text{peak}})^2}{2Z_0} = \frac{(V_{\text{RMS}})^2}{Z_0} \quad (30.1)$$

Accurate measurement becomes more difficult in the presence of complex waveforms such as modulated signals. With some power meters, peak and average power readings can vary markedly with modulation type.

One of the earliest practical RF power-measuring devices was the thermocouple RF ammeter (see Fig. 30.5). This instrument works by dissipating a small amount of power at a small resistance inside the meter and measuring the resulting heat generated with a thermocouple. A DC ammeter reports thermocouple current. Because it works on the basis of the power dissipated in heating a resistance, a thermocouple RF ammeter is inherently an RMS-reading device and performs a fairly long-term average for signals with rapidly fluctuating power levels.

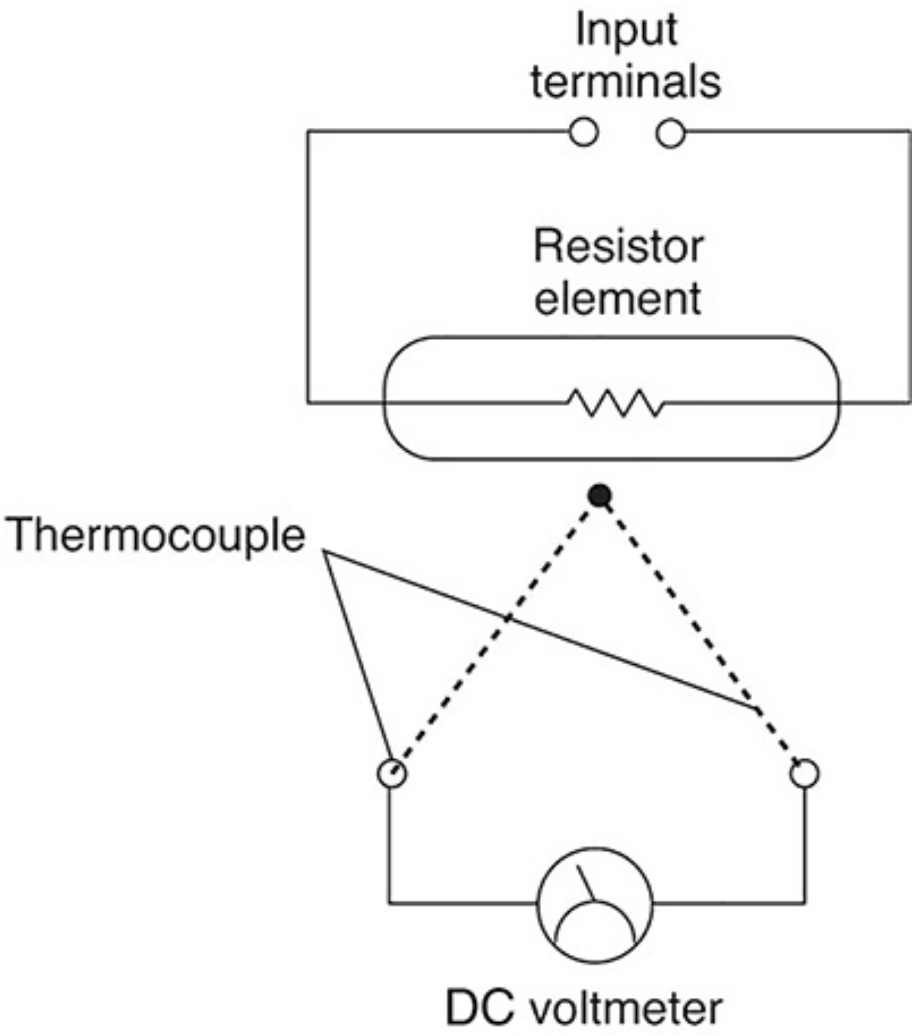
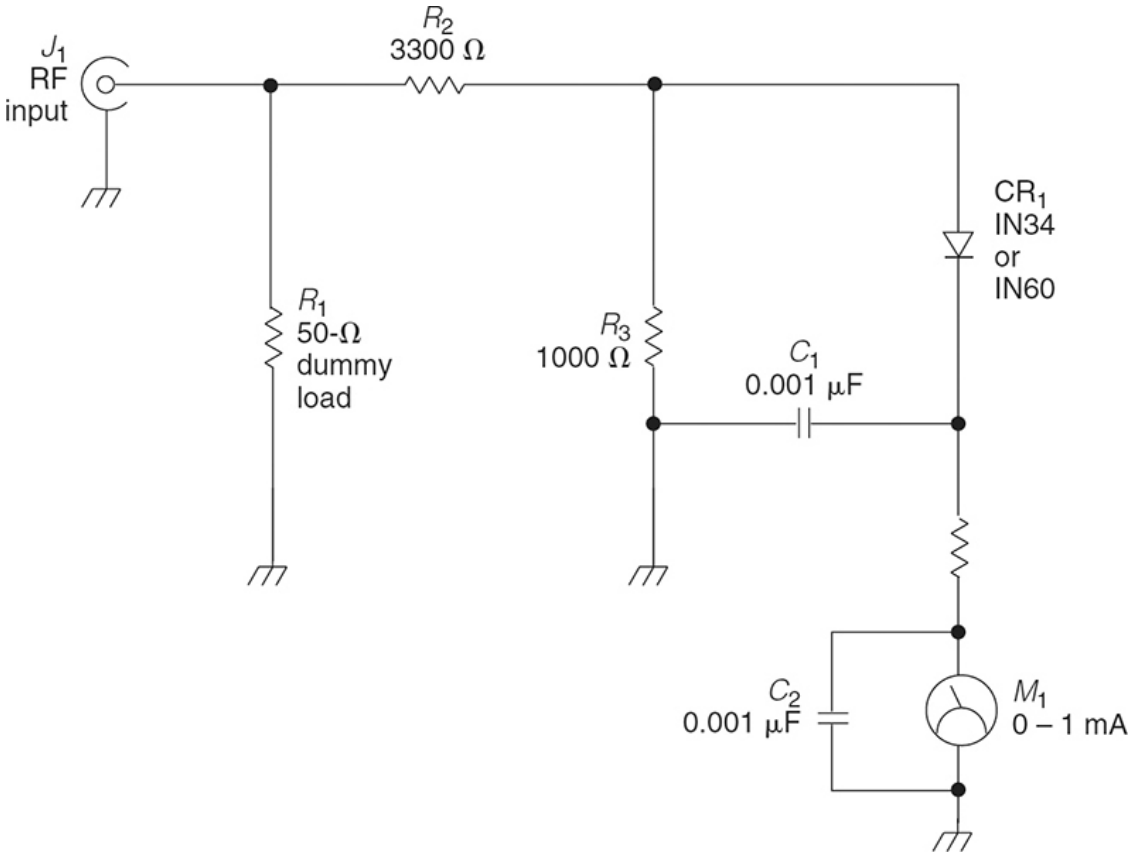


FIGURE 30.5 Circuit for thermocouple RF ammeter.

Thermocouple RF ammeters can suffer from poor accuracy over frequency. A calibrated frequency response curve can be used to correct the results. In general, at 10 MHz and higher, the readings of a thermocouple RF ammeter are only rough estimates without the original meter calibration chart.

RF power can also be measured by knowing the voltage across a load resistance. In the circuit of Fig. 30.6, the RF voltage appearing across the load is scaled down to a level compatible with the voltmeter by the resistor voltage divider (R_2/R_3). The output of this divider is rectified by

CR_1
and filtered to DC by capacitor
C_2
.



F IGURE 30.6 RF wattmeter.

Obtaining the voltage measurement from a simple diode voltmeter is valid only if the RF signal is unmodulated and has a pure sinusoidal waveform. While these criteria are almost always met with CW transmissions, they are only approximate for other waveforms, such as those characteristic of voice analog modes. If the voltmeter circuit is peak reading, as in Fig. 30.6, the peak power is

$$P = \frac{V_0^2}{R_L} \quad (30.2)$$

The average power is found by multiplying the peak power by 0.707. Some meter circuits include voltage dividers that convert the reading to RMS.

Another family of RF power meters is based on bridge circuits. Figure 30.7 shows a bridge set up to measure both forward and reverse power. There are four elements in this quasi-Wheatstone bridge circuit:
R_1
,
R_2
,
R_3
, and the antenna impedance (connected to the bridge at
J_2
). If
R_{ant}
is the antenna resistance, then we know that the bridge is in balance (i.e., the null condition) when the ratios
R_1/R_2
and
R_3/R_{ant}
are equal. In an ideal situation, resistor
R_3
will have a resistance equal to
R_{ant}
. In practice we use a compromise value such as 67 Ω for
R_3
. Such a resistor will be useable on both 50- and 75-Ω antenna systems with only small errors. Typically, these meters are designed to read relative power level, rather than the actual power.

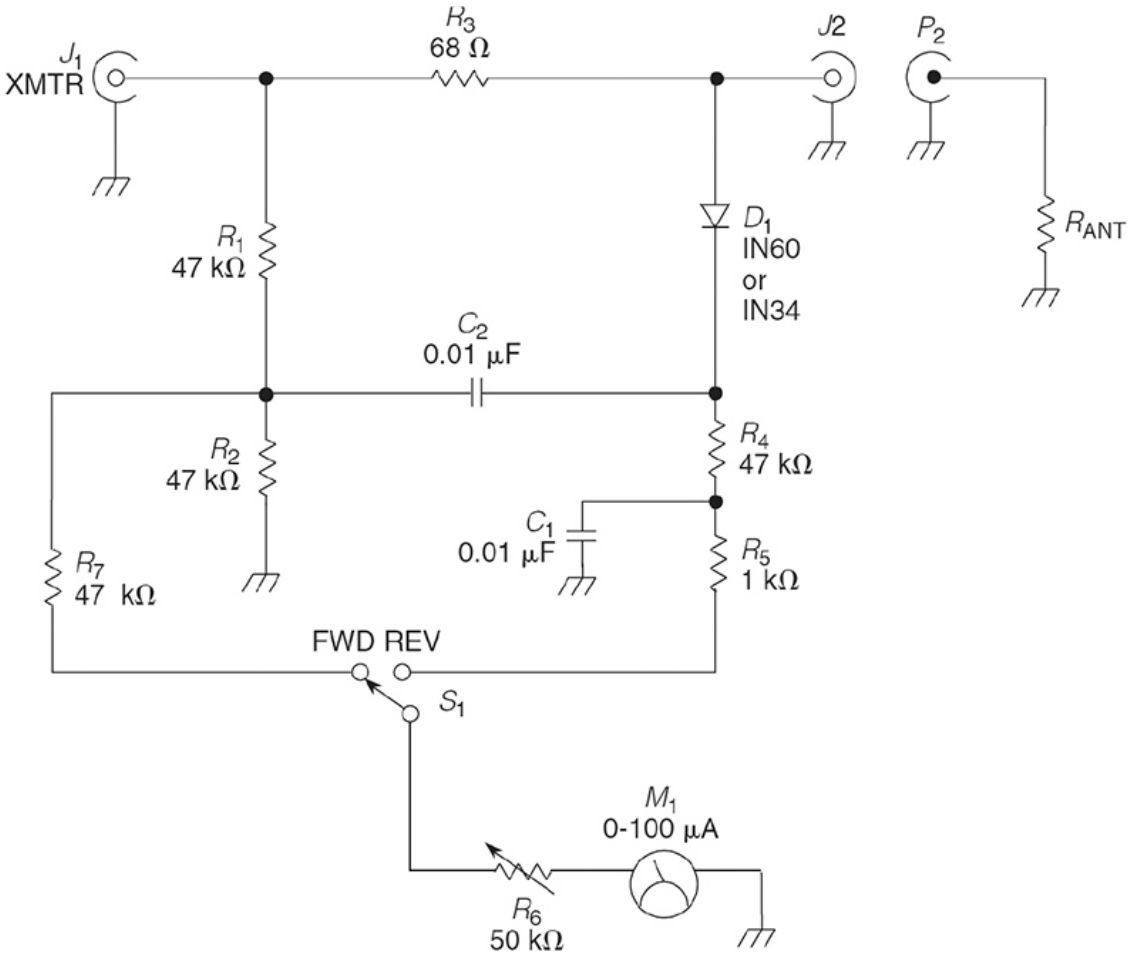


FIGURE 30.7 Bridge arrangement to measure both forward and reverse power.

This meter can be used to find the VSWR. With the switch in the forward position and RF power applied to J_1 (XMTR), potentiometer R_6 is adjusted to produce a full-scale deflection on meter M_1. When the switch is then set to the reverse position, the meter will read reverse power relative to the VSWR. The meter can include a VSWR scale with calibration specific to the circuit configuration employed. Because it dissipates a considerable amount of RF

power, the bridge of Fig. 30.7 cannot be left in the circuit while transmitting. A switch may be included to bypass the bridge.

An improved bridge circuit, shown in Fig. 30.8, is the capacitor/resistor bridge or micromatch bridge. It uses only 1 Ω in series with the line (R_i), thus dissipating considerably less power than the resistance used in the previous example and we can leave the micro-match in the line while transmitting. The sensitivity control can be used to calibrate the meter.

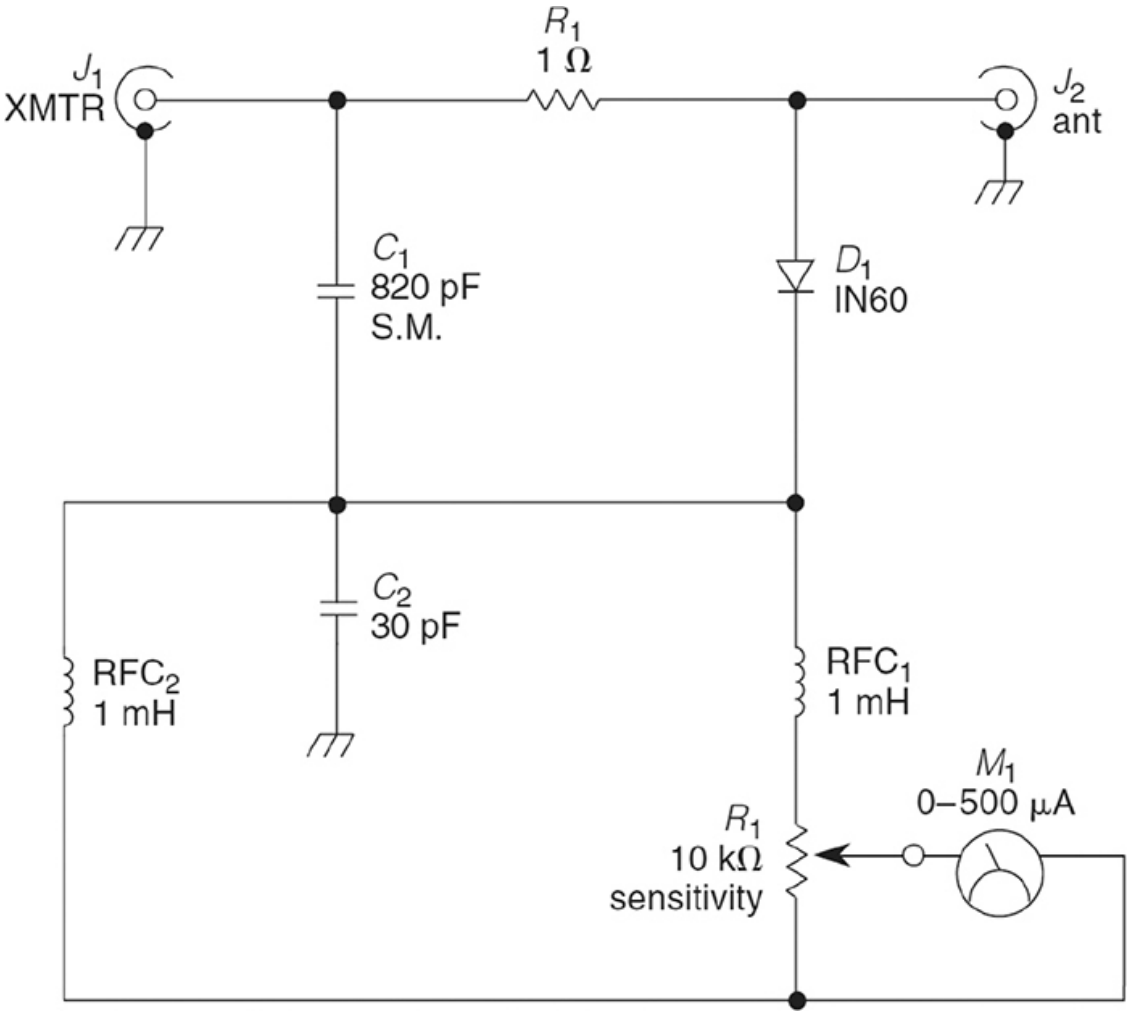


FIGURE 30.8 Micromatch wattmeter.

The monomatch bridge circuit in Fig. 30.9 is used for HF and low-VHF applications. The transmission line is segment B, while RF sampling elements are formed by segments A and C. The sensor is essentially a directional coupler with a detector for both forward and reverse directions. The resistance of

R_1

and

R_2

should match the transmission line impedance, although in many instruments a 68-Ω compromise resistance is used.

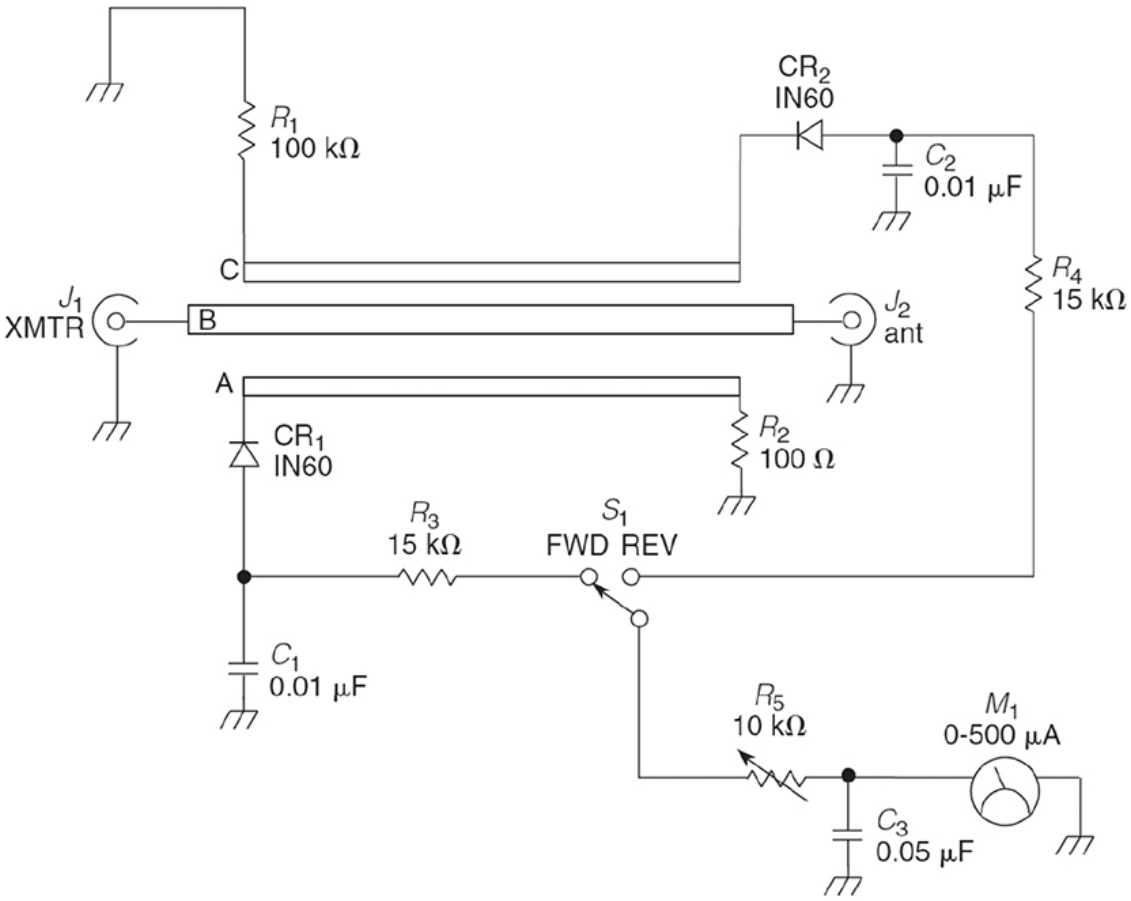


FIGURE 30.9 Monomatch wattmeter.

Another VSWR meter sensor is the current transformer shown in Fig. 30.10. A single-turn ferrite toroid transformer is used as the directional sensor. The transmission line passing through the hole in

the toroid forms the primary winding of a broadband RF transformer. The secondary, which consists of 10 to 40 turns of small enamel wire, is connected to a measurement bridge circuit with a rectified DC output.

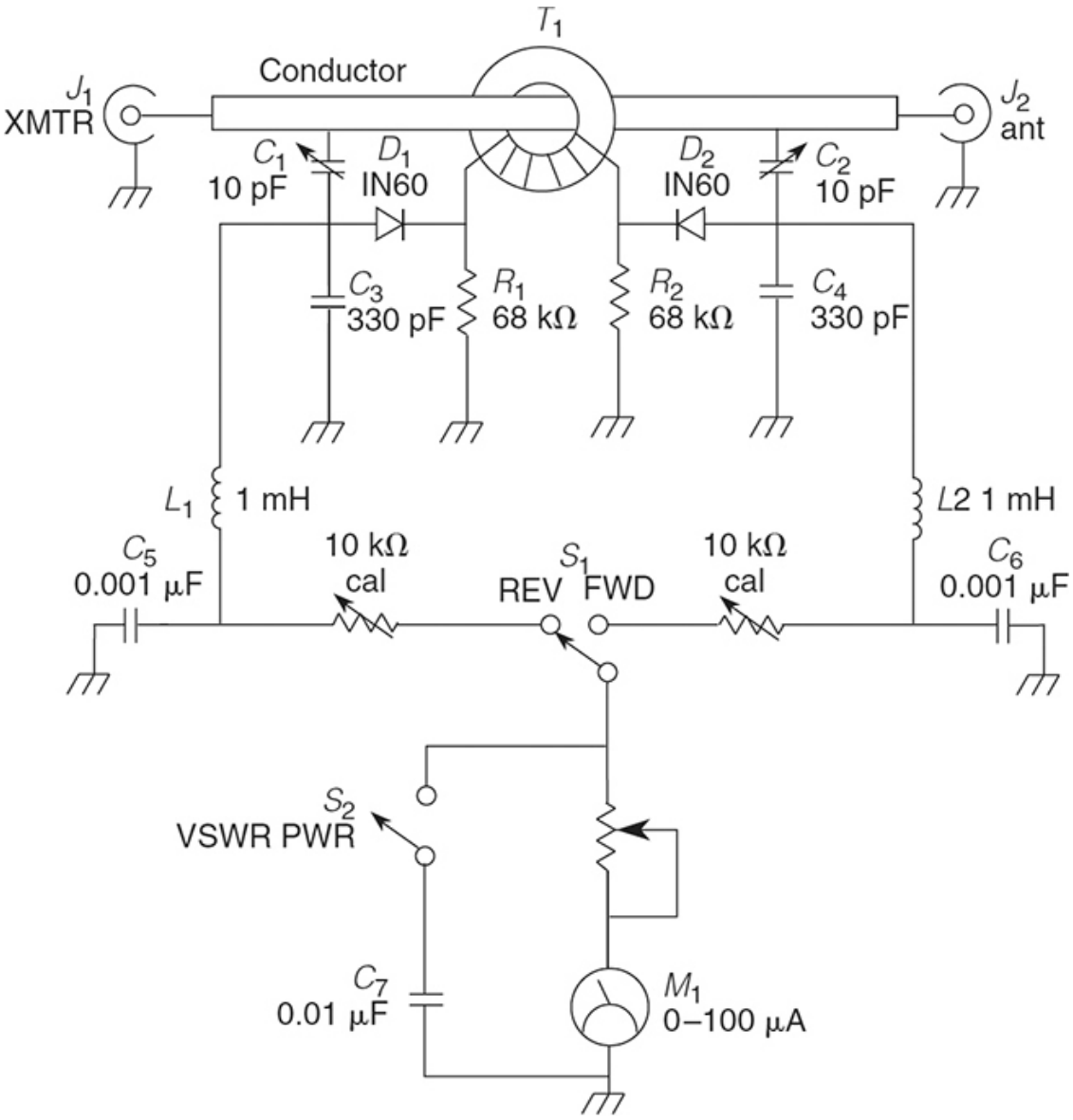


FIGURE 30.10 Current transformer wattmeter.

A Thruline meter can be inserted into the transmission line of an antenna system with so little loss that it may be left permanently in the line during normal operations. The heart of the Thruline meter

is the directional coupler transmission line assembly shown in Fig. 30.11; it is connected in series with the transmission line to the antenna or dummy load. The directional element can be rotated 180 degrees to measure either forward or reverse power levels. Each plug-in element contains a sampling loop and diode detector designed to cover a specific range of frequencies. The main RF barrel is a coaxial line segment with a 50-Ω or 75-Ω characteristic impedance. Some Thruline meters for high power applications use an in-line coaxial cable coupler for broadcast-style hardline and a remote indicator.

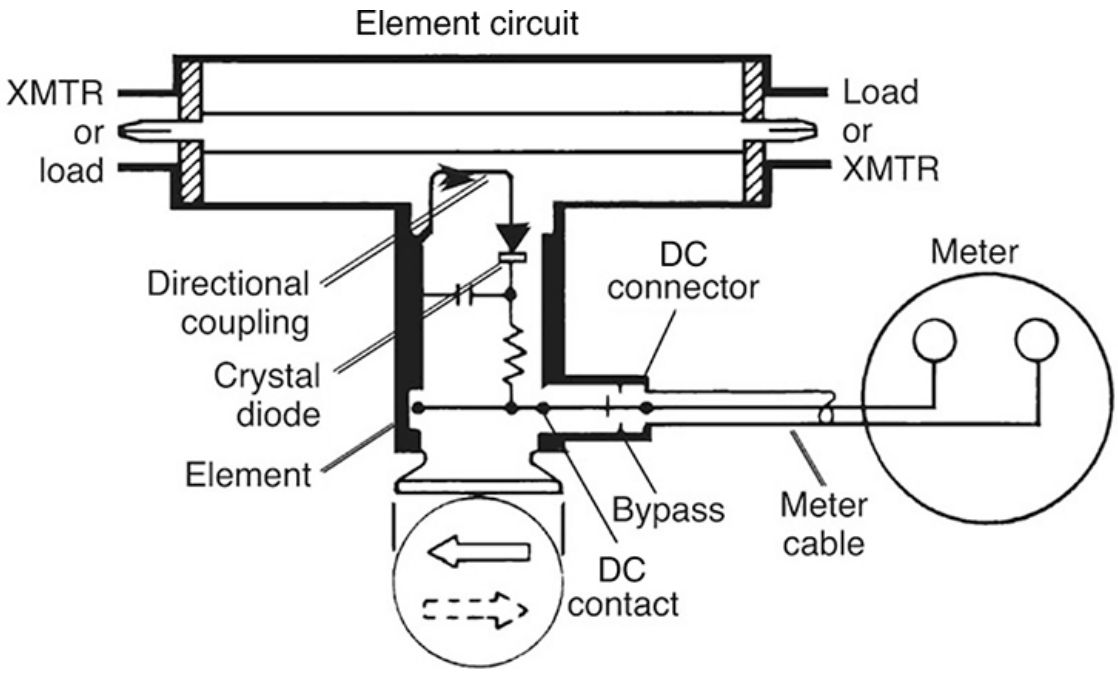


FIGURE 30.11 Thruline sensor circuit.

30.7 Testing Coaxial Cable

When you install an antenna, or do a bit of preventive maintenance, or find the antenna is not working properly, one system component it is important to check is the transmission line and connectors. A two-step process is typically employed.

The first step consists of ohmic checks at DC, as shown in Fig. 30.12. With terminals A and B open, there should not be any

resistive path across the input terminals (as shown). If a high resistance is seen, there might be some contamination in the system or the dielectric insulation has failed, permitting leakage current to flow. A low resistance indicates a short circuit somewhere along the cable's length. If the cable has been cut or the connector is faulty, a short can result. We then short terminals A and B together. A low resistance should be noted. If not, then it is likely that the center conductor is open—typically as the result of some prior excessive stress (such as extreme bending) on the cable.

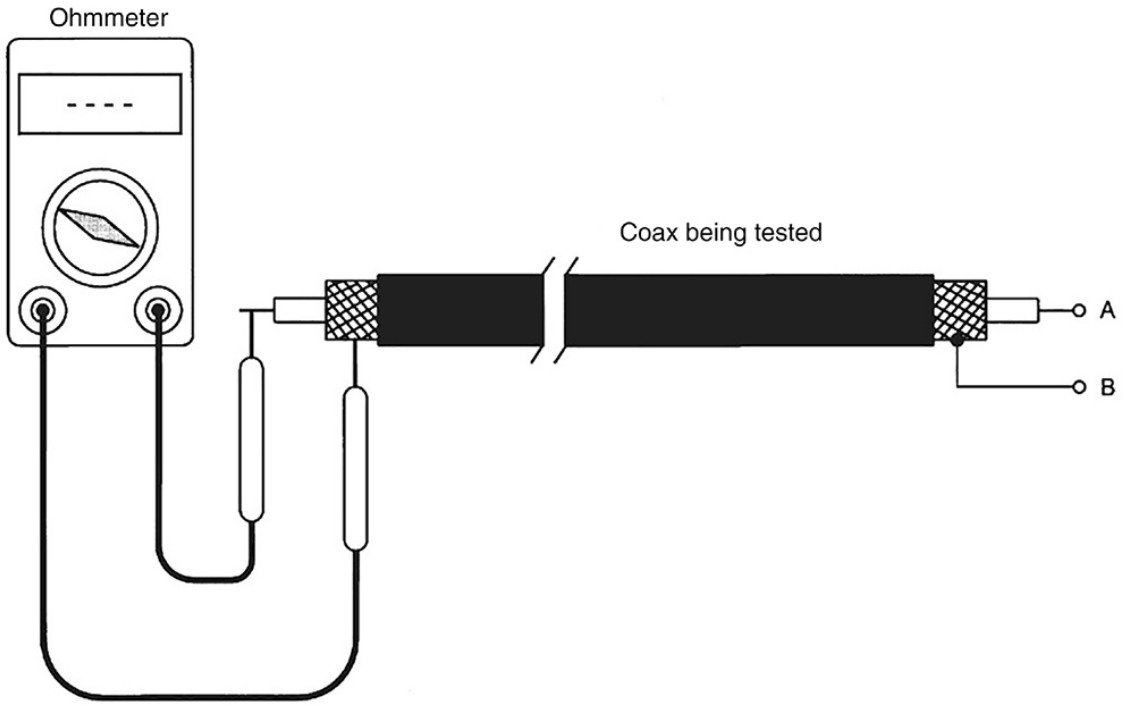


FIGURE 30.12 Ohmmeter test of coaxial cable.

Once you have completed the basic open-circuit and short-circuit ohmmeter measurements on the cable, the next step is to measure the loss of the transmission line at one or more frequencies of interest. Old cables that have not been adequately protected from the elements in prior usage or storage can develop losses that get higher as the cable deterioration progresses. Moisture is the enemy of coaxial cables, so existing cable stock should not be stored

outdoors or in damp basements or garages unless all cable ends have been sealed.

Loss in new cables increases with frequency and with the total length of cable used in any particular link between transmitter and antenna. Depending on the specific type of flexible coaxial cable employed, losses can run from as low as 0.2 dB/100 ft at 1.8 MHz (RG-8 with foam dielectric) to as much as 0.5 dB/100 ft (RG-58 and RG-59). At 30 MHz, the losses are approximately 1.0 dB/100 ft and 2.5 dB/100 ft, respectively. Common flexible coaxial cables at MF and HF are members of the RG-8 and RG-213 family. Their loss ranges from 0.25 dB/100 ft at 1.8 MHz to 1.3 dB/100 ft at 30 MHz.

Figure 30.13 shows a test setup for measuring coaxial cable loss. The first step is to determine the length of the piece of cable under test. One end of the line should then be terminated in its characteristic impedance—usually 50 or 75 Ω. The termination must be capable of dissipating the output power of the source, so usually a 100W (or higher) dummy load is employed. The simplest RF source is often your own transmitter or transceiver. Two identical RF power meters (M_1 and M_2) are used to simultaneously measure the input power to the line (M_1) and the power delivered to the load (M_2) while the RF source is activated. The difference in the two power readings is then the power lost in the cable itself. To express the loss in decibels, we use

$$\text{Loss}_{\text{dB}} = -10 \log_{10} \left(\frac{P_{M2}}{P_{M1}} \right) \quad (30.3)$$

where P_{M1} is the cable input power reading and

P_{M2}
is the cable output power. Manufacturer specifications typically give loss in dB per hundred feet. This can be calculated using

$$\text{Loss}_{\text{dB}/100\,\text{ft}} = \frac{\text{Loss (dB)}}{\text{L(ft)}} \times 100 \quad (30.4)$$

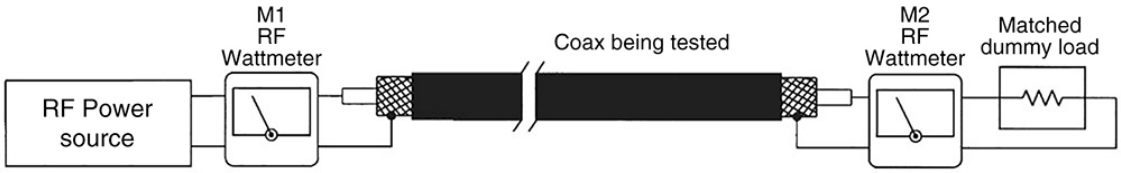


FIGURE 30.13 RF attenuation test setup for coaxial cable.

These tests can be run with a single meter by alternately inserting it in the signal path at each end of the coaxial cable being tested. This has the advantage of eliminating any errors caused by differences in the calibration of two meters but has the disadvantage of changing the test configuration in midmeasurement. When using the single-meter approach, there should be a good match between the cable and the termination. It is wise to perform these measurements on at least two different frequencies, preferably at the extremes of the frequency range over which the cable will be used.

30.8 Dummy Loads

A dummy load is a nonradiating substitute for an antenna used when measuring or testing parts of the antenna system after the transmitter. Amateur radio operators can use dummy loads to tune transmitters before switching to an antenna for final adjustments. Another use for dummy loads is to aid in troubleshooting antenna systems. Suppose we have a system in which the VSWR has become high enough to kick our transmitter off the air. Starting at the transmitter, we can disconnect successive elements of the RF transmission system and connect the dummy load at the disconnect

point. If the VSWR drops to a normal value, then the difficulty is downstream (i.e., toward the antenna). By repeating this process, we will eventually find the bad element, which is often as likely to be a coaxial cable, a connector, or a switch as it is to be the antenna itself.

Figure 30.14 shows a basic dummy load, which consists of one or more resistors connected in parallel, series, or series-parallel so that the total resistance is equal to the desired load impedance. The power dissipation is the sum of the individual power dissipations.

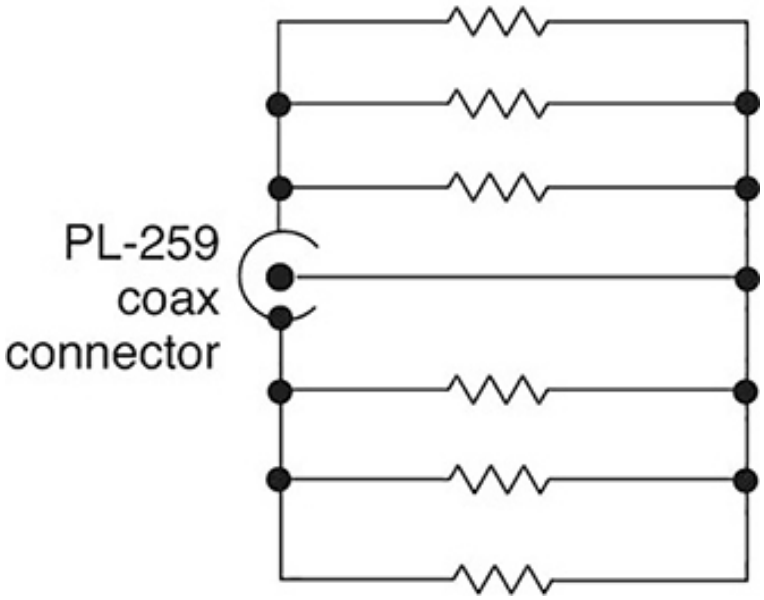


FIGURE 30.14 A basic dummy load.

Note *It is essential that noninductive resistors be used for this application. Carbon-composition or metal-film resistors are used. For very low frequency (VLF) work, it is permissible to use special low-inductance counterwound wire resistors. These resistors cannot be used for frequencies greater than a few hundred kilohertz. Homemade dummy loads should use many resistors in parallel, rather than in series, since it is easier to keep the stray inductance down in the former configuration*.

Commercial dummy loads are shown in Figs. 30.15 and 30.16. The load in Fig. 30.15 is a 5W model, typically used for servicing citizen band equipment. The resistor is mounted directly on a PL-259 coaxial connector. These loads typically work to about 300 MHz, although many are not really useful over about 150 MHz. A higher-power version of the same type is shown in Fig. 30.16. This device works in the low VHF region and dissipates up to 50 W in intermittent usage, which is perfect for servicing VHF land mobile amateur and marine transceivers.



FIGURE 30.15 CB-style 5W dummy load.



FIGURE 30.16 15W dummy load.

Another dummy load is shown in Figs. 30.17 and 30.18. The resistor is shown in Fig. 30.17 and a schematic view is shown in Fig. 30.18. The long, high-power noninductive resistor element is rated at 50 Ω and can dissipate 1000 W for several minutes. If longer times or higher powers are anticipated, forced air cooling can be applied by adding a blower. Dummy loads with additional monitoring and metering capabilities are also available.

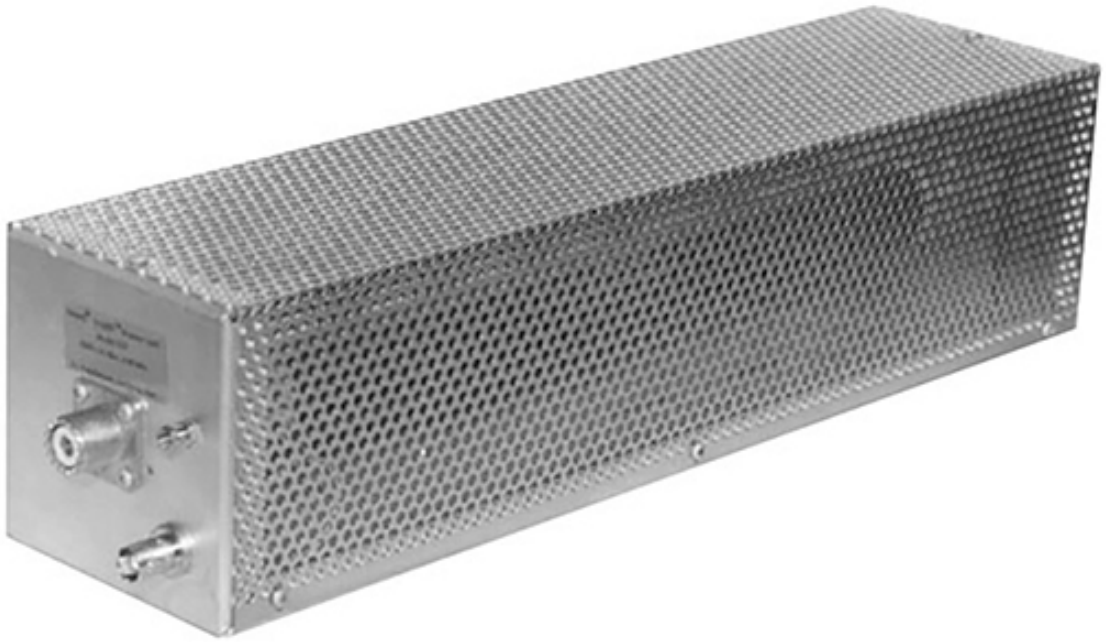


FIGURE 30.17 Drake *DL-1000* dummy load.

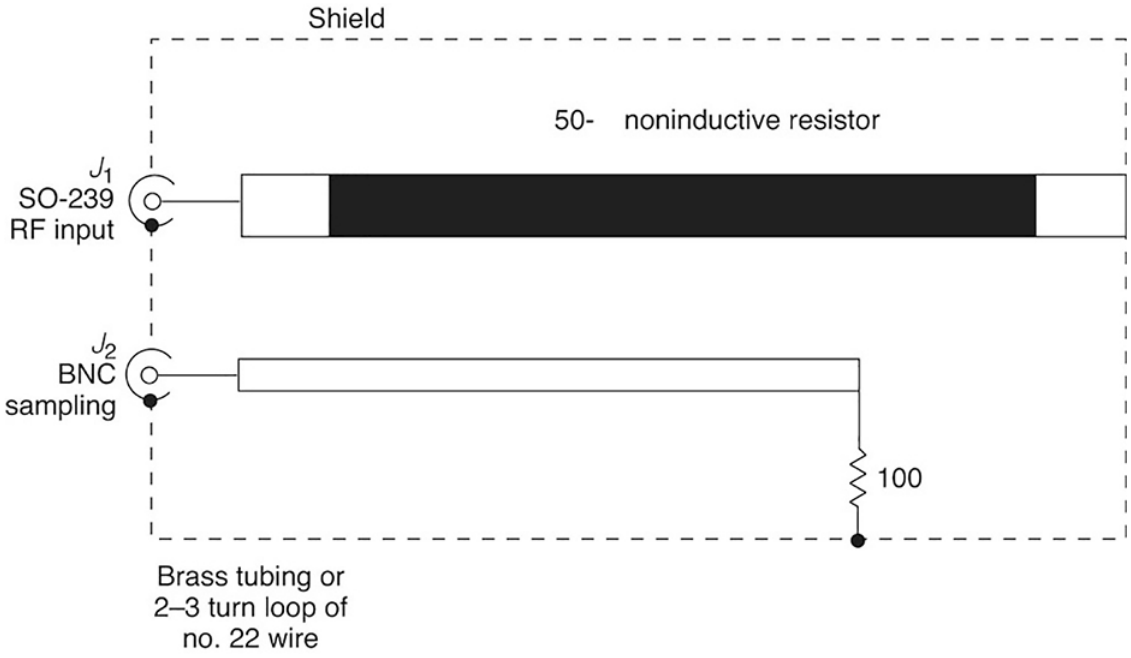


FIGURE 30.18 HF dummy load with pickup loop.

PART VII

Mechanical Construction and Installation

CHAPTER 31
Supports for Wires and Verticals
CHAPTER 32
Towers
CHAPTER 33
Grounding for Safety and Performance
CHAPTER 34
Zoning, Restrictive Covenants, and Neighbors

CHAPTER 31
Supports for Wires and Verticals

Antenna installation methods vary in complexity from those that can be performed by one person to large-scale projects that are best left to professional antenna riggers. In this chapter, we suggest techniques for erecting wire antennas and lightweight rigid verticals. In Chap. 32, we do the same for full-fledged towers. The information given herein is merely informal guidance and what you ultimately do should never be in violation of local mechanical, electrical, and other building codes or any zoning ordinances or restrictive covenants pertaining to your property.

31.1 Antenna Safety

To avoid injury or accident during installation or usage, we must first consider reliable mechanical installation and electrical safety. Even a simple wire antenna can be dangerous to erect. It is not possible to foresee all the situations you might face when installing an antenna. You must take responsibility for the risks involved in the process.

The antenna, its feedline, its supports, or any part thereof should avoid crossing over or under a power line, transformer, or other utility equipment. Users and bystanders are at risk of being electrocuted while installing or working on an antenna near an electrical line. There is never a time or situation when any part of an antenna system should be placed near enough to electrical power lines that it can come in contact with them if the antenna, its feedline, or its supports fail or if the power poles or power lines themselves come down. Avoid using a utility pole or guy wire as an attachment point for any rope, halyard, tug line, or guy wire.

Never rely on the insulation covering antenna wire or the power lines to protect you from high voltages on the utility company wires. Never assume that the power lines are insulated. Old insulation crumbles on contact with even a thin wire antenna, and even new power lines may have small breaks or weakened spots in the insulation.

Prior to starting the project, sketch the installation area to scale, as shown in Fig. 31.1, and draw boundaries of safe working areas based on the longest and tallest masts, towers, ladders, cranes, or other structures that could fall, break, or tip over. Include the area around any electric utility drop wire from the power pole to the building as part of the forbidden zone. Use stakes and colored ribbons or string to separate the "forbidden" areas from the work area. Avoid working near ravines, structures, electrified fences, and other hazards.

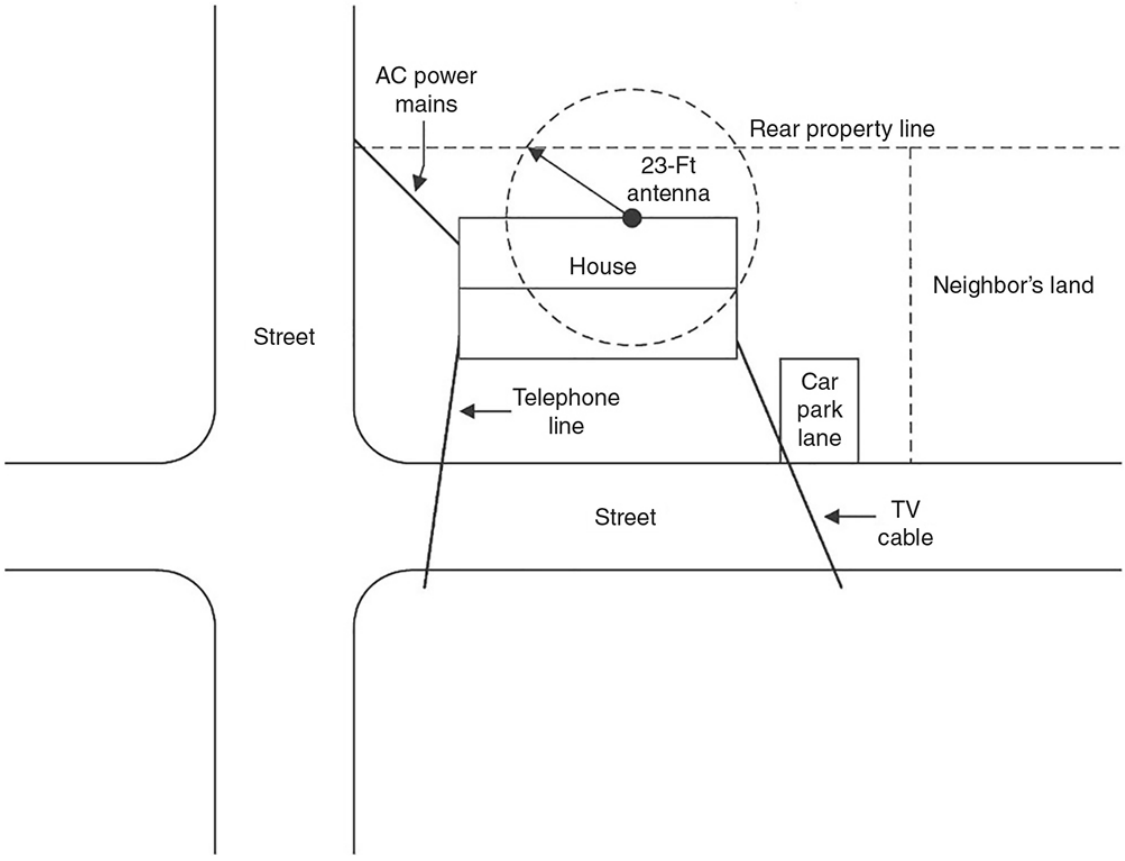


FIGURE 31.1 Antenna installation on suburban lot.

Make a list of all tools and machinery you expect to need for the project, and make sure you either own them, can rent them, or have friends who will bring or loan them. One of the leading sources of personal injury and property damage during antenna projects is failure to use the right tool for the job at hand. If electricity is required in the work area during the project, make sure that adequate three-wire 115 VAC power with functioning ground fault interruption (GFI) circuitry is available at the site. Most municipalities allow GFI protection to be inserted in one of two places: as part of the circuit breaker for the circuit in question or as part of the first outlet on that run from the breaker. GFI installation should be handled by an electrician licensed to work in your community.

If climbing is involved, make sure experienced climbers with all the appropriate climbing gear are involved. Any climbing at all, even on ladders, can be taxing. Many common antennas, including small Yagis or quads, are deceptively lightweight on the ground, but when you get up in the air even for a short distance, they become remarkably difficult to handle, especially in the presence of even light breezes.

Wind speeds and wind loads increase with height above ground, especially if there are buildings, trees, and other obstructions to slow the wind below their tops. The wind has a bad habit of coming up at the worst possible time—usually when you're about to attach the antenna to its support mast. Never tackle an antenna project with no one else around. Even if it is a project that can easily be handled by one person (stringing a dipole, for instance), make sure a family member, friend, or coworker is on the premises, within earshot, even if he or she is doing other things. One good rule is to always work under the "buddy system." Invite as many friends as are needed to do the job safely, and always have at least one assistant, even when you think you can do it alone.

There is no substitute for planning ahead. Taking time before the day of the project to define and think through the many individual tasks that together result in a successful antenna installation is the single most important thing you can do before you tackle the job or

your crew of helpers arrives on the scene. Knowledge of what you face, coupled with hard-nosed, sound judgment and some common sense, are the best tools for an antenna job.

Install only antennas of the best materials and workmanship in order to minimize the need to repair or replace them. It is not just the electrical or radio reception workings that are important but also the ability to stay up in the air in an often unfriendly environment.

When planning an antenna job, keep in mind that pedestrian traffic in your yard possibly could affect the antenna system. Wires are difficult to see; if an antenna wire is low enough for vehicles, people, or animals to come in contact with it, injury and damage can result. Take care for not only your own safety but also that of others.

Consider a typical scenario involving a four- or five-band trap vertical antenna (Fig. 31.2). It will be 18 to 26 ft tall and will be mounted on a roof or mast 12 to 30 ft off the ground. The total height above ground will be the sum of the two lengths, perhaps 30 to 60 ft overall. You must select a location at which a 30- to 60-ft metal pole can be erected—and can fall—safely. This requirement limits the selection of locations for the antenna. If the drop wire from the power pole to your house attaches to the house high on an end wall rather than coming in underground, you probably will not be able to find a proper location for a house-bracketed or roof-mounted antenna of that overall height.

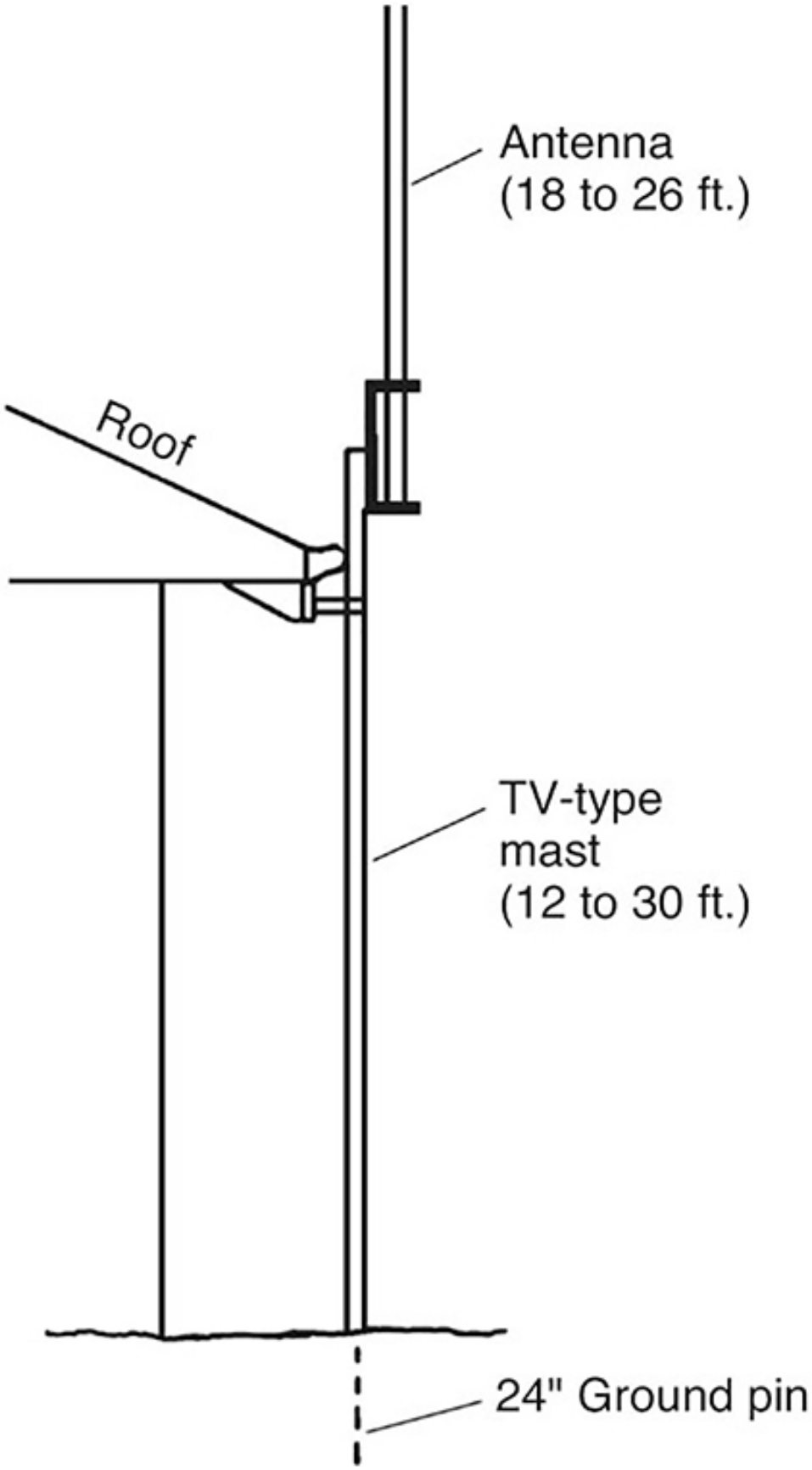

Antenna
(18 to 26 ft.)
Roof
TV-type
mast
(12 to 30 ft.)
24" Ground pin

FIGURE 31.2 Mast-mounted vertical antenna.

When installing a vertical antenna, especially one that is not ground-mounted, make sure that you have help. It takes at least two people to safely install a standard HF vertical antenna, and more may be needed for especially tall or heavy designs. Wrenched backs, smashed antennas, crunched house parts, dented vehicles, and other calamities are rare with a well-organized work party that has sufficient hands to do the job safely.

31.2 Wire Antenna Construction

Wire antennas come in many forms, but good techniques for construction and support are similar regardless of the details of the antenna. In its most common configuration (Fig. 31.3), the HF $\lambda/2$ dipole, for example, is insulated from other conducting materials at each end and supported by ropes, cables, or other wires attached to trees, buildings, masts, or any other structures of suitable strength and height in fixed positions.

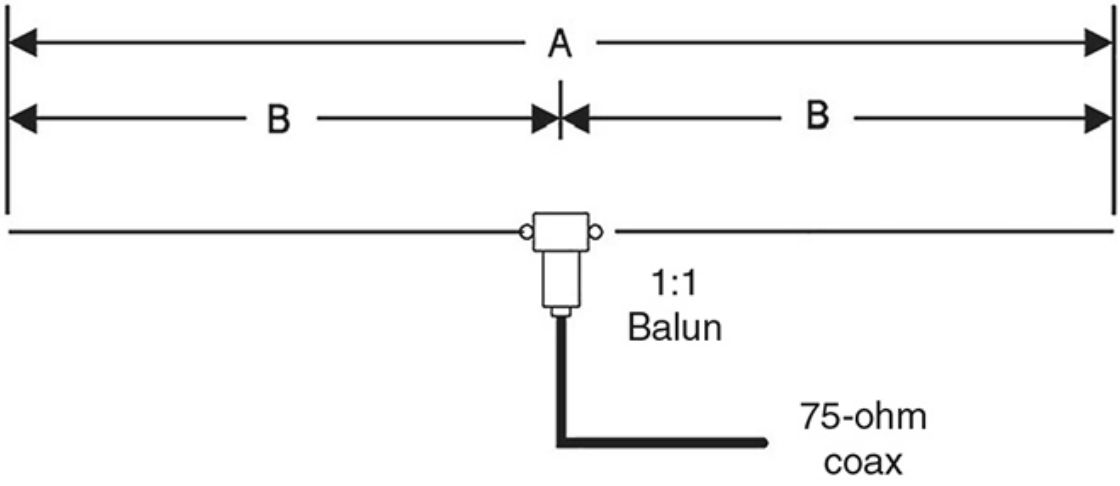


FIGURE 31.3 Typical center-fed wire dipole.

31.2.1 Center Insulators and Feedline Connection

The $\lambda/2$ dipole's nominal feedpoint impedance at resonance is 73 Ω in free space and at multiples of a quarter-wavelength ($\lambda/4$) above ground, so it is a reasonably good match to standard 50- or 75-Ω coaxial cable. When fed with coax, a 1:1 balun transformer or common-mode RF choke is often inserted at the feedpoint.

Far more important to the long-term success of a horizontal wire than a balun are the quality, strength, and weather resistance of the electromechanical connection between the antenna and the feedline in the face of prolonged exposure to the environment. In most installations, this junction is subject to temperature and humidity changes, precipitation of all kinds including ice and snow in many climates, ultraviolet rays from direct sunshine, continual vibration, and abrupt changes in tension resulting from gross motion in the wind.

A common (but poor) practice is to strip the insulation back a few inches at one end of a coaxial cable, part the braid and center conductor, and connect them with ordinary electrical solder to each side of the dipole. This type of joint can crack in a windstorm or degrade over time. Unprotected solder joints exposed to the elements eventually turn gray, brittle, and powdery and eventually fail under environmental stresses that conventional solder was never designed to resist. Solder for electrical and electronic circuits is not intended to provide a strong mechanical connection nor is it meant to be exposed to the elements. Equally disastrous, coaxial cable that is open at one end is very hard to protect from moisture ingress. Over the years, there have been many reported cases of water dripping from the indoor end of cables in the radio room.

Even if you elect not to use a balun transformer, a ready-made center insulator that accepts a PL-259 or similar connector helps minimize or eliminate many of these problems. Figure 31.4 shows a common form of center insulator for use with dipoles and other wire antennas. At the bottom is an SO-239 coaxial cable receptacle,

shielded from direct rainfall by the assembly above it. Eyebolts on the sidewalls provide mechanical strain relief for attaching the two halves of the dipole to the insulator. Although this particular style of center insulator is a compromise because it includes solder connections, if enough slack is provided in the pigtails that attach to the solder lugs, the actual solder joint is under little or no mechanical strain and is less likely to fail from stress-related mechanisms than the old-fashioned junction described above. Figure 31.5 shows a similar center insulator with a self-contained 1:1 balun.

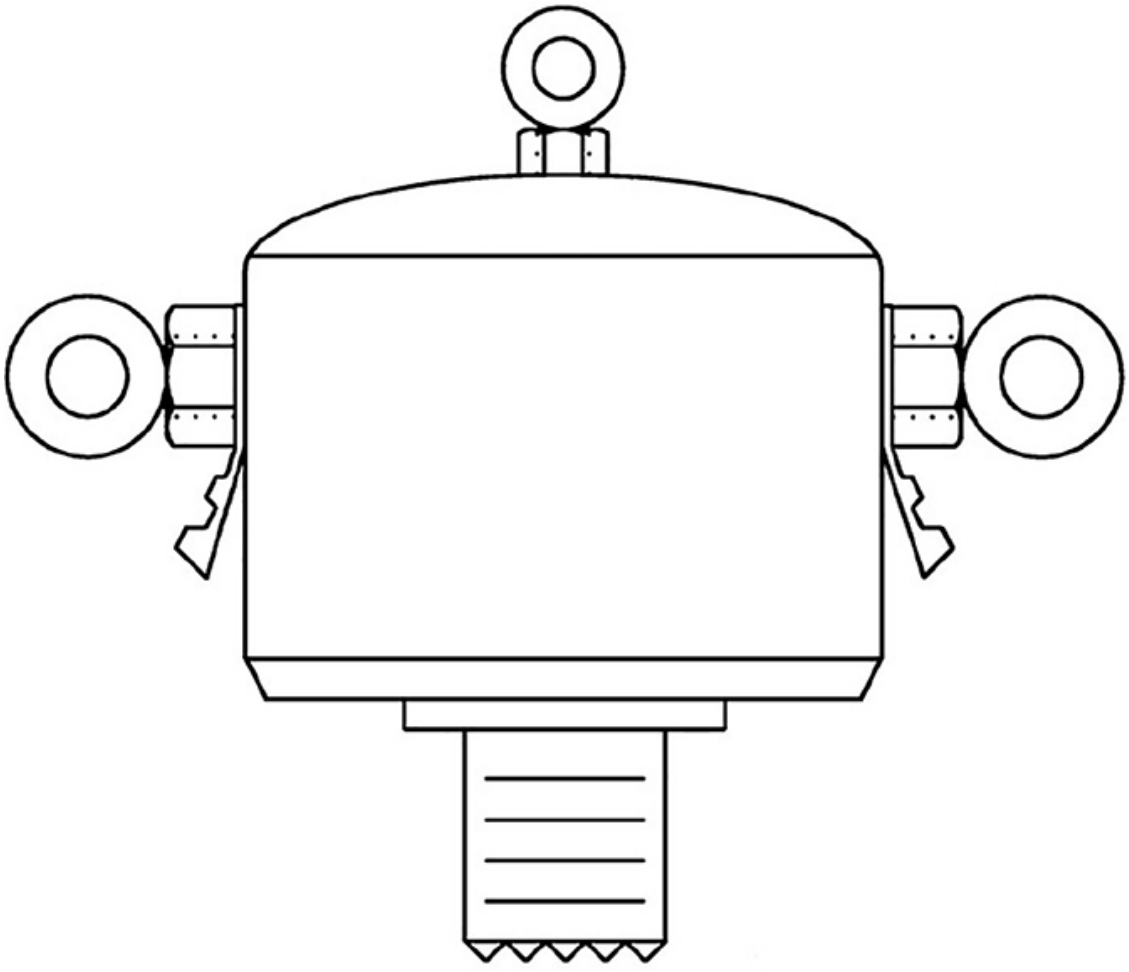


FIGURE 31.4 Center insulator.

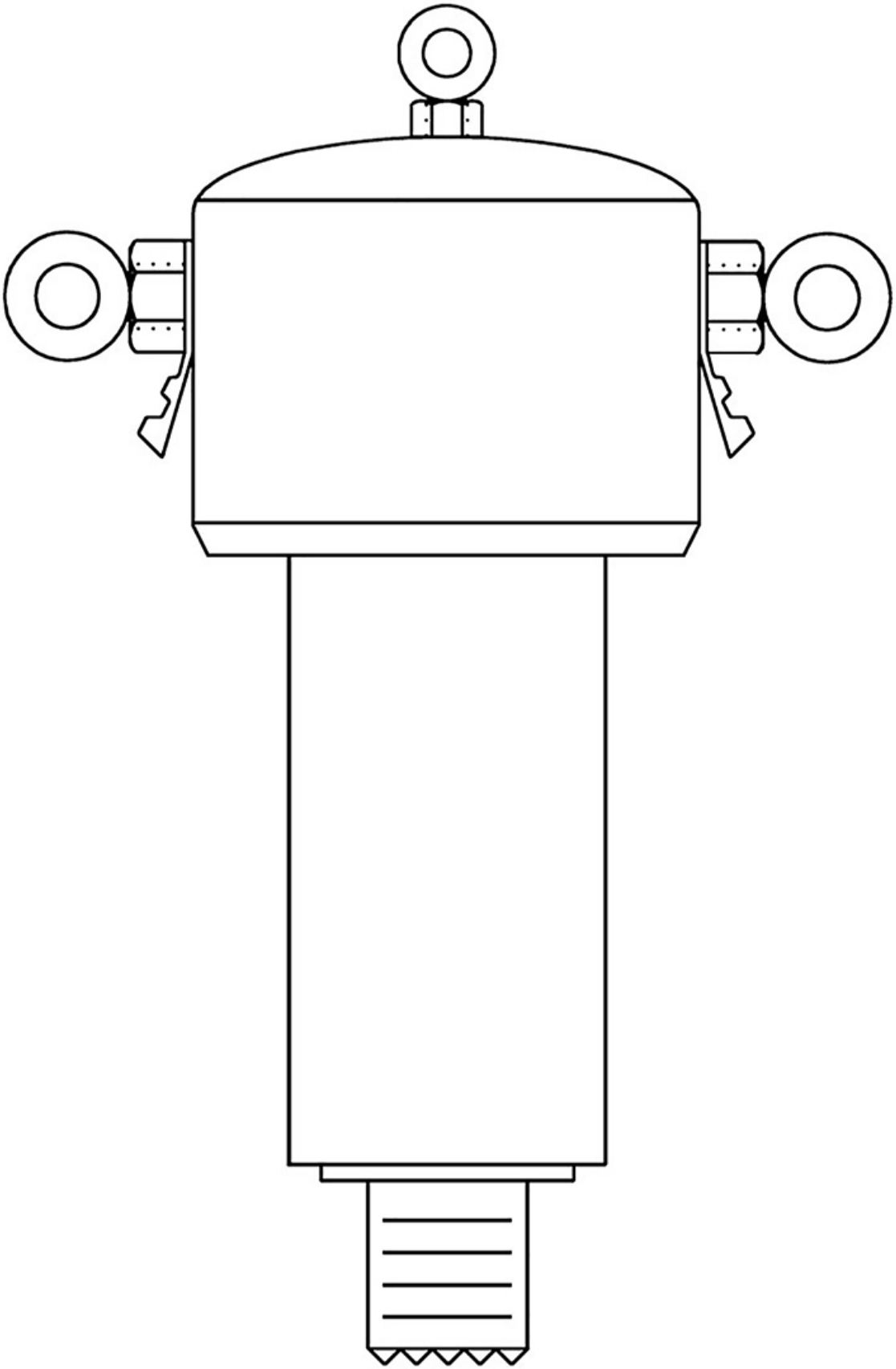


FIGURE 31.5 Center insulator with built-in balun.

A recommended way to connect a wire dipole to the device is shown in Fig. 31.6. For simplicity, only one side is shown, but the other side is identical. The

$\lambda/4$

antenna segments (length B in Fig. 28.3) are typically made of #12 or #14 antenna wire, but wire as large as #10 and as small as #18 for transmitting applications may also work. The exact gauge is not important except that larger-diameter wire is heavier and sags more and very thin wire ultimately is not strong enough to support the weight of the center insulator and feedline. In years past, stranded wire was often used, but oxidation of the strands inevitably occurs and may result in some additional loss of transmit signal and the application of large RF voltages to oxidized junctions, so it is no longer favored. Conventional house wiring (THHN) found at electrical supply stores and discount home suppliers can be used with or without its insulation, but be aware that it is soft-drawn copper and can stretch in response to excessive tension such as that encountered during windstorms by dipoles strung from trees without provision for the movement of branches. Most antenna tuners can compensate for the changing impedance caused by long-term stretching but occasional retensioning of the end support lines may also be required.

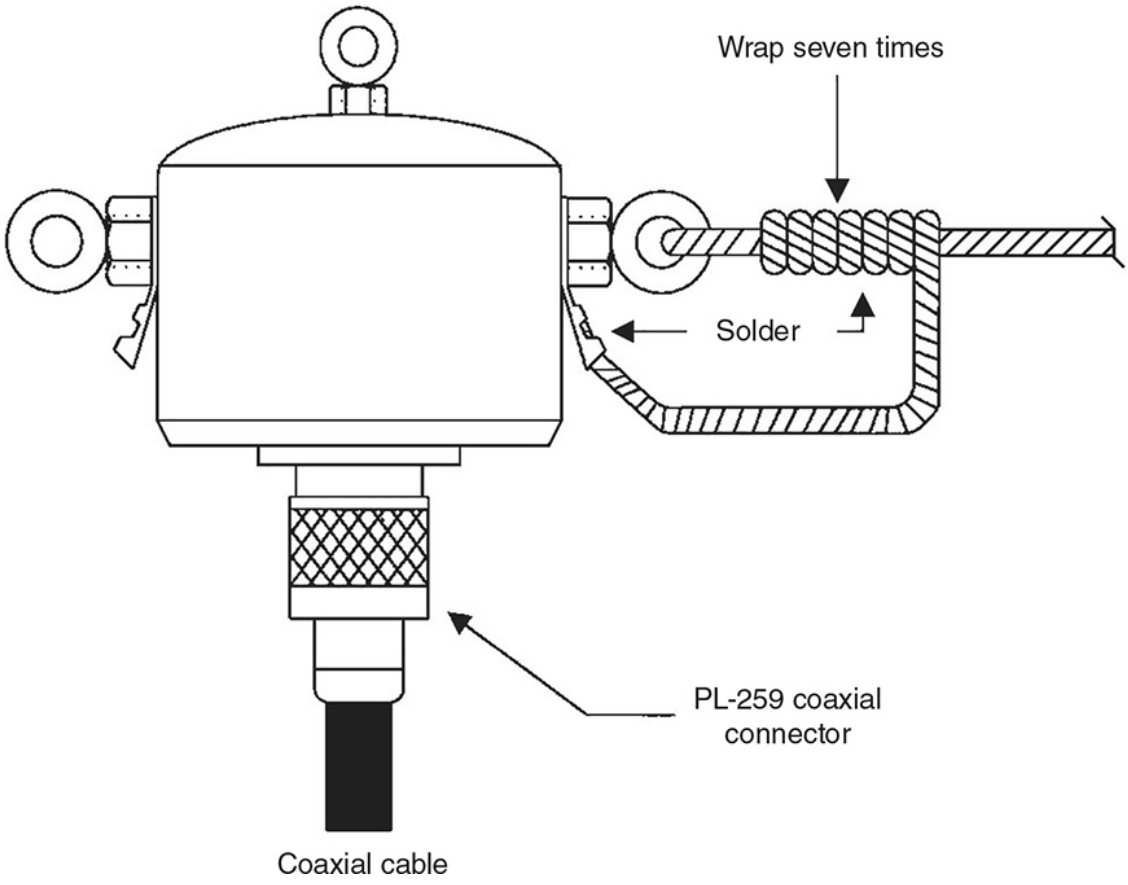


FIGURE 31.6 Center insulator connections.

To avoid stretch, some recommend copper-clad steel core wire (Copperweld is one well-known trade name), but others have encountered problems with voids in the copper, pitting and rusting, and ultimately failure of the wire from poorer quality wire. Copper-clad steel is quite springy and difficult to work with unless it is always kept under tension once unreeled.

Pass one end of one side of the dipole through an eyebolt and then double it back and twist it onto itself four to six times to ensure that no slippage of the wire through the eyebolt will occur when the antenna is erected and put under tension. Bare an inch or so of the wire at the end of the pigtail with steel wool, sandpaper, or a utility knife until the copper is clean and shiny all around the bared section. Form the short section of wire between the eyelet and the solder lug into a drip loop, and insert the end of the wire into the shank of the solder lug. Be sure to crimp the wings of the

solder lug around the wire end with a pair of pliers. While keeping the assembly perfectly still, solder the terminal shank and wire with solder having a resin core, not plumber's solder. Some people prefer to use a completely separate wire as the pigtail, but that doubles the number of solder joints. Whichever approach is used, the intent is that there be no tension at all on the pigtail when the antenna is up in the air. These steps are best performed with the dipole antenna laid out on the ground in a straight line, preferably (weather and tools permitting) in or near the area where it will be hoisted.

Note *Use only resin-core solder for antennas. Typical solder is 50/50 or 60/40 lead/tin ratio and may be labeled "radio-TV" or "electronic" or something similar. Solder alloys without any lead are available. Acid-core plumber's solder will destroy your connection within a short time.*

The life of the solder joint can be extended by painting it with nail polish or spraying it with an acrylic lacquer such as Krylon, after the joint has cooled. As with the soldering process itself, the entire assembly should not be moved or jostled until the coating has dried.

Oxidized metal cannot be soldered. Even new copper wire needs to be sanded or scraped. Bare copper wire may have a clear enamel insulation on it that isn't obvious to the eye. But all bare copper oxidizes, forming an insulating layer over the conductor. The early stages of that process result in only a slight dulling of the surface, so cleaning off the wire immediately prior to soldering it is an excellent habit to form. When preparing the center connector, it is good practice to use steel wool or sandpaper on the solder lugs to remove any oxidation that may have formed over time.

An SO-239 UHF coaxial receptacle (socket) on the bottom of the center support device shown in Fig. 31.6 accepts any feedline terminated in a standard PL-259 plug. Once that connection is made, wrap high-quality black vinyl electrical tape around the joint, starting on the vinyl cable jacket below the plug and continuing

completely over and past the PL-259 until reaching the molded case that supports the SO-239 receptacle. Half-lapping the tape like siding or roof shingles should ensure minimal water ingress at the end of the coaxial cable. To minimize water vapor ingress caused by changes in humidity, the vinyl tape should be sprayed with acrylic lacquer or painted with nail polish. Alternatively, use heat-shrink tubing and a heat gun to totally encase the connection from the threads of the SO-239 above the PL-259 to the outer jacket of the cable a half-inch or so below the bottom edge of the PL-259.

In most installations, the weight of the feedline will keep the entire center assembly oriented, as shown in Fig. 31.6, thus eliminating the need for a drip loop on the coaxial cable. If a specific installation results in the SO-239 connector being anywhere other than beneath the body of the device, the coaxial line near the connector should be routed with a half-loop that is arranged so as to send any precipitation running along the cable away from the connector.

At the top of the center insulator a small eyebolt is provided for strain-relief of the system or for suspension from a center support (such as when the antenna is an inverted-vee). One disadvantage of coaxial cable and commercial center insulator assemblies (as compared to open-wire line directly connected to the two sides of the dipole) is that both add quite a bit of deadweight to the center of the antenna. Unless some kind of auxiliary support (nylon twine, polypropylene rope, or small-diameter aircraft cable pulling up to a high point on a tower, mast, or tree) is available, the sag in a dipole for 80 or 40 m fed with RG-8 or RG-213 cable may be unacceptable to the user. If the VSWR on the line is kept close to 1.0:1 or power to the line is kept substantially below 1000 W, some weight can be removed by using the smaller-diameter RG-58 or RG-59 cables instead of RG-8—especially at 10 MHz and below, where line loss is lowest.

The longer, lower barrel of the center insulator shown in Fig. 31.5 indicates the presence of either a 4:1 or 1:1 impedance ratio balun transformer. The 1:1 size is recommended for ordinary dipoles, whereas a 4:1 unit is a good match for folded dipoles and

certain other wire antennas with feedpoint impedances substantially larger than the

Z_0

of the feedline.

Some users form the coaxial cable into a 6- to 12-in multiturn loop just below the coaxial connector, as shown in Fig. 31.7. This forms a common-mode RF choke intended to keep currents from flowing on the outside of the cable shield, thus helping to keep the feedline from becoming an unintended part of the radiating antenna. The number of turns is inversely proportional to the lowest operating frequency. For an 80-m dipole, 16 turns of 8- to 12-in diameter may be needed. Bind the turns together with black electrician's tape or some similar adhesive medium and then fasten the entire loop to the top eyelet with cord or fishing line for strain relief.

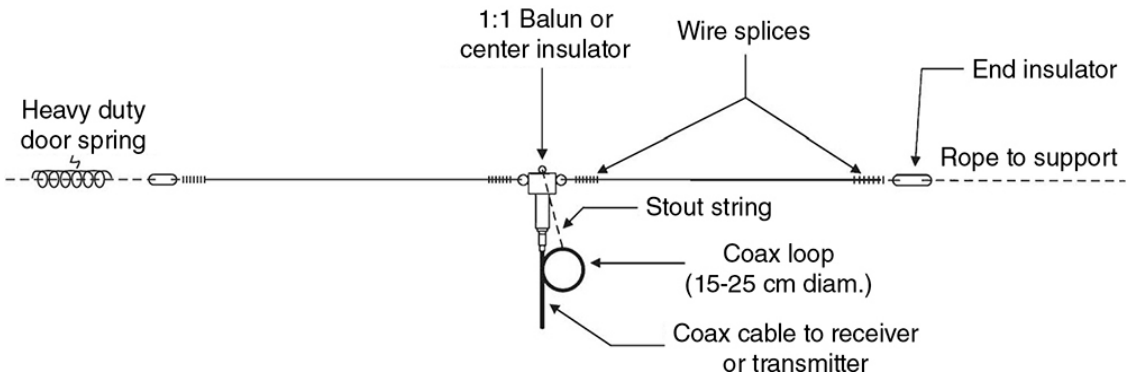


FIGURE 31.7 Dipole and wire antenna installation.

31.2.2 End Insulators

End insulators have been made from wood, glass, ceramics, plastics, and composites. At low power levels (around 100 W or less), the garden-variety polypropylene rope often used for stringing the dipole up is perfectly adequate as its own insulator. Since the wire ends of a dipole are usually high-voltage points, take care to keep the conductors away from flammable surfaces. More than one licensed amateur has set a tree on fire by transmitting it into a wire draped across tree limbs.

End insulators come in a large variety of shapes and sizes. Examples of two basic shapes are found in Fig. 31.8. Figure 31.8A depicts the standard strain insulator, and Fig. 31.8B is representative of an egg insulator. The egg insulator has two sets of wire grooves and through-holes that are orthogonal to each other (only one set is shown). The wire passes through one hole and its grooves, while the supporting rope passes through the other hole/groove set. The purpose of the orthogonal holes is to prevent the antenna from totally separating from the support line if the insulator should fracture.

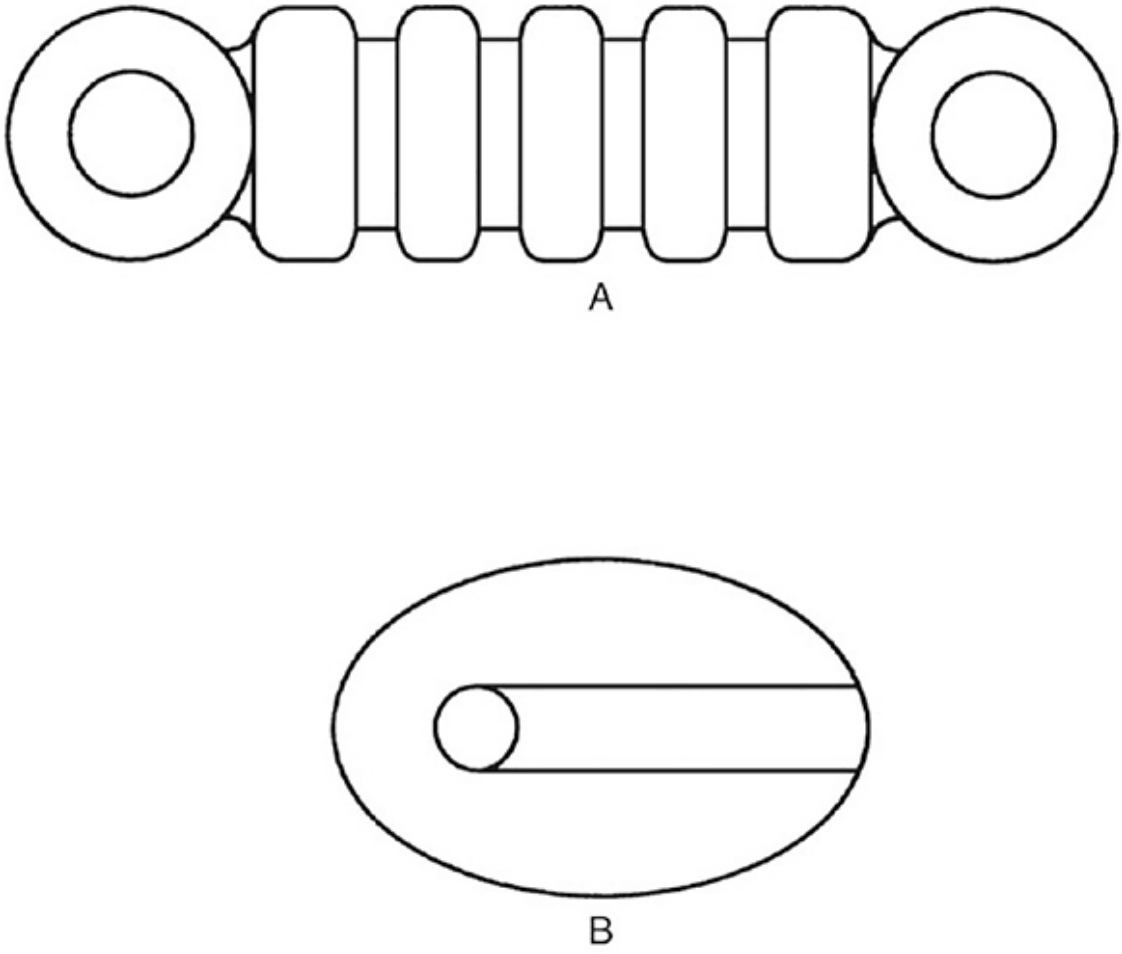


FIGURE 31.8 End insulators.

For higher power levels, the standard strain insulator of Fig. 31.8A is preferable. This style often has a number of grooves or

ridges along the body that serve to lengthen the effective electrical path between the two ends of the insulator, thus reducing losses in the insulator while increasing the breakdown voltage between its two ends. This is particularly important in locations where airborne contaminants gradually collect on the surface of the insulator.

Other materials for fabricating strain insulators include nylon, plastics, and treated hardwoods. Short lengths of hardwood stock can be drilled at each end for wire or rope, then treated with an exterior urethane or other protectant.

Strictly speaking, the egg insulator in Fig. 31.8B is not designed to be an end insulator for transmitting antennas. The distances between the antenna wire and the support line are small and subject to breakdown at high power, since at least one end of a wire antenna is a high-voltage point. Egg insulators are okay to use at low power levels or for "receive only" and temporary installations, but so is polypropylene rope. On the other hand, the egg insulator is potentially much stronger than a comparable strain insulator because the interweaving of the wire and/or rope through their respective holes puts the egg in compression rather than tension. For that reason, the egg insulator and its rectangular sibling seen on power pole guy wires are the only types to use when it is necessary to electrically break up long guy wires into multiple shorter sections.

Figure 31.9 shows how to make connections through an end insulator. The electrical connections to the antenna wire are the same as for the center insulator. For transmitting antennas avoid leaving any sharp points sticking up. The ends of a $\lambda/2$ dipole experience high RF voltages, and sharp points tend to cause corona sparking. Cut off any extraneous leads that form sharp points, and then smooth the whole affair down with solder tinning.

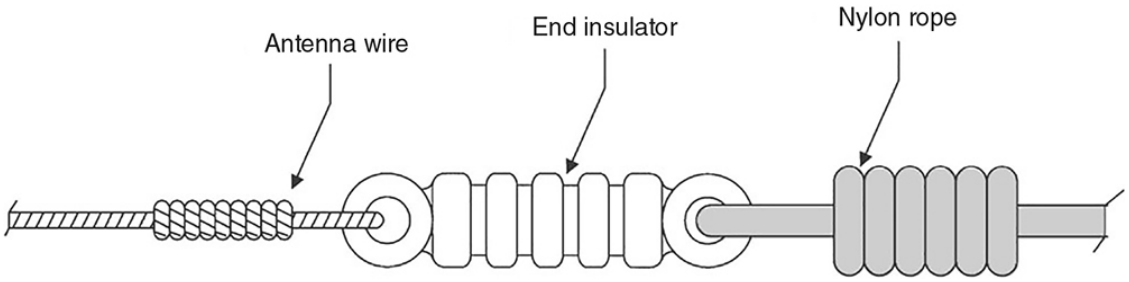


FIGURE 31.9 Connections to end insulators.

The rope is dressed in a manner similar to the wire and is knotted with a self-tightening knot that clamps down on itself as the support line is tensioned. Some use the hangman's knot. The bowline is useful if the line is always in tension, as the knot can be broken and untied later.

31.2.3 Folded Dipole Construction

As discussed in Chap. 8, the folded dipole (Fig. 31.10) consists of two close-spaced conductors shorted together at both ends, but only one of the conductors is split in the center for the feedpoint. The free-space feedpoint impedance at resonance is around 290 Ω, so it makes a good match to 300-Ω, television antenna twin-lead. Often the antenna itself is cut from the same roll of twin-lead. Alternatively (as shown in Fig. 31.10), a 4:1 balun transformer at the feedpoint provides a match to 75-Ω, coaxial cable. A preassembled folded dipole and feedline rolls up into a small storage space, weighs very little, and can be rapidly deployed in an emergency or portable scenario.

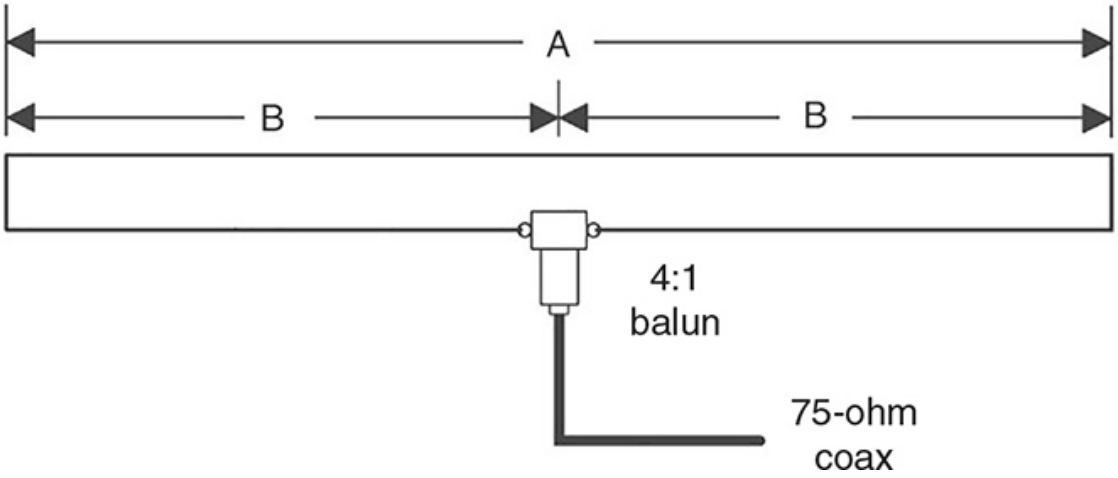


FIGURE 31.10 Folded dipole.

Neither the conductors nor the dielectric of ordinary TV twin-lead are particularly strong, so some ingenuity is required to keep 300-Ω, twin-lead from breaking. Figure 31.11 shows a possible solution.

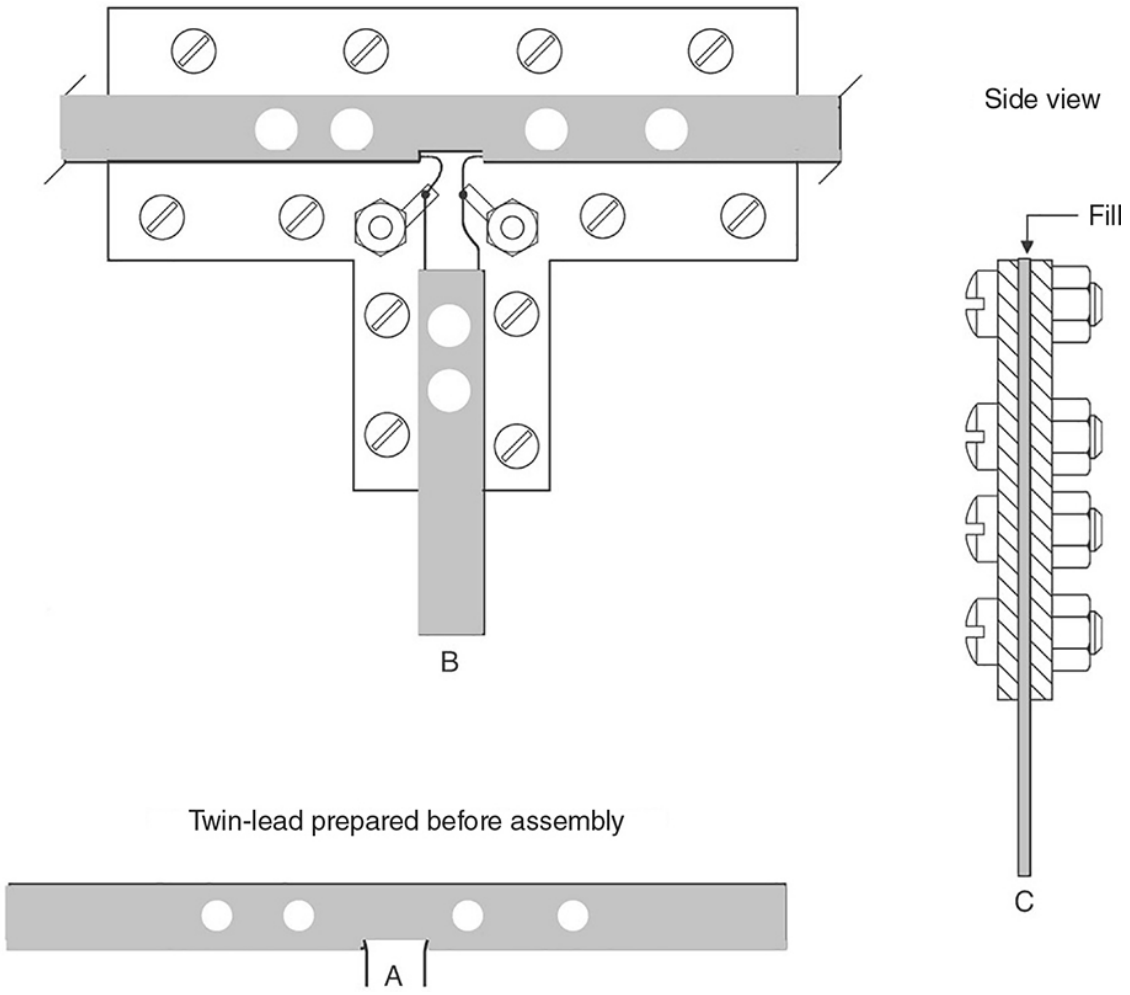


FIGURE 31.11 Center insulator for ribbon cable folded dipole. (A) Cable preparation. (B) front/back view. (C) Edge view.

The 300-Ω TV twin-lead is prepared as shown in Fig. 31.11A. One conductor of the twin-lead is cut at the planned feedpoint, and the insulated wire freed from the center dielectric material a distance of about 0.5 to 0.75 cm in both directions, using a match or small diagonal side cutters. Use the same tools to strip the insulation from the two wire ends just created. A hand punch—either a large-size leather punch or a paper punch— places two or three holes in the center insulation on either side of the break.

The center insulator is shown in Fig. 31.11B. It is made from a piece of strong plastic, Lucite, or other insulating material. Two

identical sections, front and back, are needed. A number of 5-mm holes are drilled into both pieces at the points shown to clamp the twin-lead. A pair of solder lugs provides a mechanical tie-point for connections between the antenna element and the transmission line. The nuts and small bolts that fasten the two halves of the insulator are made of nylon to eliminate any possibility of short-circuiting the two terminals, or the two sides of the twin-lead that form the antenna.

A side view is shown in Fig. 31.11C. Note that the twin-lead causes a gap that can catch water and makes it possible to break one or both insulators by overtightening the nuts. To prevent these failures, a gasket of similar material is glued into the space as filler. The gasket material should be the same thickness as the twin-lead. Rather than burden this antenna with the weight of coaxial cable, the twin-lead feedline should be run to an ATU for balanced lines.

An end insulator for twin-lead folded dipoles is shown in Fig. 31.12. It is constructed in a manner similar to the center insulator. The clamping fixture can be fabricated from metal such as 3- to 6-mm brass or copper stock. The ends of the antenna wires are shorted and soldered together, so use of a conducting material at the ends is not a problem.

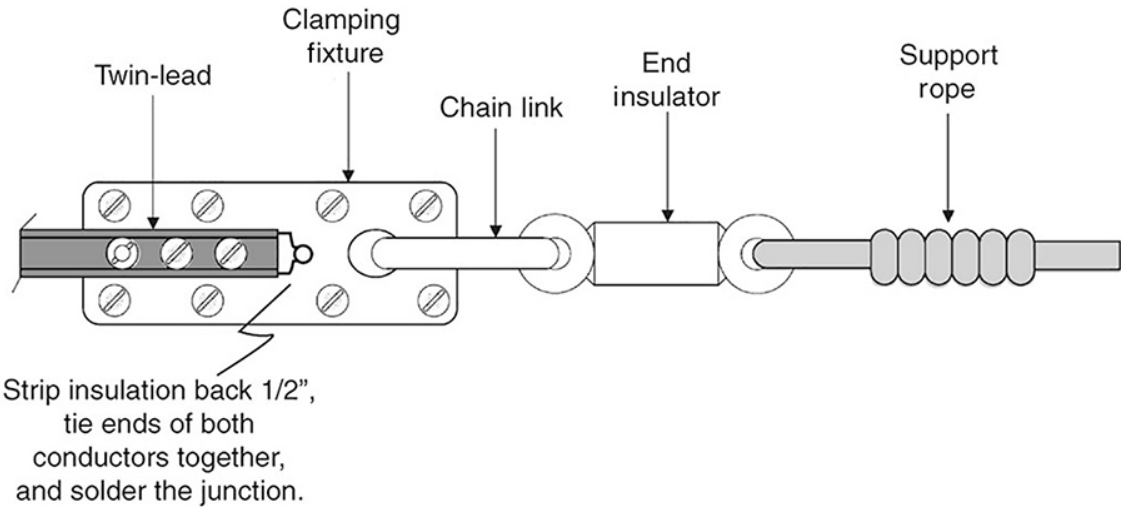


FIGURE 31.12 Twin-lead folded dipole end insulator.

An end insulator is then used with a rope in the normal manner. A chain-link section or carabiner is used to connect the clamping fixture and the end insulator. Alternatively, the clamping fixture can do double duty if it is the same material as the center insulator. In that case, the other end insulator and chain link can be ignored. Bevel and polish the rope hole on the clamping fixture to prevent chafing of the rope, or insert a guy wire thimble of an appropriate size through the hole in the fixture.

31.2.4 Supporting the Wire Antenna

Figure 31.7 shows the dipole installed. The ends of the two wire elements are connected to end insulators, which are in turn supported by rope to a vertical support (e.g., mast, tree, roofline of the house). Although shown perfectly horizontal here, the actual installation will droop in the center due to the weights of the center insulator or balun, the coaxial cable, and the wire itself.

If either end of the dipole is supported by a tree limb or anything else that can move in the wind, add a door spring just beyond the end insulator in one support line. The strength of the spring should be chosen such that the normal tension in the support lines on a calm day causes the spring to be less than halfway extended. For an 80-m dipole with substantial feedline weight, a residential garage door spring is most likely to be appropriate.

31.3 Trees, Masts, and Other Supports

The end supports of the antenna can be anything that supplies height and is strong enough: a mast, a tree, or the roofline of a building. Figure 31.13 shows the use of a mast, but other forms of support can be used as well. To facilitate easy raising and lowering of the antenna for maintenance and tuning, the support rope should not be tied off at the top, as is true in all too many installations; instead, it should be brought down to ground level, just like the halyard of a flagpole. Be sure to provide enough slack to make lowering the antenna feasible. Once a halyard is created, it

can be lengthened temporarily, as needed, to allow the antenna to be pulled sideways (i.e., along its axis) during the raising or lowering process.

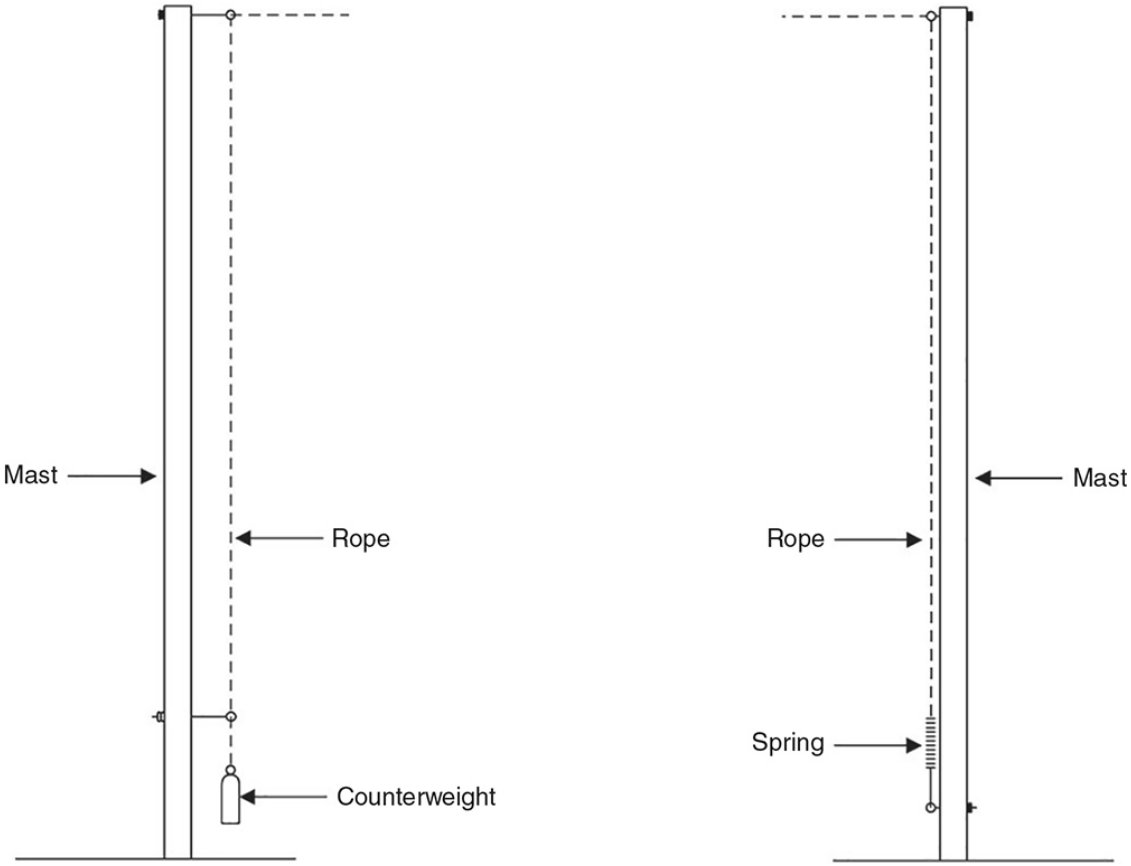


FIGURE 31.13 Mounting mast schemes.

The ideal support is arguably a climbable tower or a building with an elevator to the roof, but most dipoles and other wire antennas are strung from tall trees. To throw a line over a tree, a baseball with a hole drilled through its center can be used. Tie one end of some light nylon twine through the center and a half-circumference of the ball. Lay out the twine on clear ground in front of you before throwing the ball.

Some techniques use a projectile trailing fishing line behind it as it is launched. The fishing line can be wound on a fishing reel. Attach the reel near one end of a 2- or 3-ft length of wooden dowel and push the other end into the earth at about a 45-degree angle

so it is aimed up high in the targeted tree. Once the line is up and over the target tree, attach nylon twine to the far end and reel the fishing line back in. Then attach the halyard or an intermediate line to the end of the twine that has been pulled through the tree top. Many people have had success with a slingshot, or a bow and arrow with the tip removed and added weight.

For greater heights and better accuracy, borrow or purchase a bowfishing rig. The most important element is the compound bow with a tapped hole and matching metal post for direct mounting of a spinning reel. To get the fishing line to drop straight down the back side of the targeted tree, halt the flight of the arrow just after it passes the top of the tree with a slight turn on the handle of the fishing reel. Always make sure the area around you and well beyond the targeted tree is clear of humans and animals before launching any projectiles. Even blunt-tipped arrows or hex nuts can cause serious injury. If all else fails, pay a tree-climbing professional to place a permanent halyard mount near the top of your chosen trees.

31.4 Halyards and Support Lines

There are many choices for materials to use as a halyard or support line. Of paramount concern is long-term exposure to sun, wind, and moisture. Natural cordage is a poor choice. Nylon rope is too stretchy. Various "poly" compounds, including polyester and polypropylene, are worth considering, as is braided Dacron.

The safe working range for rope is typically less than 15 percent of its rated breaking strength. For an antenna that experiences up to 100 lb of tension during windy weather, a 750-lb-rated rope should be used. Choosing the best diameter for your cordage is a trade-off among required strength, size, and cost of any pulleys and what is comfortable for pulling by hand. For supporting a typical dipole, somewhere around 5/16-in is a good starting point.

Virtually all cordage will eventually unravel at a cut end unless the end is treated. Laid or twisted three-strand rope can be eye-

spliced or back-spliced, but braided ropes will need to have their ends fused (melted), taped, or whipped with twine. Fusing can be done with a soldering gun that has a special tip, a propane torch, or a match.

The antenna support line is connected to a spring on one end (see Fig. 31.7) or a counterweight on the other end. The spring should be stout enough so that it is partly extended on a calm day with the antenna support lines normally tensioned. Ideally, with no wind the spring should be extended less than 50 percent of its total possible expansion. For an 80-m dipole and about 100 ft of open-wire line, both constructed from #10 enameled copper, a stiff garage door spring seems to work best.

The counterweight should be just enough to balance the weight of the antenna for a reasonable amount of sag along the wire—probably between 15 and 40 lb for the 80-m dipole and OWL described in the previous paragraph. As wind causes the supporting tree limbs to move up and down or back and forth, the counterweight rises and falls, reducing the chances of snapping a support line or the antenna wire itself. Locate the counterweight where it cannot harm anyone or anything should the rope break. It's helpful to keep the counterweight in balance at a point no more than a few feet above the ground.

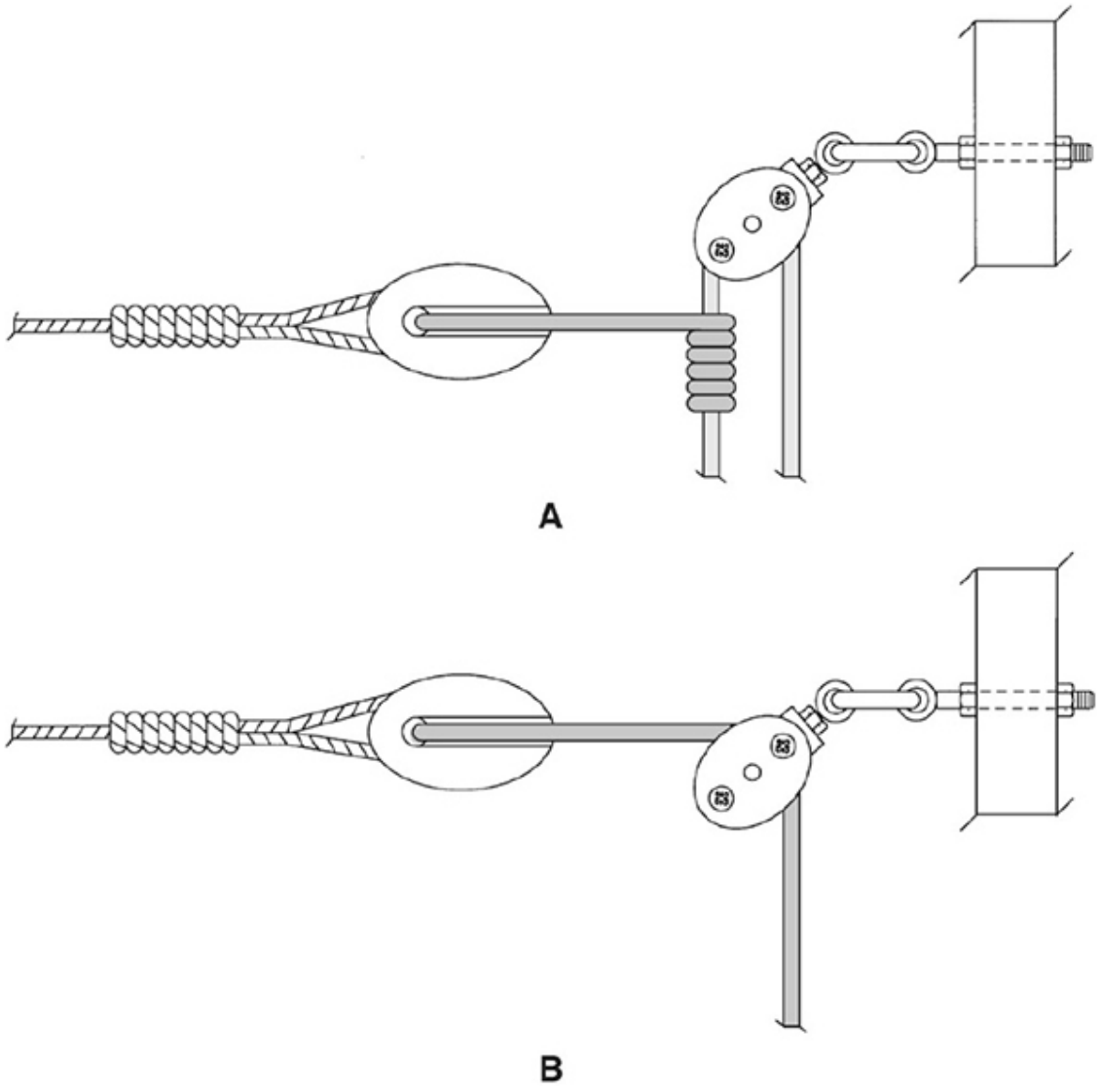
Don't underestimate the value of making the dipole out of ordinary soft-drawn copper (house wiring, in other words). Although you may periodically have to retension the support lines, such wire is more able to withstand the shock of wind gusts snapping the tops of trees around than copper-clad steel is. Stretch in the antenna wire may change its resonant frequency a little bit, but dipoles that are longer than

$\lambda/2$

exhibit a bit more gain in their main lobe.

Figure 31.14 A and B show two methods for making the connections at the top of a supporting mast. Although egg insulators are shown, any kind can be used. A pulley is mounted at the top of the mast with a link section and a stout eyebolt or screw-

eye that passes through the mount. A simple screw-eye will not harm a tree, but avoid wrapping wire or rope around a tree trunk or branch as this can kill the tree. Make the downhaul rope a closed loop so you can raise and lower the antenna.



F IGURE 31.14 Two alternate pulley systems for wire antenna mounting.

A pulley can be used to change the direction of a support line's path. The wrong device will tend to rust if made of ferrous metal or crack if made of plastic. The size of the pulley and the diameter of the cordage running through it must be correctly matched to avoid

jams. The smaller the diameter of the pulley sheave, the shorter the life of the rope that lies in its groove. Figure 31.15 shows a simple replacement for temporary installations or lightweight dipoles. A U-bolt is fastened to the top of the support mast. The rope can be passed through the U-bolt as in the pulley system but without fear of jamming. Install the U-bolt all the way up to the threads so that the rope has no chance to chafe against them.

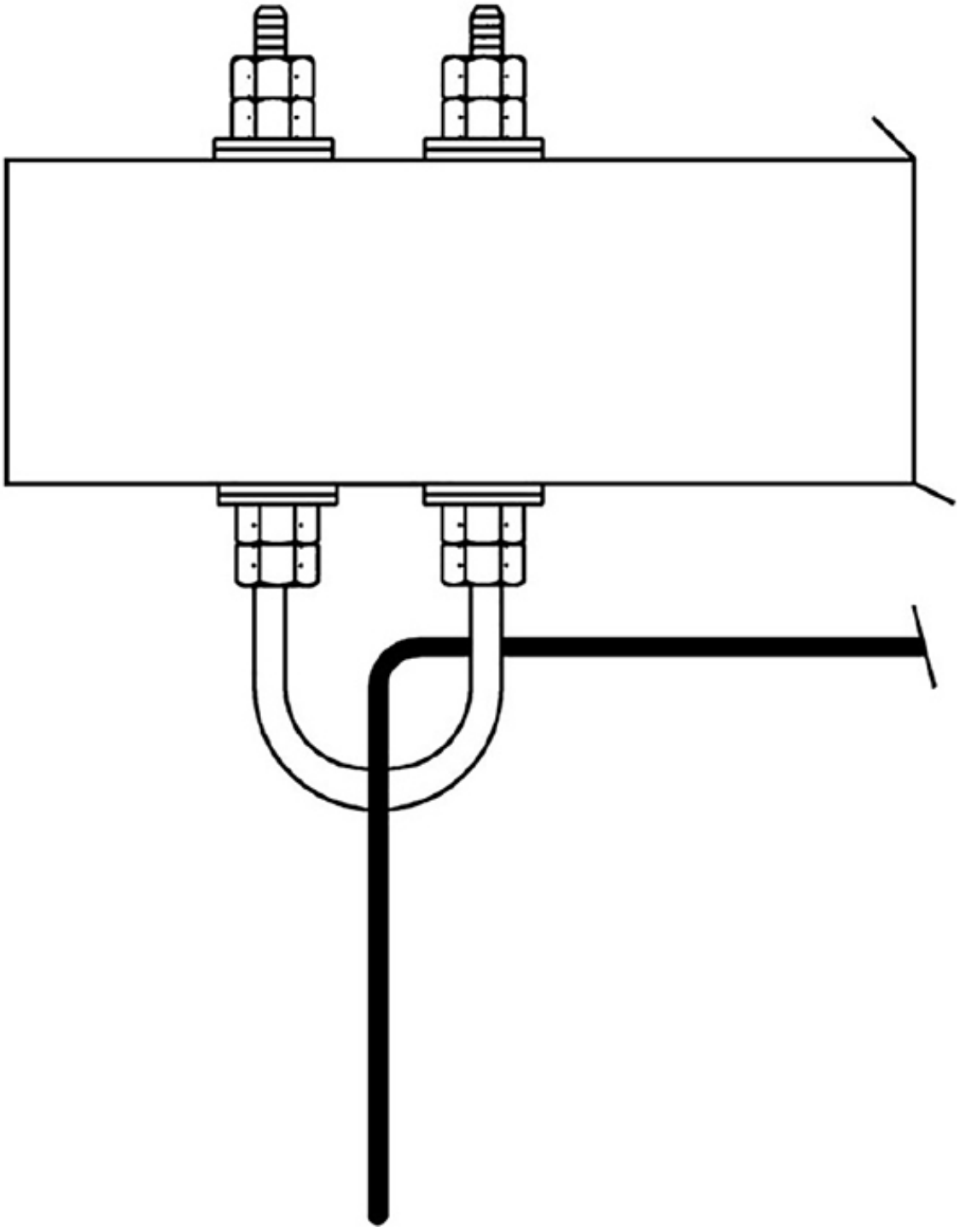


FIGURE 31.15 U-bolt end mounting.

31.5 Strain Relief

For long dipoles with heavy center insulators and coaxial cable feedlines, the ideal installation includes a high center support capable of supplying strain relief and thus minimizing sag at the dipole center. Such a support also reduces tension on the dipole wires, end insulators, and halyards or hoist lines at both ends. Most commonly it is used with inverted-vee antennas, where the center of the antenna is pulled up high on the only available tall support.

Figure 31.16 shows an alternate method of strain relief for the center insulator or balun to reduce total sag from the height of the dipole ends to its center. A strong messenger cable is run between two end supports and tensioned for minimal sag. By suspending the center of the dipole from the messenger, forces on the various electrical and mechanical components of the dipole are substantially lessened. If the messenger cable is a conductor (stranded galvanized steel guy wire material is a favorite), it should be broken up with insulators to prevent detuning the wire antenna. The messenger cable can be run at right angles to the antenna wire using different end supports, in which case the need for insulators disappears.

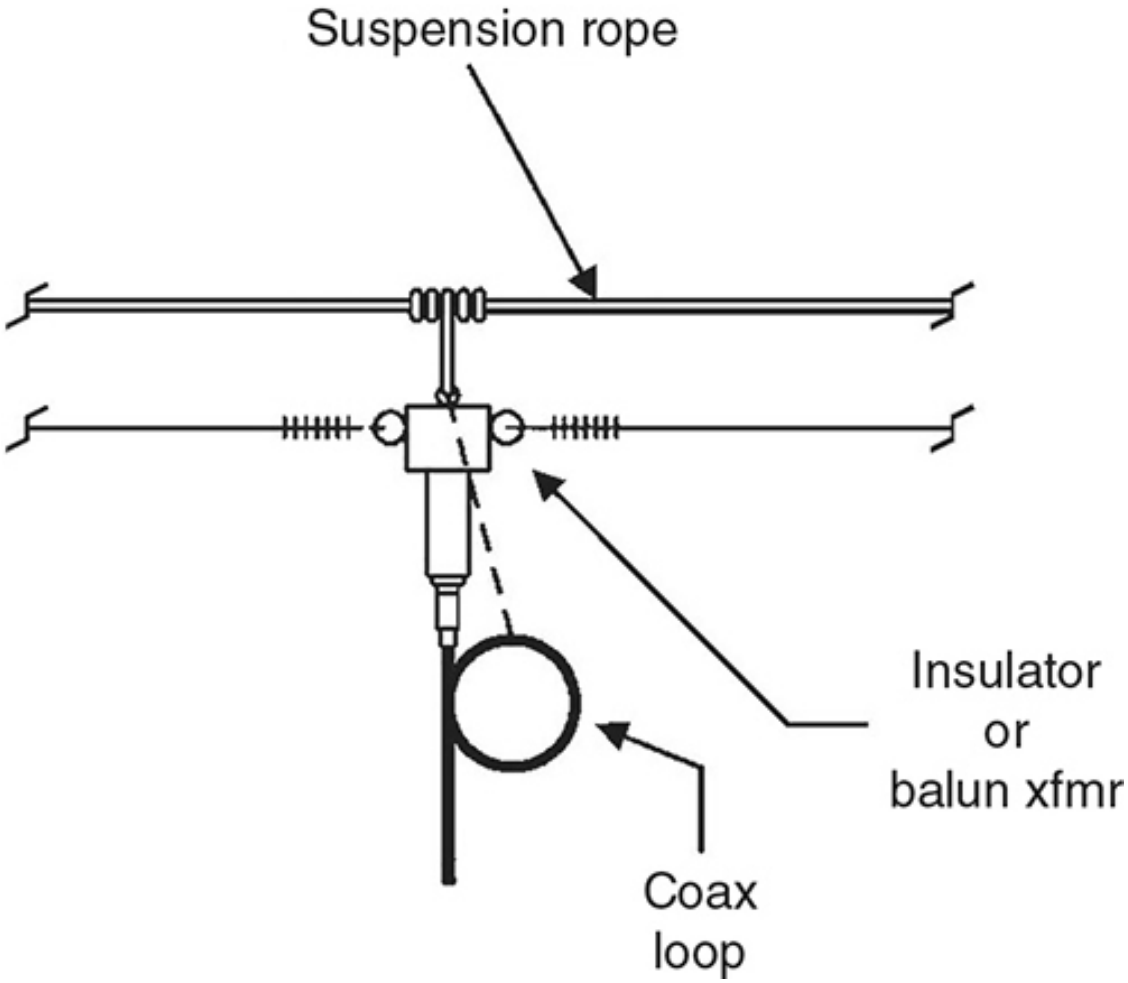


FIGURE 31.16 Strain relieving the center insulator or balun.

31.6 Installing Vertical Antennas

At frequencies above 5 MHz, erecting and supporting HF verticals—both ground-mounted and elevated ground planes—is a relatively easy task for two people, provided the vertical is made of aluminum tubing strong enough to support itself during the brief period when it is being tilted from a horizontal to an upright position. At lower frequencies, verticals should be treated as towers (Chap. 32) or erected with the help of extensive rigging and a larger crew.

The Hy-Gain *14-AVQ* or *AV-14AVQ* has been a popular multiband vertical monopole. By using three parallel-resonant traps and a small capacitive top hat, the antenna provides a physically short

vertical (18 ft overall) that is an electrical quarter-wavelength on 40, 20, 15, and 10 m. (The design is essentially unchanged from the original, pre-WARC band, model, so there is no guarantee of low SWR on the 17- and 12-m bands.) It can be ground-mounted or elevated atop a mast as in Figure 31.17. In either case, the antenna requires radials, but the number and length are different for the two cases. When ground-mounted, at least 16 (nonresonant) radials that are roughly as long as the vertical is tall (18 ft) should be laid out in a circle around the base of the antenna. When elevated, however, three (resonant $\lambda/4$) radials for each band, equally spaced around a 360-degree circle, are sufficient.

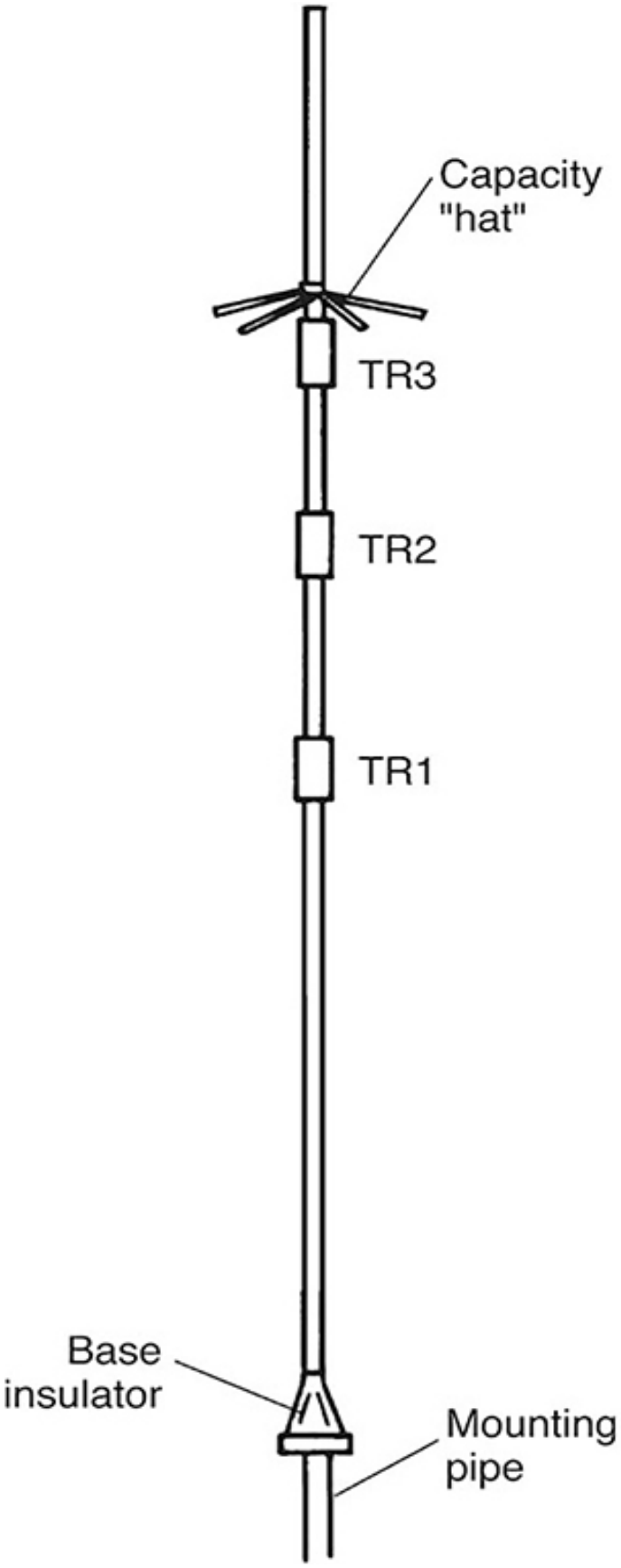

Capacity
"hat"
TR3
TR2
TR1
Base
insulator
Mounting
pipe

FIGURE 31.17 Mounting commercial trap vertical.

Ground-mount installation of the *AV-14AVQ* and similar verticals is simplicity in itself: Drive a 3-ft length of 1 1/2-in OD steel post into the ground, stand the assembled vertical upright, and clamp its base to the post with the U-bolts supplied. Attach radials to the base or add a ring of copper wire or soft copper plumbing tubing around the base as an attachment point. Install a lightning ground (see Chap. 33) of multiple 8-ft ground rods in a circle with a radius of about 8 ft centered at the base of the vertical, and connect each ground rod directly and separately to the base with #6 or heavier gauge wire. Finally, connect 52-Ω coaxial cable at the bottom of the base as shown in Fig. 31.18 or 31.19 and weather-seal the connection with heat-shrink tubing or a half-lap winding of electrical tape (as described earlier in this chapter) and a spray coating of clear acrylic lacquer. Although Hy-Gain does not require or suggest the use of guy wires for the vertical itself, a single set of lightweight, nonconducting guys such as nylon twine or 1/4-in poly rope tied just above or just below the middle trap can be used unless your installation location is such that the antenna is at greater risk from people tripping over the guy ropes than it is from high winds.

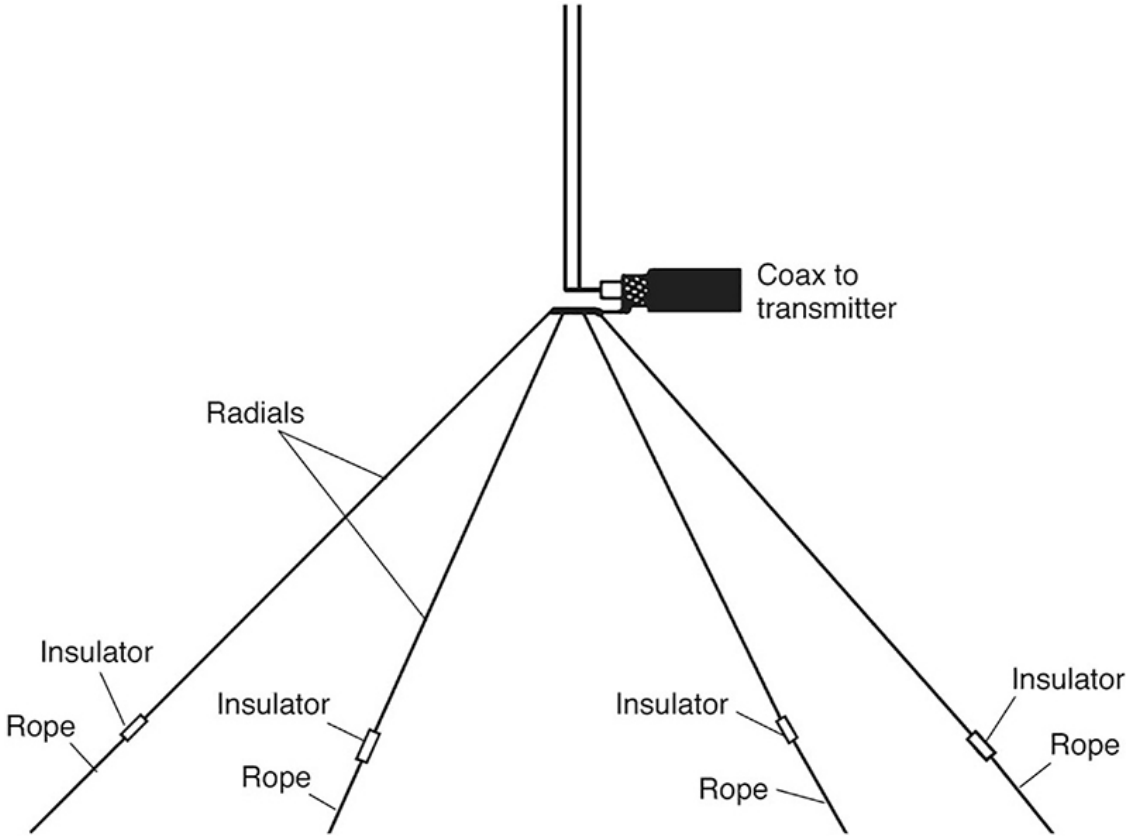


FIGURE 31.18 Feeding the mast-mounted vertical antenna.

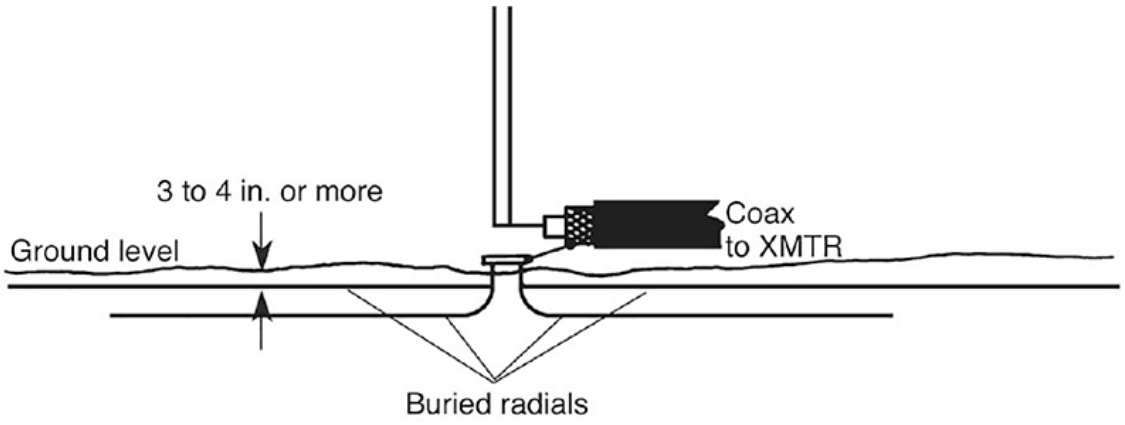


FIGURE 31.19 Feeding the ground-mounted vertical antenna.

An alternative to a telescoping metal mast is a pivoted wood mast, as shown in Fig. 31.20. At a minimum, one set of guy ropes just below the top of the mast is sufficient, although a set of

lightweight guy ropes at the midpoint of the vertical eases wind-loading stresses on the antenna.

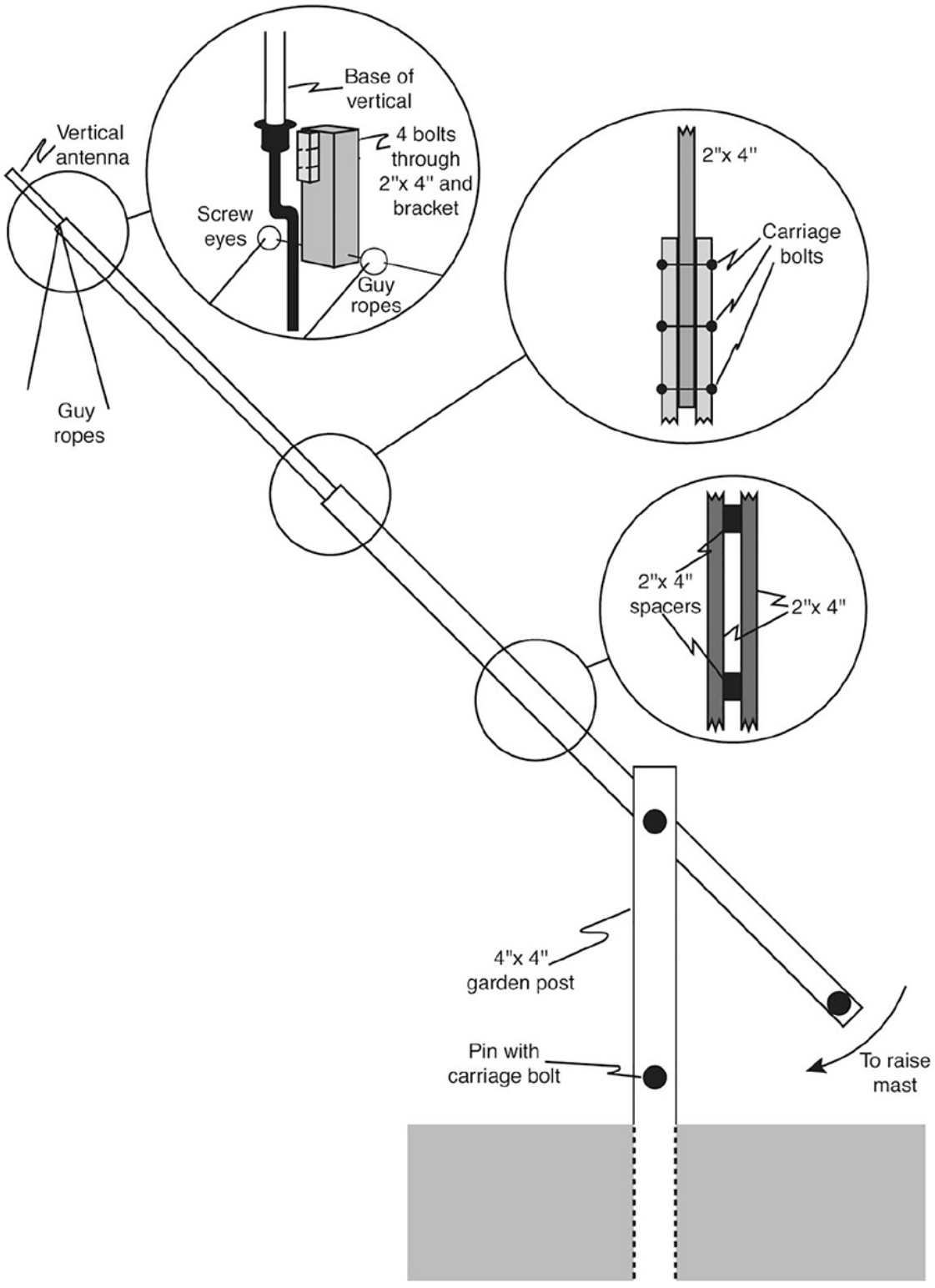


FIGURE 31.20 Pivot mast.

Although a ground-plane vertical does not require any additional RF ground system other than its elevated radial system, it does require a lightning ground at the base of the mast. It is a good idea to also ground the coaxial cable outer conductor there as it comes down from the connector above and heads off to the radio room. A simple way to do that is to use two separate pieces of coaxial cable—one just long enough to run from the SO-239 at the base of the vertical to ground level, and the other to run from the base of the mast to the radio equipment—and join them at the base with a double female (or barrel) adapter mounted on a small piece of aluminum mechanically connected to the center of the lightning ground system. Alternatively, commercial lightning suppressors that house the barrel, a discharge circuit, and a convenient lug for the ground connection can be used.

Because the normal stresses on guy lines for verticals erected in this fashion are much lower than the tension on halyards and other support lines for horizontal wires, smaller-diameter rope is appropriate. 1/4-in poly is about the maximum that would ever be needed for verticals of this size.

When a vertical is placed atop a mast of any significant height, the support needs to be more substantial than just a pipe driven into the ground. Not only the vertical but the mast itself contributes to the total moment arm that exerts a sometimes unexpectedly large sideward or overturning force on the base.

Figure 31.21 shows a standard chain-link fence post used as a support for either a vertical antenna or a mast (which could be metal, PVC, or wood). The typical fence post is 1.25 to 2.00 in OD and made of thick-wall galvanized steel. Such a pipe makes a rugged installation. Some users believe that the pipe will last longer if painted along its entire length (especially the underground portion) with a rust-inhibiting paint such as Rustoleum. If your intention is to directly mount the vertical to the metal fence post, make sure the pipe you choose is a smaller outside diameter than the maximum mast size accommodated by the base bracket and U-bolts of the antenna you're putting up. Otherwise, fabricating a suitable adapter can turn out to be the most time-consuming part

of the entire project. On the other hand, if the fence post is supporting an intermediate mast of some height, the larger diameters are more appropriate.

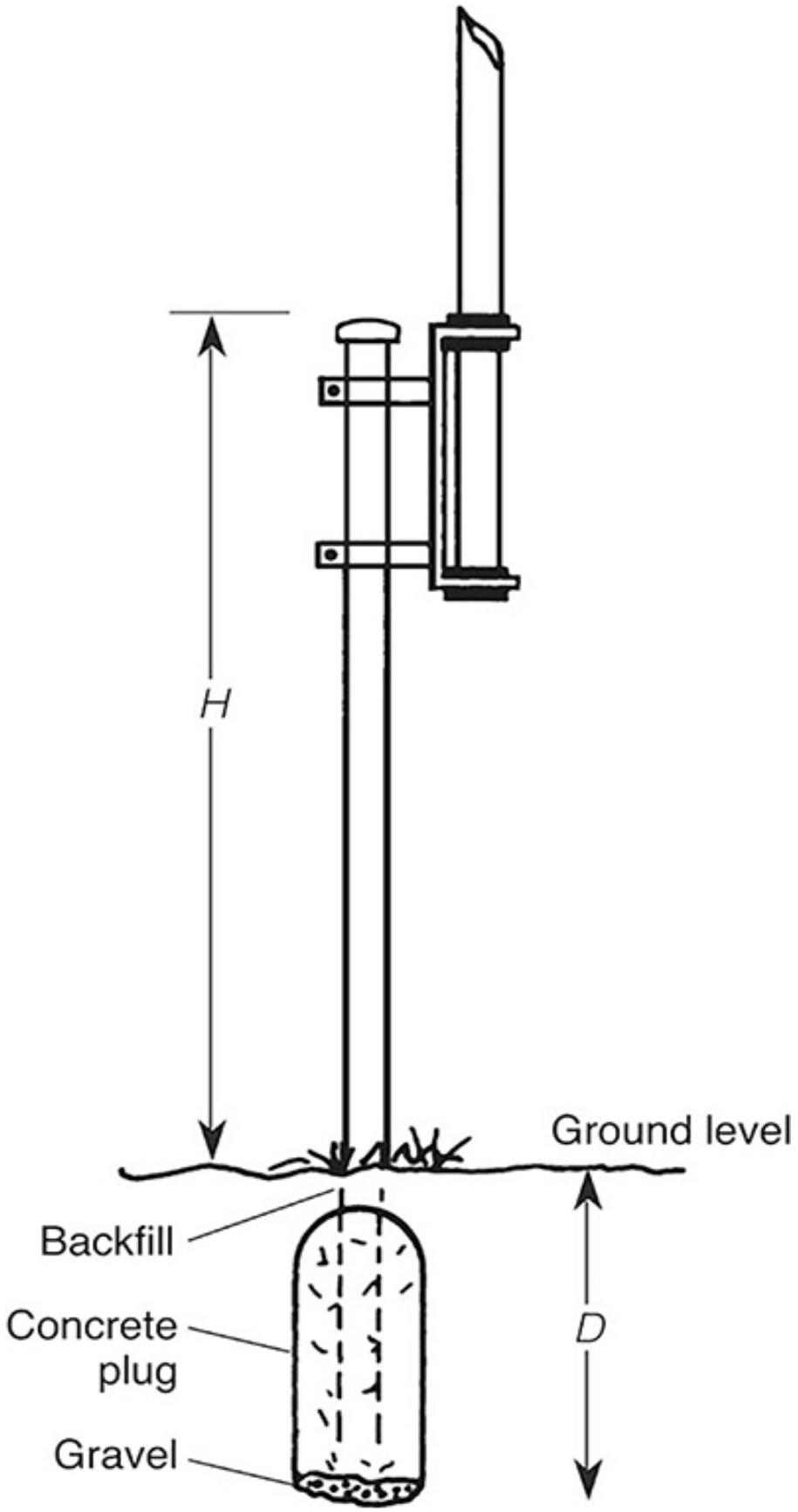

H
Ground level
Backfill
Concrete plug
Gravel
D

FIGURE 31.21 Fence post mounting.

The beauty of the fence post as an antenna support is that it is relatively easy to install, and supplies are obtained easily from hardware stores. The post is mounted in a concrete plug set at depth *D* (in Fig. 31.21) in the ground. To avoid snapped ropes or other damage from frost heaves, *D* should be a function of local climate and local building codes and determined by the 100-year depth of the frost line in your area, a figure familiar to local contractors and code enforcement officials. Shallower depths may be legal for fence posts, but installation of a vertical antenna (or mast) on top of the fence post may change the mechanical situation considerably.

Dig the hole with a post-hole digger or an earth-auger-bit tool. The latter can often be rented in gasoline engine, electric, or manual versions from tool rental stores. Once the hole is dug, place about 4 in (or local requirement) of clean gravel (available in bags from hardware stores) at the bottom of the hole. Insert the post and plumb it (i.e., make it vertical in the hole) by adjusting a single set of guy ropes. Fill with your favorite brand of ready-mix concrete to a level just below the average surface around the hole. Cover with straw or burlap, and keep the top of the concrete moist for a few days. Do not subject the post to side loads for at least 4 days or whatever interval the concrete manufacturer recommends. Concrete needs time to cure; its strength increases exponentially, reaching 99 percent of its ultimate capability in 28 days, but a 7-day hiatus should be adequate before completing the antenna installation. Once the concrete has cured, top it off with soil and sod or garden chips.

Regardless of the height of its base, never attach the antenna base bracket to the mast or other support with just a single mounting point or U-bolt. Always use at least two-point mounting to prevent the antenna from pivoting and shearing off the mounting hardware.

Figure 31.22 shows how to mount a

2×4
mast to a fence post. The lumber should be kiln-dried and pressure-treated. Attach the
2×4
to the fence post with U-bolts. A
2×4
scrap is used as a wedge or shim to take up the difference between the post and the mast. In some installations, the U-bolt will go around the perimeter of the
2×4
, while in others a pair of holes can be drilled in the
2×4
to admit the U-bolt arms. The U-bolt must be at least 1⁄4-in #20 thread, and either a 5/16-in bolt or a 3/8-in bolt is highly recommended if the mast supports any significant weight. Some installers place a cinder block or large brick at the base of the
2×4
to bear the static and dynamic loads (i.e., gravity and wind) that eventually can cause the mast to list because of eccentric loading on the footer. An added reason is to keep the butt end of the
2×4
off the ground and thereby prevent or slow down rot. Alternatively, a pressure-treated
4×4
of the sort used to support decks (Fig. 31.23) can be used.

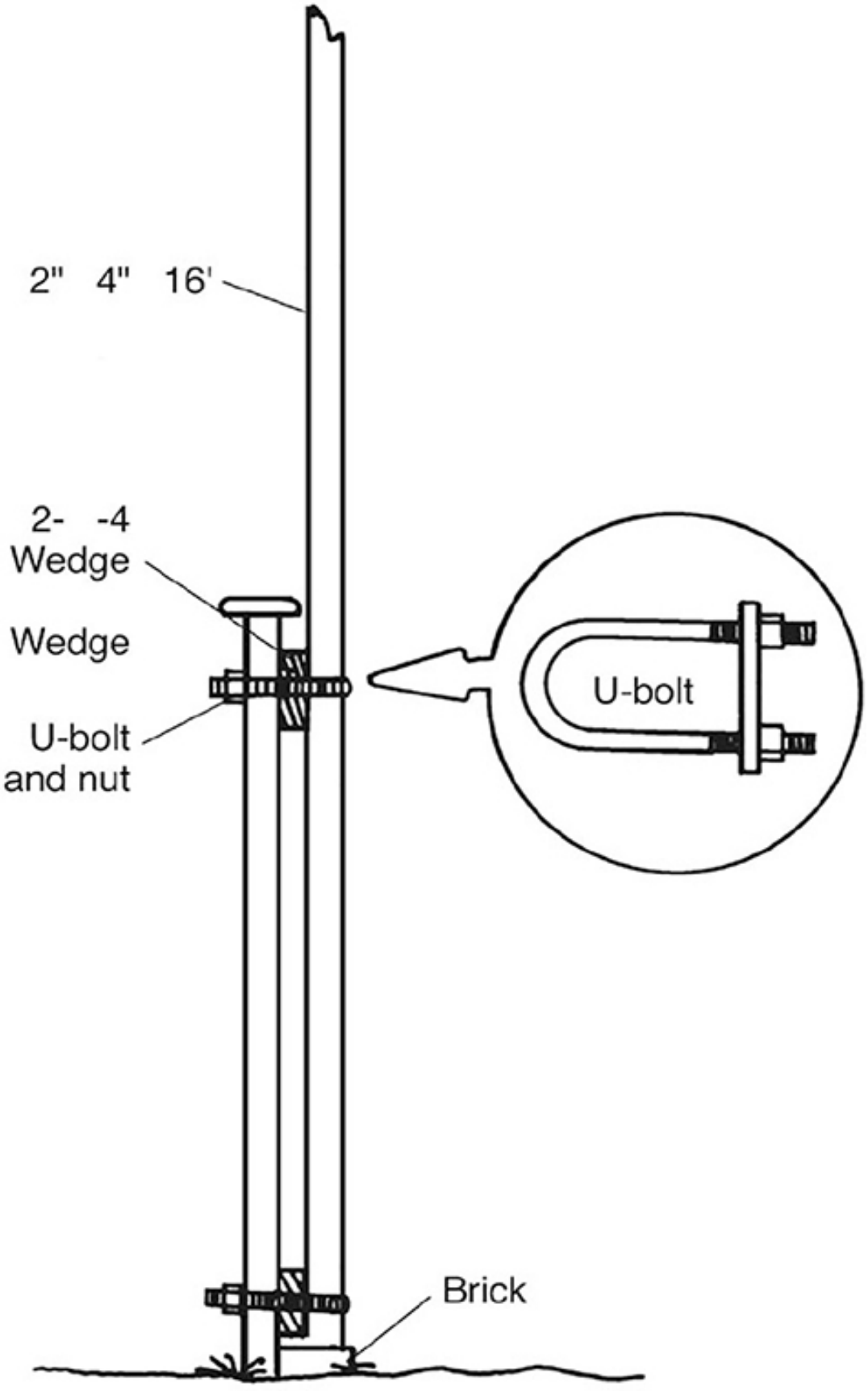

2" 4" 16'
2- -4
Wedge
Wedge
U-bolt
and nut
U-bolt
Brick

FIGURE 31.22 Mast mounted to a fence post.

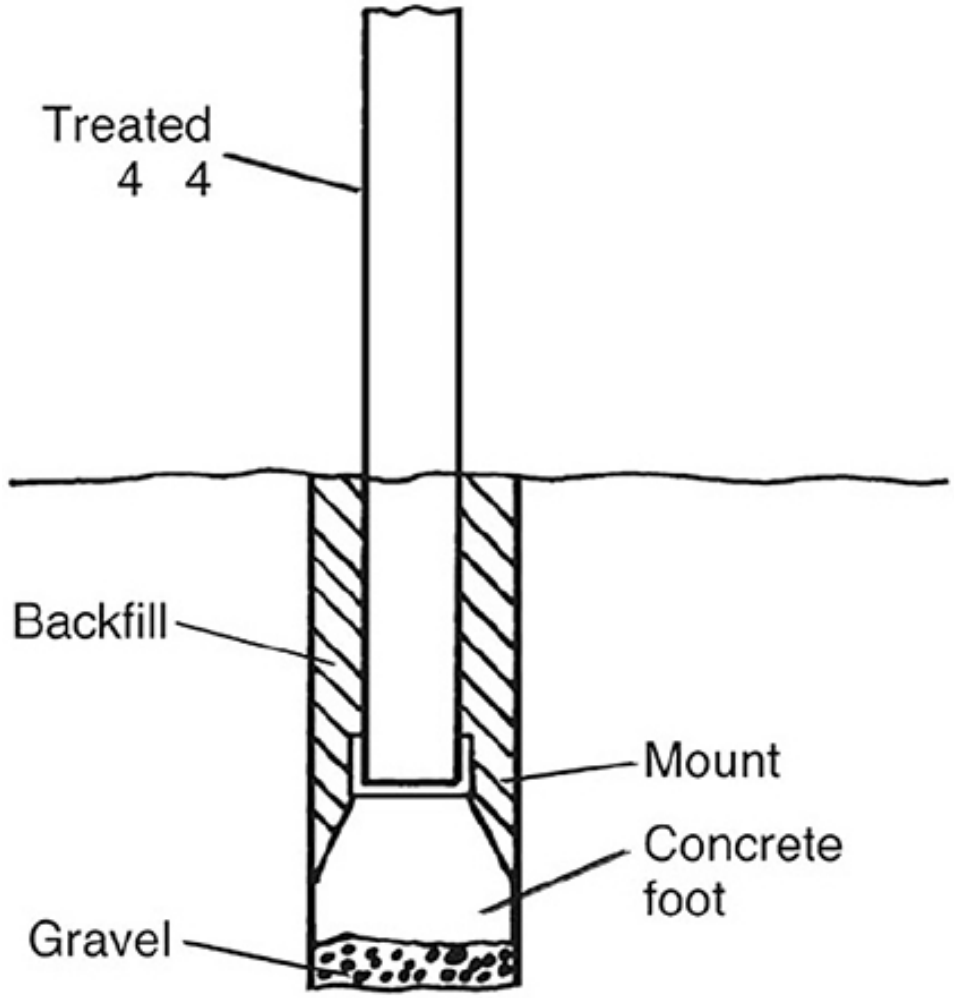


FIGURE 31.23
4×4
antenna support.

Even if ground-mounted, full-length
$\lambda/4$
verticals (35 ft, more or less) for 40 m require guying unless they are made of extremely strong, self-supporting members. A full-size 40-m vertical is akin to a very tall flagpole and requires similar construction unless it is guyed at one or two levels. Even with guying once it is fully vertical, the tubing must be strong enough to

not bend or crease while being pulled or pushed up from horizontal to vertical. The problem is even more severe for $\lambda/4$ 80-m ground-mounted verticals. Irrigation pipe, either butt-joined or telescoped with reducers, is often employed. Antennas of this length require larger ground crews or a cherry picker for successful installation, and the use of techniques covered in Chap. 32 is appropriate.

CHAPTER 32
Towers

When the height or strength required for a specified antenna installation exceeds the capabilities of trees or cordage, safely and reliably maintaining antennas at the desired height requires a tower. Typically, we call upon towers whenever rotatable HF antennas such as Yagis and quads are to be installed, although UHF, VHF, and even some small HF Yagis are sometimes mounted (if only for temporary events) on well-guyed heavy-duty masts. We use towers to support very high wires, especially for 160 m and 80 m, since few of us have access to trees in excess of 100 ft tall.

The purpose of this chapter is to familiarize the radio enthusiast with mechanical and safety considerations that are part of any tower project associated with putting an antenna up high in the air and keeping it there through episodes of severe weather. There are many antenna installation tasks that can be done by amateurs possessing the proper knowledge, the right tools, and a healthy respect for the safety of people and property. When planning an installation, be aware that the information given herein is merely informal guidance.

In the United States there is a wide variation from state to state and town to town in how local municipalities deal with towers. Regulations may range from strict codes to near total disinterest in what happens on private property. Similarly, there is little consistency in the way insurance companies handle the existence of a residential tower or losses resulting from the fall of a tower. Regardless of the stance taken by your local officials or your insurance carrier, you should have your planned tower installation reviewed by a professional mechanical or structural engineer licensed to practice in your state. Become familiar with and abide

by any local, state, and federal mechanical, electrical, and building codes that pertain to your project. Follow the manufacturer's instructions for storing, handling, installing, and using the materials and structures that are part of your project. It is not worth the risk to skirt any of these regulations or recommendations.

Because amateurs have multiple bands of frequencies ranging from MF to UHF and beyond, and multiple modes of propagation, including line of sight and ionospheric skip, tower requirements vary widely. Businesses and amateurs employ towers as short as 6 or 8 ft (tripod towers designed to straddle peaked rooftops) and as tall as 300 ft. Most nonbroadcast tower installations fall within the 30- to 100-ft range. Taller towers generally improve radio communication quality, but cost roughly scales with the square of the height and the required safe fall zone space increases with height as well. Proximity to airports and local zoning and building code restrictions also limit the feasible height of private antenna towers. The last two factors are discussed in more detail in Chap. 34.

32.1 Types of Towers

Several different forms of tower construction are used by amateurs and other services. The majority in the United States consist of lattice structures with a triangular cross section, although a few employing a square cross section (called windmill towers) are in use. A small but structurally impressive segment of the tower market consists of hollow tubes and solid surface poles with square or circular cross sections. At least one commercial model, *Big Bertha*—available in heights up to 142 ft—is a thick-walled round pipe that is driven by an electric motor at the base. In this model, the entire tower (called a rotating pole) is turned via chain drive from the motor.

Ignoring roof-mount tripods and house-bracketed towers, both of which are severely limited in maximum height, the following list of

common tower types is arranged in order of increasing cost for towers in the 70- to 100-ft range:

- Guyed triangular lattice sections
- Guyed triangular lattice section foldover
- Triangular lattice section crank-up (some with tilt-over option)
- Freestanding nonrotating pole (mall parking lot lamppost style)
- Freestanding tapered triangular lattice sections
- Guyed triangular lattice section rotating tower (or a hybrid)
- Freestanding rotating pole

32.2 Tower Fundamentals

If the wind never blew, a tower's role would be limited to supporting the weight of antennas, rotator, and mast. For most antenna and tower installations, deadweight is a secondary concern compared to the forces that result when strong winds act on the tower, antennas, rotator, mast, cables, and other accessories located above ground level.

The first requirement of a tower installation is a strong, solid base. Consider the free-standing tower of Fig. 32.1. The purpose of this tower is to support and allow rotation of a Yagi beam at a desired height in the presence of a wind. The wind is assumed to be applying a perfectly horizontal force to one side of the entire system. In response, the tower and antenna will move sideways, in the same direction as the wind, unless prevented from doing so.

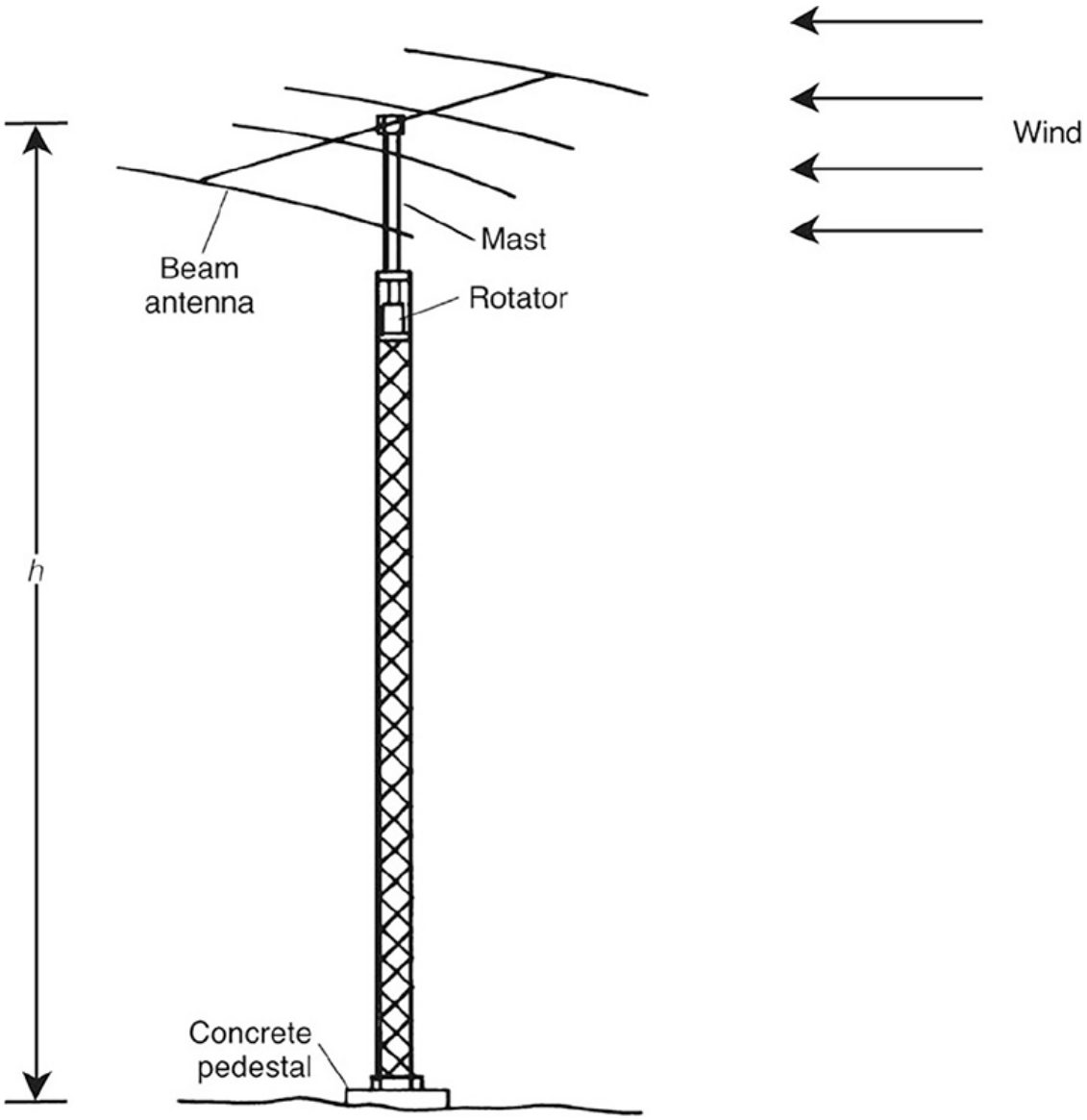


FIGURE 32.1 Antenna atop freestanding tower.

With the base securely anchored, the tendency of the tower is to overturn. The tower acts as a lever. Ignoring the force of the wind on the tower, the force of the wind blowing on the antenna at the top of the tower creates a moment arm or torque on the base of the tower with a magnitude equal to the height of the tower multiplied by the horizontal force of the wind pushing on the antenna:

$$M_{\text{ant}} = H_{\text{ant}}\ F_{\text{ant}} \quad (32.1)$$

The force is the wind pressure multiplied by the effective projected surface area of the antenna exposed to that pressure:

$$F_{\text{ant}} = P_{\text{wind}} A_{\text{ant}} \quad (32.2)$$

Different surface contours having the same area develop different resistance, or drag, in response to a given wind speed. The force on a flat plate squarely facing a breeze is greater than the force on a long piece of cylindrical tubing in the same breeze. Surface areas for tower-mounted antennas manufactured in the United States are usually given in equivalent or effective square feet. A common convention is to convert the actual area of one face of each cylinder into an equivalent flat plate area, often by multiplying the cylindrical surface area by 0.6, to account for the reduced drag when air flows past a cylindrical tube, as compared to a flat plate.

For most types of antenna, the equivalent area that directly faces the wind varies with the wind direction and the compass orientation of the antenna. In general, the number of interest is the worst case or maximum equivalent area. For a multielement Yagi, that may occur when the beam is facing the wind or when it is broadside to the wind or when turned at some angle between those two extremes. The pressure resulting from wind is given by

$$P_{\text{wind}} = K v_{\text{wind}}^2 \quad (32.3)$$

where

v_{wind}

is the wind speed in miles per hour and *K* is a constant equal to 0.00256 when using mph and the wind pressure is in PSF.

Of more interest to us is something called flat plate impact pressure (

P_{flat}

). This is the net pressure increase experienced by a flat plate squarely facing into a steady wind at a given speed.

P_{flat}

is proportional to the square of wind speed, but the constant is different from the *K* in (32.3) and depends on a number of geometrical characteristics, not just of the plate but of the environment around the plate. The equation for

$$P_{\text{flat}}$$

may vary with the underlying assumptions embedded in a particular analysis. For some years, curves relating the flat plate impact pressure of flat microwave antenna panels to true wind speed often used a

$$K_{\text{flat}} = 0.004$$

(see Fig. 32.2). Today the determination of the K-factor to be used may be conducted on an installation-by-installation basis. The most important fact to take away from Eq. (32.3) or Fig. 29.2 is that pressure on an antenna and the horizontal force on a tower is proportional to the square of the wind speed.

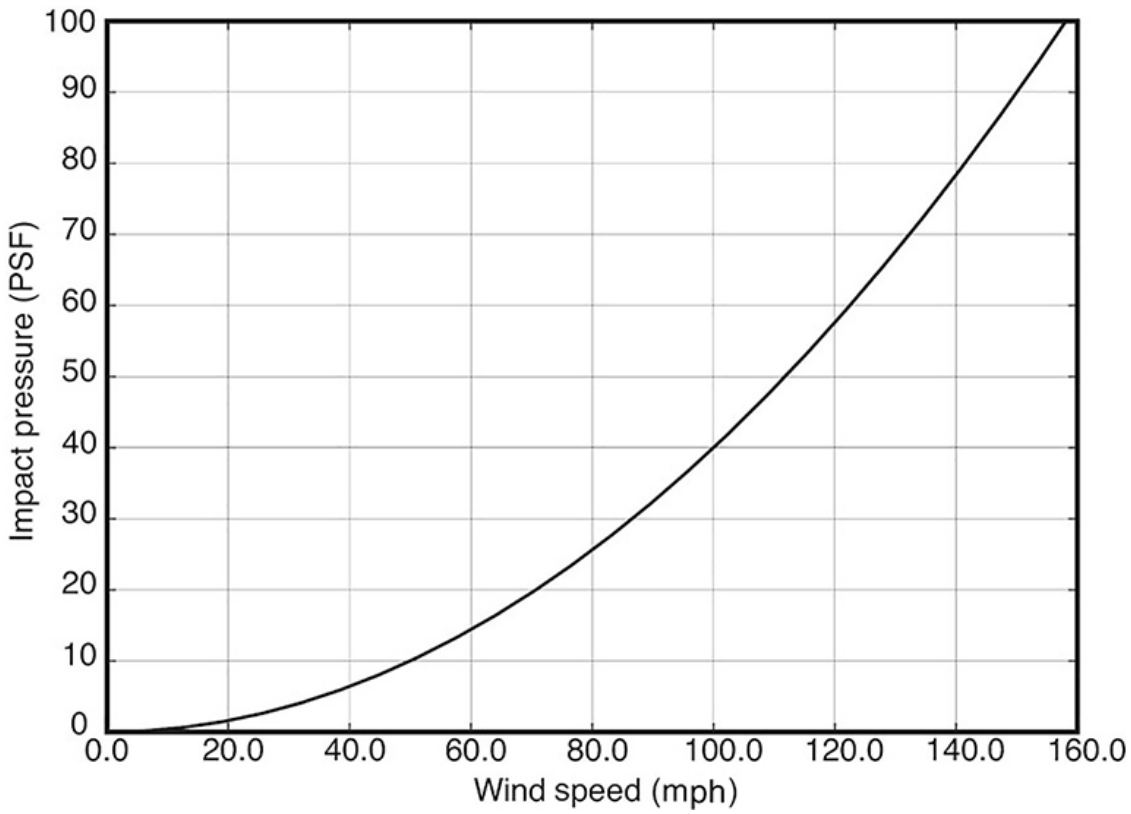


FIGURE 32.2 Pressure versus wind speed.

In 2005, the Telecommunications Industry Association (TIA) issued Revision G of TIA-222, their Structural Standards for Antenna Supporting Structures and Antennas. Determination of the appropriate wind speed ratings to apply to a tower or antenna in Revision G represents a substantial and significant departure from the methods of earlier revisions. The focus has shifted to wind speed defined by a peak 3-s gust. Although there is no precise 1:1 equivalence between these numbers, tower and antenna installations designed for approximately 70 mph using the earlier Revision F standard seem to be comparable to designs based on a 90 mph peak 3-s gust standard. Further, Revision G introduces and requires certain adjustments (usually additions) to peak wind speed numbers because of terrain and/or reliability requirements for the support structure.

The standard provides a baseline wind speed specification for each county in the United States. Most local building codes officials can provide you with the current requirements for their own counties. The baseline wind speed rating for flat, open countryside across the vast majority of the interior 48 states is 90 mph, but this number must be adjusted upward with icing, with increased tower height, and whenever unusual topographic features (hills, mountains, cliffs, ocean coastlines, etc.) are encountered. Overall, the baseline speeds vary from a low of 85 mph along the Pacific coast states to a high of 150 mph along the coastlines of south Florida and some of the other states bordering the Gulf of Mexico. TIA-222 wind speed ratings are a part of local building codes in many municipalities.

Example 32.1 A typical HF short-boom trap-triband HF Yagi might have a published effective wind surface area of between 5 and $6\ \mathrm{ft}^2$. The Cushcraft *A-4S*, for instance, is specified at $5.5\ \mathrm{ft}^2$. This figure is determined by the manufacturer from calculations on both the cylindrical parts (boom, elements, struts, traps, etc.) and the flat plate parts (brackets, end faces of cylindrical items) at different

orientations with respect to the wind. Next, we use Fig. 32.2 to obtain the impact pressure for a target wind speed. Let's suppose that we conclude that 100 mph is a conservative value for the highest wind speed we expect our tower installation to survive. From the graph or the equation we obtain
$P = 40$
PSF.

Solution Suppose we add
$0.5\ \text{ft}^2$
for a rotator housing and other components at the top of the tower. Then our total wind load (ignoring the tower itself) is
$40(\text{psf}) \times 6.0(\text{ft}^2)$
, or 240 lb of horizontal force. If the antenna is mounted at the top of an 80-ft tower, for instance, the total moment arm, measured from the base of the tower, due to the force of the wind on just the antenna and rotator is calculated from (32.1) as 19,200 lb-ft. Because the bottom of this tower is rigidly held by its base, the tower will resist overturning. To do that, the moment arm is converted back into a compressive force in the downwind or leeward side of the tube if it's a single rigid member or in the downwind leg if it's a triangular lattice structure. That force can be calculated by working backward through a second moment arm equation involving the cross-sectional dimensions of the tower. Specifically, we take the distance of any leg or outer wall from the centerline of the tower and divide that distance into the 19,200 lb-ft. For a tower whose legs form a triangle 18 in on a side, the calculation is:

$$\begin{aligned} F_{\text{comp}} &= \frac{M_{\text{ant}}}{R_{\text{tower}}} \\ &= \frac{19,200}{0.866} \qquad (32.4) \\ &= 22,170\ \text{lb} \end{aligned}$$

Mounting a small HF Yagi antenna atop a freestanding 80-ft "stick" has converted a sideward force of 240 lb into a vertical compressive force nearly 100 times greater. For this reason, while antennas are made of aluminum, towers are almost always made of steel.

For comparison, the maximum allowable axial compression in a side rail (or tower leg) for Rohn 45 tower sections, popular with amateurs and business band users, is slightly less than 8,000 lb. Thus, it is not

allowable to put an antenna of this size on top of 80 ft of freestanding Rohn 45 tower sections. The tower and its accessories all have surface areas, so the wind loads of the tower and accessories on it (rotator, brackets, feedlines, control cables, work platforms, etc.) have to be added to the antenna wind loading to obtain the total real-world moment arm and force on the tower legs. When that is done, it can be seen that the maximum allowable height for a freestanding Rohn 45 with no antennas or accessories except perhaps a small UHF whip might be more like 50 ft, depending on the maximum wind speed anticipated.

Although the preceding calculation is approximate, it provides a basic understanding of the key structural issues germane to communications towers. Larger antennas and/or taller towers result in larger compressive forces at the base of the tower. Starting from an understanding of the lever action that is involved, we can now more easily visualize ways to improve the antenna-carrying capacity or increase the height of our antenna support:

- Use thicker tower legs of the same material, or make the tower members out of stronger materials. (Both approaches are important contributors to the strength of freestanding rotating poles.)
- Increase the distance from each leg to the centerline of the tower, especially toward the base of the tower. This is the principle behind tapered self-supporting towers such as the Vesto, the Rohn *SSV* series, some rotating poles, and most crank-up towers such as the EZ-Way products.
- Reduce the moment resulting from wind pressure on the antenna and tower members with the judicious addition of guy wires that counteract the force of the wind. This is the principle behind all guyed towers.

Caution *The calculations given in this section are simplified in order to emphasize basic structural strength issues in towers. No attempt has been made to include the effects of wind gusts,*

sudden shifts in wind direction, uneven terrain, updrafts and downdrafts, icing, mechanical resonances, and a host of other effects that a thorough design must take into account. Any new tower installation project should be conducted in accordance with the determinations of a professional engineer (PE) licensed to practice in your area. In many municipalities, certain paragraphs of the local building code may apply and periodic inspections may be required during the course of the project. Many building code enforcement offices will accept preexisting standard installation drawings and calculations provided by commercial tower manufacturers if they have been signed and stamped by locally licensed PEs, so that avenue should be explored before incurring the expense of designing a tower installation from scratch.

32.3 Guyed Towers

One way to minimize the moment arm at the base of a tower from wind blowing against an antenna at the top of the tower is to directly resist the force of that wind through some physical mechanism. The most efficient way is to oppose the wind by pushing on the antenna in a direction opposite the wind or by attaching a guy line to a firmly fixed support on the windward side of the antenna. If the guy wire is adjusted to be taut with no wind present, then any increase in force on the antenna when the wind comes up will be met with an increase in tension in the guy wire, with the result that the moment arm at the base of the tower is substantially reduced or eliminated.

In the ideal installation, we would attach the far end of the guy wire to a support at the same height as the antenna so that the guy wire was perfectly horizontal. But if such a support existed, we might as well attach the antenna itself to it and eliminate the tower entirely.

Since our objective in putting an antenna on top of a tower is to get the antenna up above surrounding objects, we must resign ourselves to connecting the far ends of guy wires to attachment

points substantially lower than the antenna itself. Often the best we can do is to attach the guy wires to restraints or anchors at ground level.

Figure 32.3 is a schematic of a very simple guyed tower. As before, a

6-ft^2

antenna sits atop an 80-ft tower, which is guyed at two levels (40 and 80 ft), but this analysis is going to discuss the effect of only the top set of guys on stresses in the tower members. Again, let us assume that the tower members are thin and generate only small wind loading of their own. Assume also that the guy wires or guy lines are made of material that does not stretch under tensions encountered here, and that they are attached to earth anchors 60 ft out from the base of the tower. From the trigonometry of 3-4-5 right triangles, we know that each top guy wire is 100 ft long.

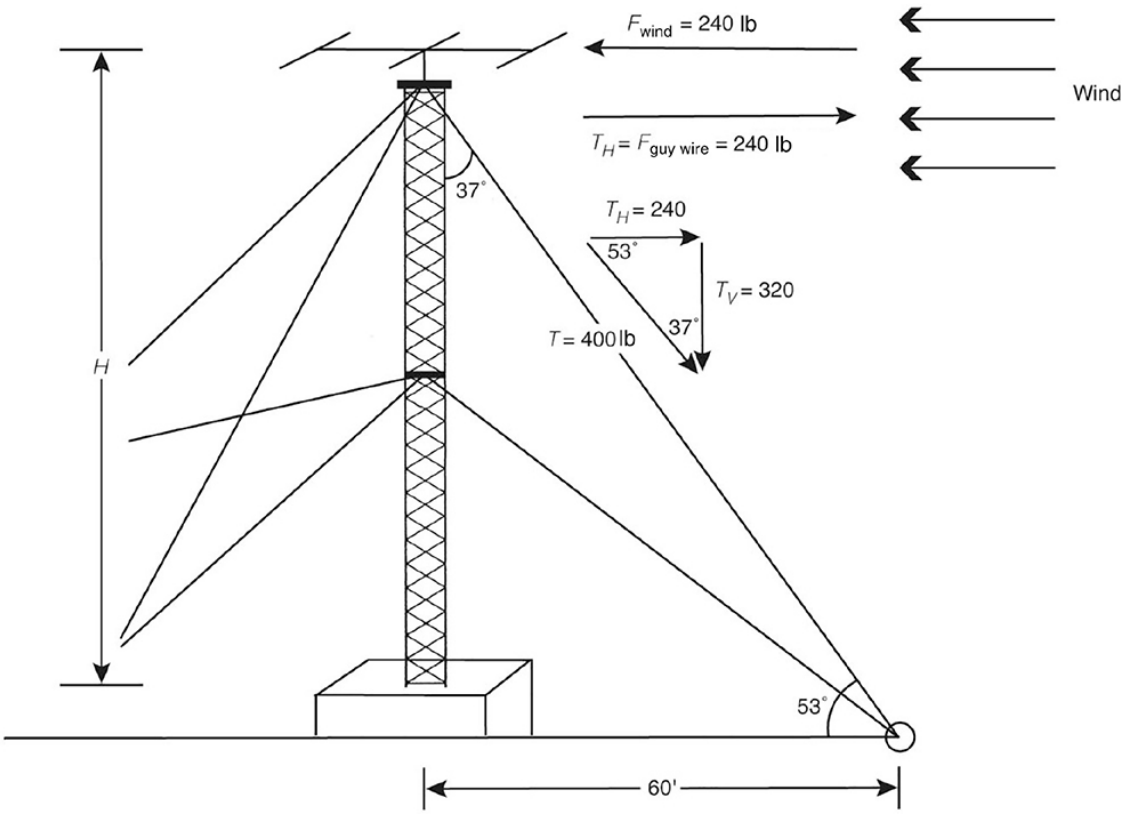


FIGURE 32.3 Forces on a guyed tower.

When the wind blows from right to left in Fig. 32.3, it pushes on the antenna and tries to move it to the left. The antenna doesn't move because it is restrained from doing so by the top guy wire that is to windward (i.e., on the right-hand side) of it. For an object to remain motionless in one place, the vector sum of all the forces acting on it must be zero. Thus, if the antenna stays in place, it must mean the horizontal force of the wind is exactly balanced by the horizontal force exerted on it by the guy wire. But the guy wire is not horizontal, and since the tension in the wire has to point along the axis of the wire, we have to break the tension up into two components (trigonometry again): a vertical component (T_V) and a horizontal component (T_H). It is the horizontal component of the tension that must balance the horizontal push of the wind. Thus, if the force on the antenna from the wind is 240 lb, T_H must also be 240 lb, acting on the antenna in the opposite direction.

Again using the trigonometry of right triangles, if $T_H = 240\ \text{lb}$, $T = 400\ \text{lb}$ and $T_V = 320\ \text{lb}$. But why are these of interest? *T* (400 lb) represents the total force (tension) in the guy wire itself. There are at least two reasons to care how large it gets:

- The maximum *T*—multiplied by a suitable safety factor—determines the strength required for all the guy wire components being procured for this installation.
- The earth anchor must be designed to withstand a maximum force of *T*—multiplied by a suitable safety factor—imposed on it by the guy wire.

T_V
, on the other hand, is a downward force that is applied to the tower legs. In other words, a diagonal guy wire creates a downward force on the tower that is proportional to the magnitude of the horizontal wind force. That is the penalty we pay for not being able to employ perfectly horizontal guy wires. But the benefits are enormous—now, instead of dealing with compressive forces in the leeward tower leg of over 22,000 lb, we are imposing a paltry 320-lb downward force on that leg. Our tower is no longer overloaded and we can consider adding more antennas or going higher.

Caution *Once again, this is an approximate analysis of a complex mechanical design problem. We have not considered many factors that add to the compressive force on the tower members, including the deadweight of the rotator and antenna, as well as both the weight and wind load of the tower itself, the guy wires, and any cabling or ice buildup, nor have we considered transient loadings from shifts in wind speed or direction, the impact of the initial tensioning of the wires, the interplay between any stretch, however slight, in the guy wire material, and the reappearance of a fraction of the original freestanding moment arm. Professional engineering is needed to perform a complete tower analysis and design.*

Several guyed tower configurations are commonly seen. Figure 32.3 shows the basic form—multiple tower sections (each usually 10 or 20 ft in length) having the same cross-sectional dimensions, braced by multiple levels of guy wires. This type of tower costs the least, and many installers and owners feel it is also the most comfortable platform to climb and work atop. In addition to the standard sections used as building blocks, each tower typically requires a base section and perhaps side-mounting hardware or a special top section designed to accept a thrust bearing for the antenna mast and a rotator mounting platform. The maximum possible height varies with the strength of the sections; Rohn 25, for instance, can be erected to 200 ft, while Rohn 45 is spec'd up to

300 ft. But in areas with higher wind zones, or with antenna loads greater than those assumed by Rohn in its standard configurations, those maximum heights may end up being reduced by 50 percent or more.

Towers are guyed at more than one height to keep side stresses on the tower legs from growing beyond a certain value as distance above or below a guy station increases. Rohn 25, for instance should be guyed every 25 to 30 ft, while Rohn 45 can be guyed at 30- to 35-ft intervals and Rohn 55 at 35- to 45-ft intervals for some manufacturer-recommended configurations, dependent on total antenna wind load, maximum wind speed, and other installation factors.

Note *These guy spacings are taken from the Rohn data sheets for a sampling of specific configurations. They may not be appropriate for your installation, antenna loads, and wind speeds. When in doubt, follow the recommendations of the manufacturer to the letter. If you make any changes, make them on the conservative side and get them approved by a professional engineer.*

Caution *The guy wire spacings previously discussed are for towers that have been completely installed. During the process of erecting a tower, additional temporary guy wires, at closer spacing intervals, may be needed.*

Caution *Rope of any kind has too much stretch to be suitable as a guy wire for a steel tower. The only suitable materials for permanent guys are those specified by the manufacturer; typically that will mean only stranded EHS guy wire or Phillystran nonconductive cable.*

32.4 Bracketed Towers

Although triangular lattice towers with a constant cross-sectional area (such as Rohn 25 or 45) are designed for guyed installations, short versions of these towers (say, 20 to 40 ft) can be erected with

no guys at all (freestanding mode). If bracketed properly to a strong structure such as the gable end of a home, generally at a point 18 ft or more above the base, the allowable unguyed tower height can be increased by nearly the height of the bracket.

If a freestanding 30-ft tower comprised of Rohn 45 sections is capable of handling a certain wind load, bracketing the tower at 22 ft may allow the tower to be extended to about the 50-ft level for the same antenna load. But if that specific combination of heights and antenna area is not explicitly given in the manufacturer's data sheets for this tower model, your installation plans should be reviewed by a licensed professional engineer.

Bracketed towers share a disadvantage with rooftop (tripod) towers: audible noise inside the structure. Wind-induced vibrations in the antenna and tower, as well as vibration from any rotator and thrust bearing motion, are transmitted mechanically through the attachment hardware to the framing of the structure. A bracketed 40-ft tower bears a resemblance to a violin or guitar. There is additional noise from cables loosely tied to the side of the tower and the brake mechanism in some antenna rotators. Vibrations can be dampened with rubber shims between the tower and the brackets (a discarded inner tube is an excellent source of rubber stock for this application).

Often overlooked is the need for strength in the structure the bracket is attached to. Most residential exterior walls are designed primarily to support lots of weight directly above and they have relatively little resistance to horizontal forces. A wall should be reinforced specifically to withstand the sideward force of high winds on a bracketed tower and antenna installation. This is most easily done at the time the house is constructed. Under no circumstances should you expect attachment to an existing eave or soffit to be adequate. This is another area where expert help is particularly important.

32.5 Foldover Towers

A sketch of a foldover tower is shown in Fig. 32.4. This type of tower uses the same triangular lattice construction as the fixed, guyed model but is hinged at a point between one-third and two-thirds the total height. The tower must be guyed at the hinge level, and, to avoid interference with the counterbalance boom that goes up in the air as the top half of the tower is cranked down, four guy wires, uniformly spaced at 90-degree intervals, are required. A waist-high winch spools aircraft cable in or out to raise or lower the upper part of the tower between its fully upright and foldover positions. A safety bolt at the lower end (or farthest point) on the counterbalance boom secures the tilt assembly once erected, thus minimizing end-of-travel "slop" and eliminating the need to rely on the cable to keep the upper portion perfectly vertical when fully erected.

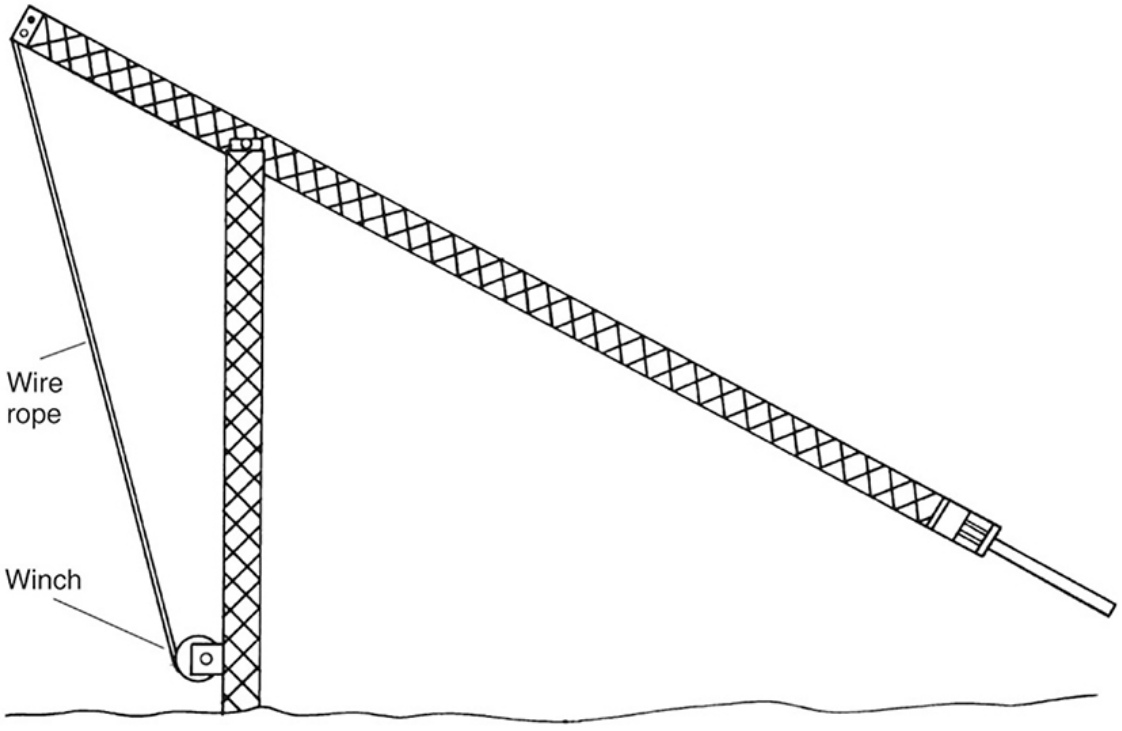


FIGURE 32.4 Foldover tower.

At one time Rohn sold foldover towers in both 25G and 45G ratings, but they are no longer in their catalog. Occasionally one can be found on the used market. Alternatively, if the special

hinged section and counterbalance boom can be procured, the remainder of the tower is conventional Rohn product that is still available new.

The advantage of a foldover is that the top of the tower can be brought close to the ground and all work on antennas, rotator, and cables can be done while standing on the ground or a stepladder. There are some disadvantages to this system that are not present in a conventional guyed tower:

- During the process of raising and lowering the tilting portion of the tower, the winch cable and upper tower sections experience maximum stress when the tilt assembly is horizontal. This puts a severe weight and wind load limit on what can be installed on top because side load limits for triangular lattice tower sections are much lower than their longitudinal compressive and tensile limits. In addition, the winch cable itself has a safe working limit that must not be exceeded.
- The manual winch represents a potential safety hazard. With a full load at the top of the tower, a person of normal build may have difficulty turning the crank smoothly.
- If the winch assembly gets rusty or gummed up such that the safety catch does not properly engage the teeth on the crank assembly, the winch handle can freewheel if the operator's hand slips while cranking the antennas up or down. Even if the safety catch does engage, any lag in the catching mechanism can result in a lot of weight on a very large moment arm coming to an abrupt stop, with the possibility of damaging parts of the tower or, worse yet, snapping the cable and seriously injuring the operator or helpers.
- As marketed, the Rohn foldover had no guy wires at the top. There is merit to adding a set, depending on how many sections have been installed above the hinge and

the total wind load of the antennas. The owner needs to factor their added weight into the maximum weight calculations for the tilt process.

- Some antenna rotators especially those with internal ball bearings, may not take kindly to being tilted partially upside down, depending on internal play and retainer spacings.

32.6 Crank-up Towers

Figure 32.5 shows a crank-up tower. In the example shown, a wide base section supports a smaller moveable upper section. Commercially available models usually have two to five sections. In all of these, a manual or a motorized winch is used to lift the upper sections into fully extended position. Typical two-section towers can be raised from a low position of 20 ft to a height of 35 ft or so. When the tower is fully extended, some models incorporate provisions for locking the sections in place with steel bolts, steel bars, or other mechanisms, but these options are by no means universal. A lot of shear force is applied to these bolts, so it is wise to use several hard stainless steel bolts for the lockdown.

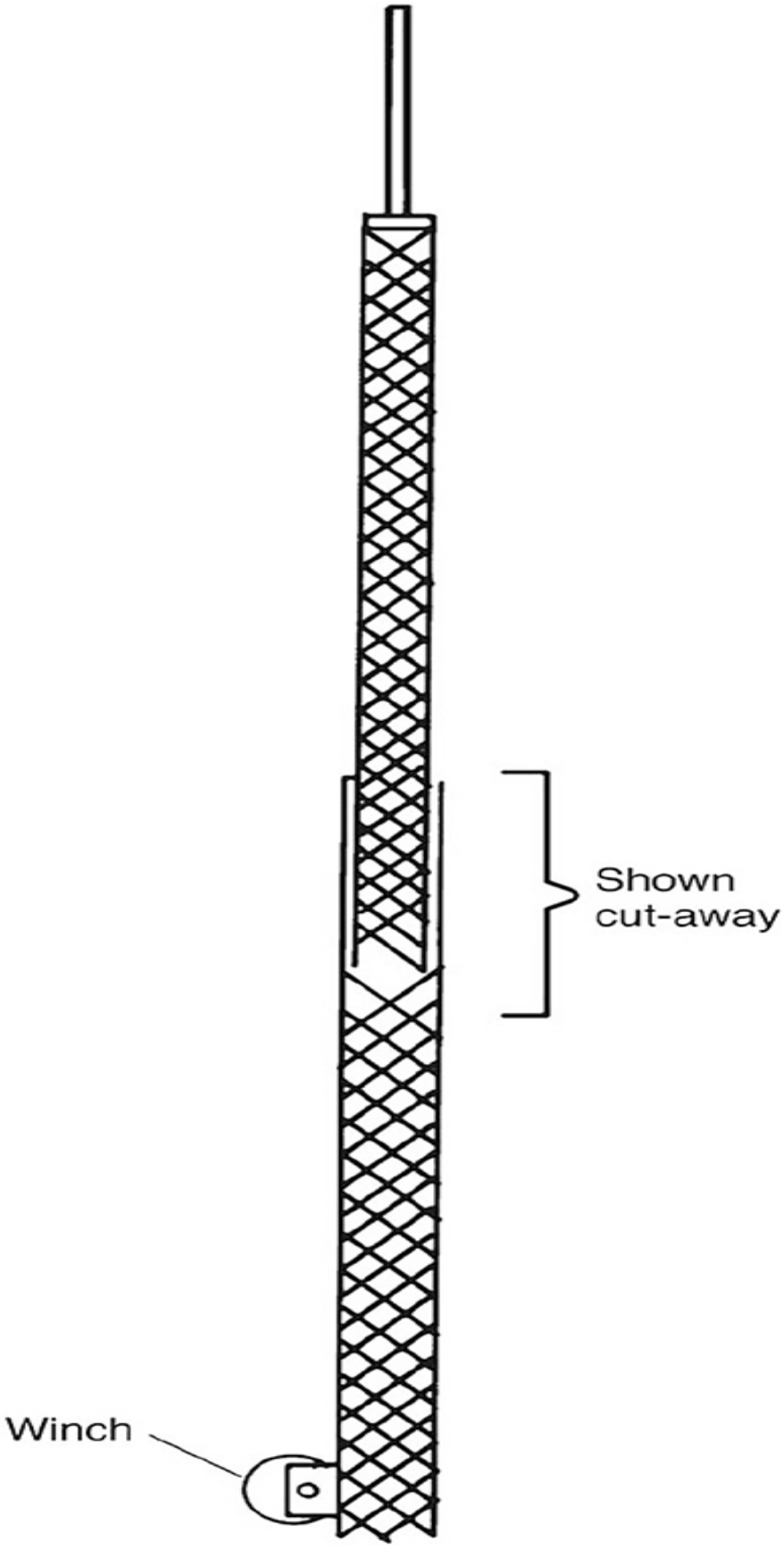

Shown
cut-away
Winch

FIGURE 32.5 Crank-up tower.

Caution *There is a serious danger inherent in the design of the crank-up tower. If a cable supporting extended sections breaks while you are working on or near the tower, one or more sections will come plummeting down the shaft formed by the sections below it and shear off any arms, legs, or other body parts that get in the way. Whenever you work on this form of tower in its extended position, either fully or partially, use steel fence posts or similar pieces of metal as a safety measure (Fig. 32.6); at least two should be used, and both should be attached securely at both ends with rope so they can't be knocked out accidentally. These pipes are used in addition to, not instead of, the normal bolt fasteners that keep the antenna tower erect.*

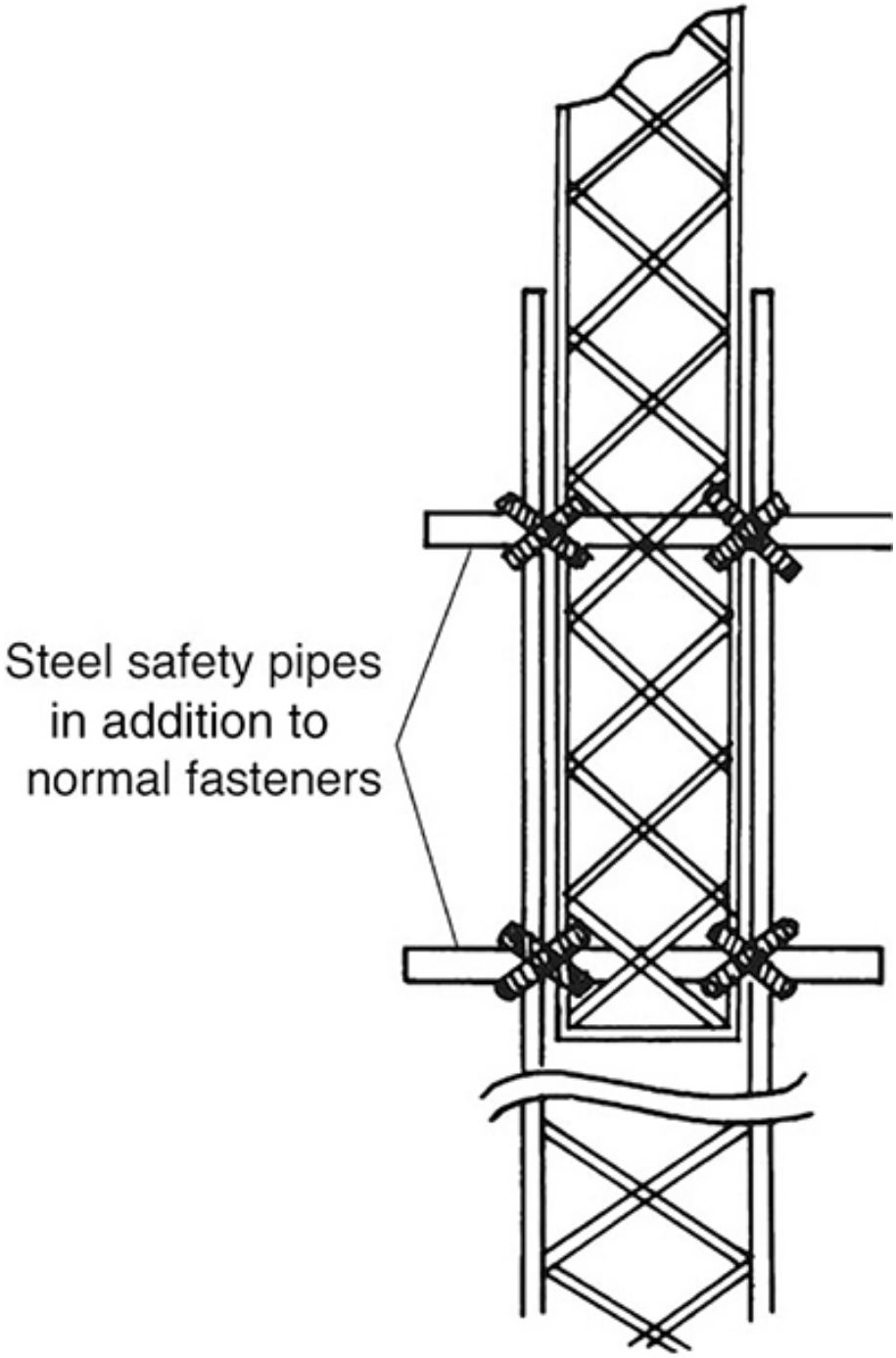


FIGURE 32.6 Safety precaution.

Some crank-up towers are motorized so that the tower can be raised or lowered without anyone having to stand at the base. Some crank-ups also tilt over, so only a stepladder is needed for installing, removing, or adjusting a typical HF Yagi. Between these

two optional features, quite a few of the potential hazards associated with crank-up towers are eliminated or substantially reduced. Of course, crank-up towers with either or both of these features cost quite a bit more.

Towers require periodic maintenance and examination, but crank-up towers as a category require more than most others. In particular, the cables that perform the raising and lowering functions must be frequently inspected and lubricated with compounds specified by the manufacturer of the tower.

For crank-up towers, the specified maximum wind load when the tower is fully extended is so small as to effectively preclude the installation of even a modestly sized HF beam. Prospective owners might wish to consider whether they want to be off the air every time the wind gets above 50 mph or whether they want to go outside and crank it up every time they operate.

32.7 Self-Supporting Towers

As mentioned earlier, self-supporting towers may be designed and constructed with a single fixed diameter or cross section from top to bottom or with a tapered cross section. The greater the taper, the more constant and smaller the compressive forces in the downwind leg or side of the tower will be from top to bottom.

32.8 Tapered Cross-Section Towers

Old-fashioned windmill towers seen on many farms are an excellent example of the basic design that governs the construction of self-supporting towers. The slope of the sides of the tower determines the ratio of compressive force in the leeward tower leg at the bottom of the tower to the force exerted on the antenna by the wind. When there is little taper (i.e., very little difference in tower cross-sectional area between top and bottom), the multiplier is quite high. When the slope of the sides or legs is pronounced, the multiplier is much lower. The value of sloping the legs of a tower can be understood by considering the value of spreading your feet to brace yourself against a strong wind.

Historically of primary use in commercial applications and occasionally in residential wind generator systems, the Rohn *SSV* series of tower sections has found favor with increasing numbers of radio amateurs. The current family of 16 different sections—each 20 ft long—in principle allows for construction of towers up to 320 ft in height. A heavy-duty version of the SSV family replaces the bottom eight sections with beefed-up alternates. A shorter tower can be selected from many different sets of consecutive sections, so if your objective is to build a 100-ft self-supporting tower, the SSV gives you nearly two dozen options to choose from, depending on the maximum wind loading. You can customize the strength and cost of your tower based on the maximum wind load you anticipate.

Self-supporting towers are expensive. Even though there are no guy wire costs, the superior strength of the sections makes them more expensive than guyed tower sections per foot of height, and the size of the concrete base (which takes the place of the guy wires in resisting any tendency for the tower to tip over in a strong wind) is massive. Installation of these 20-ft sections of tower will undoubtedly require professional rigging companies, and the interface between the top of the tower and typical amateur

rotators, masts, and thrust bearings will usually entail custom metalworking and welding at a local shop.

One problem with tapered self-supporting towers is the relative difficulty of climbing them. The climbing spikes commonly seen attached to one leg of the tower can be too small to be comfortably grasped. Often a matching metal ladder is available.

AN Wireless Tower Company is another manufacturer of tapered self-supporting towers for the amateur radio community. Like Rohn, their product consists of a family of sections with some heavy-duty options possible. AN Wireless's sections are 10 ft long, and the maximum height currently marketed is 120 ft.

A lighter-duty tapered tower family of 8-ft sections, originally offered by Spaulding and later by Rohn, can occasionally be found on used markets. Although inexpensive, these have some disadvantages. The side rails are made of U-shaped steel channel instead of cylindrical tubes. Consequently, they are painful to grip while climbing or when working at the top of the tower. The cross-bracing is comprised of diagonal steel strips on edge. Since there are no horizontal surfaces to stand on, climbing and working on one of these towers is rough on the feet. Invariably the used tower has rust on it. If so, it must be sandblasted and regalvanized or painted after being inspected for soundness. Hardware securing the cross-braces to the side rails must be removed, discarded, and replaced with new high-strength nuts and bolts. These tower sections also came in a heavy-duty version (the *HDBX* series). This is the only version that is truly adequate for even small amateur HF Yagis; the standard-duty model is suitable only for VHF and TV antennas.

32.9 Rotating Poles

Another form of self-supporting tower is the rotating pole. As compared to almost any other kind of tower, the rotating pole provides the greatest degree of flexibility in placing multiple high-gain antennas (notably long-boom Yagis) at specific heights on the

tower for optimum stack spacing and full 360-degree (or more) rotation. A typical rotating pole is supported by a bearing below ground level and is rotated by chain drive at ground level with a motor mounted on the same concrete pad that holds the tube in which the bearing is located.

A well-known rotating pole is the *Big Bertha*. The current *Big Bertha* design is available in heights up to 142 ft. A typical antenna complement is a three-element 80-m Yagi at the top, a pair of four-element 40-m Yagis at the midpoint and near the top, and three long-boom six-element 20-m Yagis spaced appropriately for maximum stack performance. The Tri-Ex *Sky Needle* was another rotating pole, available in 70 and 90 ft heights. These can occasionally found used but dismantling and transporting one can cost as much as three or four guyed towers.

A guyed rotating pole is less expensive than a *Big Bertha* (but still more costly than a conventional guyed tower). In some cases, enterprising amateurs erect their own towers, using standard Rohn 45, 55, or 65 sections and guy ring assemblies purchased separately. Others buy the complete tower and often its installation, as well from a vendor. Compared with a self-supporting rotating pole, the advantages of a guyed rotating pole are:

- It is far less expensive.
- It can be erected to greater heights.

Disadvantages include:

- Guy wire attachment points may interfere with optimum placement of antennas.
- The guy rings can be a challenging maintenance item.

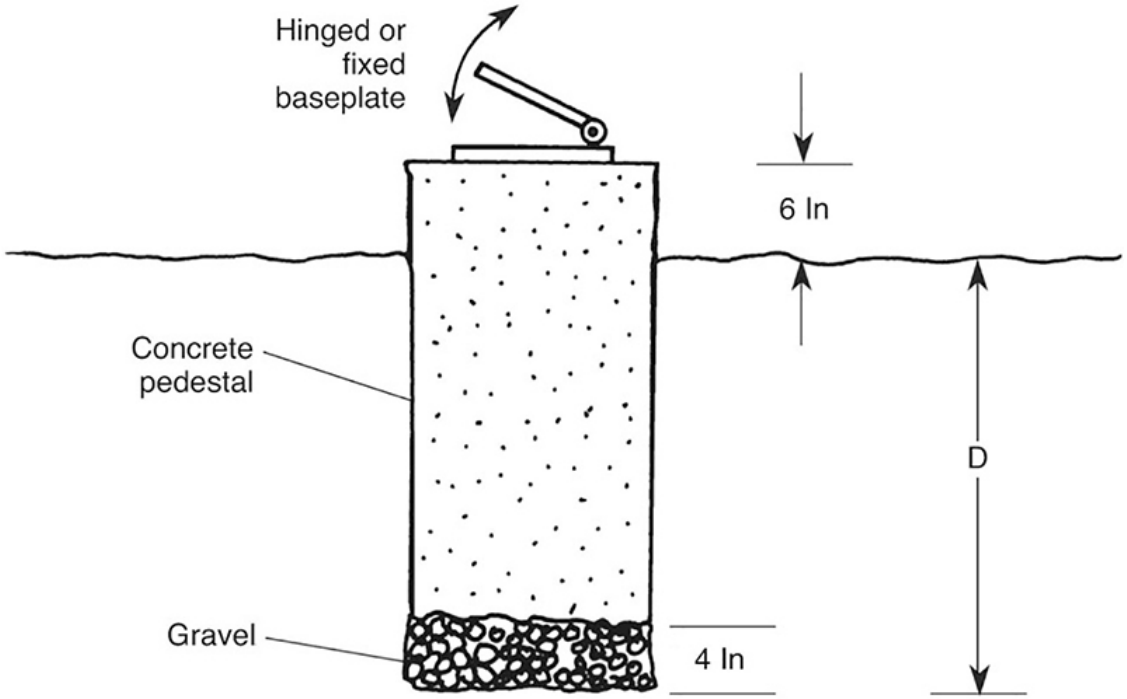
A hybrid guyed rotating pole consists of a fixed lower tower portion with a rotating portion above. This is sometimes done to reduce the overall cost of the installation when the height of the lowest rotatable antenna on the tower is a substantial percentage of total tower height. The disadvantage is that the rotator and associated drive system must be located at the junction of the fixed

and rotating segments, many feet above ground, making rotator maintenance more difficult.

32.10 Tower Bases

Regardless of its exact configuration, a communication tower is typically installed atop a buried pedestal made of concrete (Fig. 32.7). The base serves multiple purposes:

- It keeps the tower from sinking into the ground.
- It keeps the bottom of the tower from moving.
- In conjunction with a layer of stone beneath, it allows proper drainage of tower legs.
- For a self-supporting tower, it resists any overturning moment.
- It provides a physical barrier between the tower legs and corrosive soils.
- It may provide part of the lightning ground system for the tower.



F IGURE 32.7 Ground pedestal.

Base construction must follow any relevant local codes, but there are some general considerations to keep in mind. The surface area of the top of the pedestal should be as specified by the manufacturer, but the smallest practical or useful base is probably at least 3 ft by 3 ft. If a backhoe makes the hole, the base is likely to be square, but augers can also be used, especially if a circular hole is desired or necessary.

Caution *If the hole for a tower base or guy anchor is going to be more than 3 ft deep, do not dig it by hand, and do not leave it unattended. The sides of many excavated holes have collapsed unexpectedly.*

Regardless of manufacturer specifications, local codes, or other at-a-distance requirements, it is unlikely you will need less than 1.5 cubic yards of concrete for your tower base. Holes in the ground tend to be bigger than planned, forms notwithstanding. Don't forget to add to that total for any concrete guy anchor stations you

may have. Presuming your site will allow reasonable access to the tower hole by a concrete truck or a backhoe, order no less than 2 yards of premixed concrete from your local supplier. Further, such a small order will undoubtedly be the last load delivered from that particular truck, and you're apt to get shorted a little bit—it goes with the territory. Most important, mixing bagged concrete by hand is hard and takes a long time. The strength and permanence of your base can be compromised when its pour is stretched out over the long interval it takes to process that many bags by hand. Concrete can generally be ordered and poured at any time of year. Additives can be used to improve various properties of the concrete. When you order concrete, ask for an extra-strength mix, which might be designated as 3000 PSI.

Regardless of the manufacturer's minimum specifications, the depth of the pedestal and the depth of the gravel or stone base should reflect the local frost line. The concrete pedestal should extend below the local frost line even if local law permits less. In northern parts of the United States, this will put it at least 4 or 4 1/2 ft below the surface. Beneath the base it is important to have a layer of stones for proper drainage; layer thickness will be a function of the type of soil encountered.

Most manufacturers specify that the concrete pedestal include rebar reinforcing steel. Refer to manufacturer information before pouring the concrete. Make sure you know exactly what needs to be embedded in the concrete as part of the tower base before summoning the concrete truck. There are many forms of tower bases and at least an equal number of custom inserts for the concrete. Examples include:

- Short section of tower (typically used for constant-diameter guyed triangular lattice towers)
- Pier pin (also used for constant-diameter guyed triangular lattice towers)
- Hinged baseplate (often used for guyed triangular lattice towers up to 40 ft)

- Tilted flanges (for each leg of a tapered self-supporting tower such as Rohn SSV)
- Custom leg brackets (often used for smaller self-supporting towers)
- Single I-beam or mounting post (for crank-up towers)
- Special bolts supplied by the tower vendor

Some professional engineers and tower installers believe the base and its reinforcing metal should be part of the tower grounding system to divert lightning energy into the earth. Bases designed for this are said to provide an Ufer ground. Others prefer the traditional approach of multiple ground rods surrounding the tower base in all directions, with heavy-gauge ground cables running between the bottom of the tower and the rods. See Chap. 33 for additional discussion of grounding techniques.

Standard concrete reaches 99 percent of full strength about a month after it is poured. The curing process is exponential, so many experienced contractors will begin to attach hardware within a week of the pour, making sure no extreme stresses are applied to the concrete pad, either directly or through the lever arm of installed tower sections. Commonly, the first 20 ft of tower is held plumb with temporary guy wires during the pour and cure cycles.

32.11 Guy Wires and Anchors

If you are installing a guyed tower, you may be using concrete pedestals for the guy anchors. This requires a minimum of three, but possibly four if you're putting up a foldover or even six if the tower is very tall or you want some redundancy in your guying system. There are three reasonable guy anchoring approaches to choose from:

- Conventional concrete block guy anchor
- Earth screw-in anchors (called screw anchors for short)

- Special screw eyes held in boulders with equally special expanding cement

Examples of inappropriate and unacceptable guy anchors are:

- Houses, garages, sheds, or other buildings
- Utility poles
- Trees or other living things
- A spike hammered into the top of a tree stump
- A steel or iron post anchored in concrete

Occasionally a steel post anchor appears to be an attractive way to deal with a difficult guying situation, such as when the lower end of a guy wire needs to cross over a driveway, parking area, small shed, or sidewalk. Typically, the plan is to use a 10-ft-long steel tube or I-beam of unknown strength, with 3 or 4 ft of it buried in a concrete pad and 6 or 7 ft sticking up in the air. The guy wire is then attached to a hole drilled near the top of the post. Unfortunately, the guy wire creates significant wind-generated stress on the post. The horizontal component of a 1000-lb tension in the guy wire is 707 lb. This acts with a moment arm of 6 or 7 ft around the point on the I-beam where it enters the concrete. This overturning moment of 4000 lb-ft or greater can become a 20,000-lb compressive force in the post, bending it right at the top of the concrete. The odds are high that the post is going to bend or break during the first windstorm.

The easiest, least expensive, and fastest way to create guy anchors in most soils is with screw anchors. Figure 32.8 shows a typical device, available from some tower manufacturers (such as Rohn) and from utility company suppliers such as Graybar or regional companies. One excellent anchor is the Chance 6-ft rod with 6-in auger. Some professional tower installers refuse to climb towers guyed with screw anchors. Compared to concrete pedestals, earth screw anchors have at least two shortcomings:

- The holding power of the soil may be unknown or may vary with flooding or weather conditions.
- The speed and likelihood of corrosion resulting from direct contact with soil of unknown characteristics is uncertain.



FIGURE 32.8 Typical screw anchor.

Charts relating soil type to screw anchor holding power are available. Some have inserted a screw anchor in their soil and measured the tension required to pull it out with a tractor or 4WD SUV. However, holding power is not the only important soil parameter. Soil acidity and the types of trees in the guy anchor area may have a major effect on anchor lifetime. To combat this, some users of screw anchors coat them with a tarlike substance or use other types of electrochemical corrosion prevention before installing them.

Screw anchors suitable for modest amateur tower installations can often be installed by hand, by screwing them into the ground

with a long pinch bar or other strong steel rod threaded through the eye of the rod. Many tool rental outfits have gasoline-powered drivers that make the job easier. In general, the harder it is to insert the screw anchor, the harder it will be for stressed guy wires to pull it out.

Whether screw anchors or concrete pedestals are used, it is important to apply only longitudinal forces to the anchor rods, which have relatively little strength off-axis. Some commercially available rods can be bent sideways by hand. Each guy rod should be installed at an angle that makes the combined force on it from all attached guy wires as close to in line with the axis of the rod as possible. If the guy wires attached to various heights on the tower are all tensioned to the same specification, the angle of the rod should be equal to the average angle of the guy wires—typically in the 45-degree range. Little more of the rod than the knuckle or attachment eyelet should be sticking up beyond the concrete pad or out of the earth. Failure to follow these rules will subject the anchor to the same kind of stresses discussed previously in conjunction with the steel post anchor.

Each guy wire should have a turnbuckle at its lower end, just above the guy anchor. Two special tools are helpful for tensioning guy wires:

- A come-along for maintaining a temporary tension in a guy wire while securing the wire
- A Loos tension gauge for tensioning the guy wires following installation

The Loos can be obtained from Champion Radio Products or may be found at sailing accessory stores.

Two acceptable types of guy wire material are EHS stranded galvanized steel and *Phillystran*. The advantage of *Phillystran* is that it is nonconductive and long lengths of it do not need to be broken up by insulators to avoid resonances. When *Phillystran* is used, the bottom end of each guy wire must end in a length of EHS steel to avoid possible breakage from abrasion, vandalism, or animals.

Steel guy wires come in different diameters, each corresponding to a tensile strength rating. Generally, guy wires are tensioned (on a calm day) to some percentage (10 percent is common) of their maximum strength rating. That adds to the vertical force on the tower legs but ensures that the wires are not slack. Guy wires that are not properly tensioned can allow tower tops to be snapped back and forth by wind gusts.

It's a good idea to insert heavy-duty insulators of the proper compression strength in guy wires if there's a chance the tower will ever be used as a vertical antenna (typically at 160 or 80 m). Retrofitting a tower installation with insulators is difficult. The tower itself can be directly grounded at its base and fed with shunt-feed techniques described in Chap. 11.

Ideally, all guy wires at a given height (or station) on the tower should come away at the same angle with respect to the tower or ground. When the land surrounding the tower base is not level, or when there is less space on one side of the tower than the other for positioning guy anchors, this isn't always possible. The top of the tower is in force equilibrium (i.e., stationary and vertical) only when the horizontal forces on it are equal and opposing. If guy wires come away at different angles with respect to the tower axis, it will take different tensions, *T*, in the three guy wires to produce identical horizontal components, T_H. In other words, don't try to make the tensions measured by your Loos gauge in the three guy wires identical if they come away from the same tower attachment height at different vertical angles. Trigonometry is your friend when determining what tension to expect in each guy wire.

Despite being made of rigid, strong steel sections, a tall, slender tower is capable of (often disconcerting) sideward movement far above its base. Blind tensioning of guy wires can lead to incorrect results, since applying too much tension to one guy wire will simply result in pulling the tower sideways at that guy wire's attachment point until the horizontal forces at that guy station are again in

equilibrium. Tensioning should always be followed by a check of the entire tower for trueness, such as by taking a carpenter's level up the tower. During the tensioning process, standing or lying at the base of the tower and sighting straight up each face in turn should provide a sense of straightness (but not whether it's truly vertical) to guide tensioning while minimizing the number of trips up and down with the level.

Steel guy wires have long been secured with cable clamps, but grips have also become popular, thanks to the speed and ease of their installation. Regardless of which method of securing guy wire ends you choose, there are many ways to fail. Follow the manufacturer's recommendations or refer more advanced books on the topic.

32.12 Erecting the Tower

A triangular lattice tower can be raised using a heavy rope over a high support, sections can be installed one by one, or the tower can be held in place from a cherry picker or helicopter while it is bolted to its base.

Figure 32.9 shows the first method. The tower is laid out on the ground and then raised with a heavy rope over a high support (such as the peak of the house roof). This method is extremely dangerous unless certain precautions are taken:

- An adjustable-height support must be placed beneath the tower. As the tower is raised, the support is continually moved by a helper so that it is in contact with the underside of the tower until the tower has rotated past the critical angle.
- Temporary guy wires that can be quickly lengthened or shortened by other helpers are needed on each side of the tower (i.e., at the two ends of the back wall of the house).

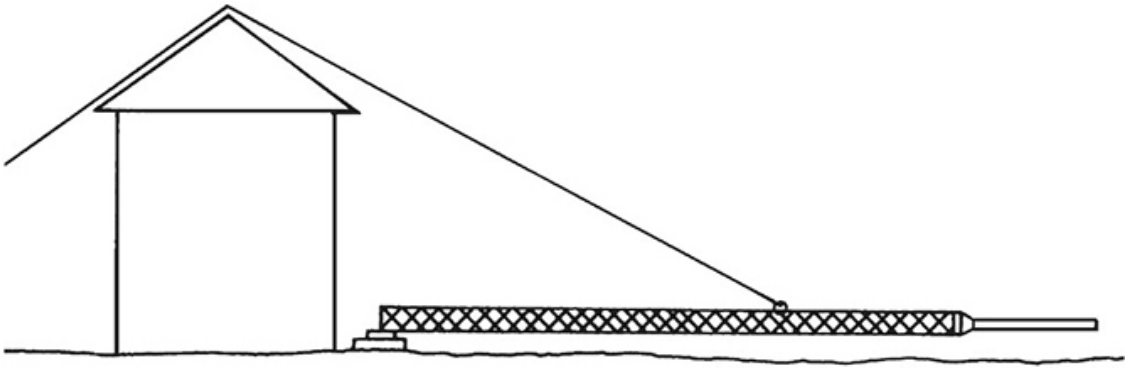


FIGURE 32.9 Dangerous method for erecting tower.

Towers over 40 ft tall should not be erected this way.

The section-by-section method of raising a lattice tower is shown in Fig. 32.10. A "gin pole" is required for this job. The gin pole is a length of thick-walled aluminum pipe, perhaps 2 ft longer than the length of the sections to be installed, and fitted with a pulley at the top and a pair of clamps at the bottom. A long piece of suitably strong rope such as that used by technical rock climbers (avoid nylon rope) is run through the pulley. Ideally the line should be twice the height of the tallest tower to be erected with the pole.

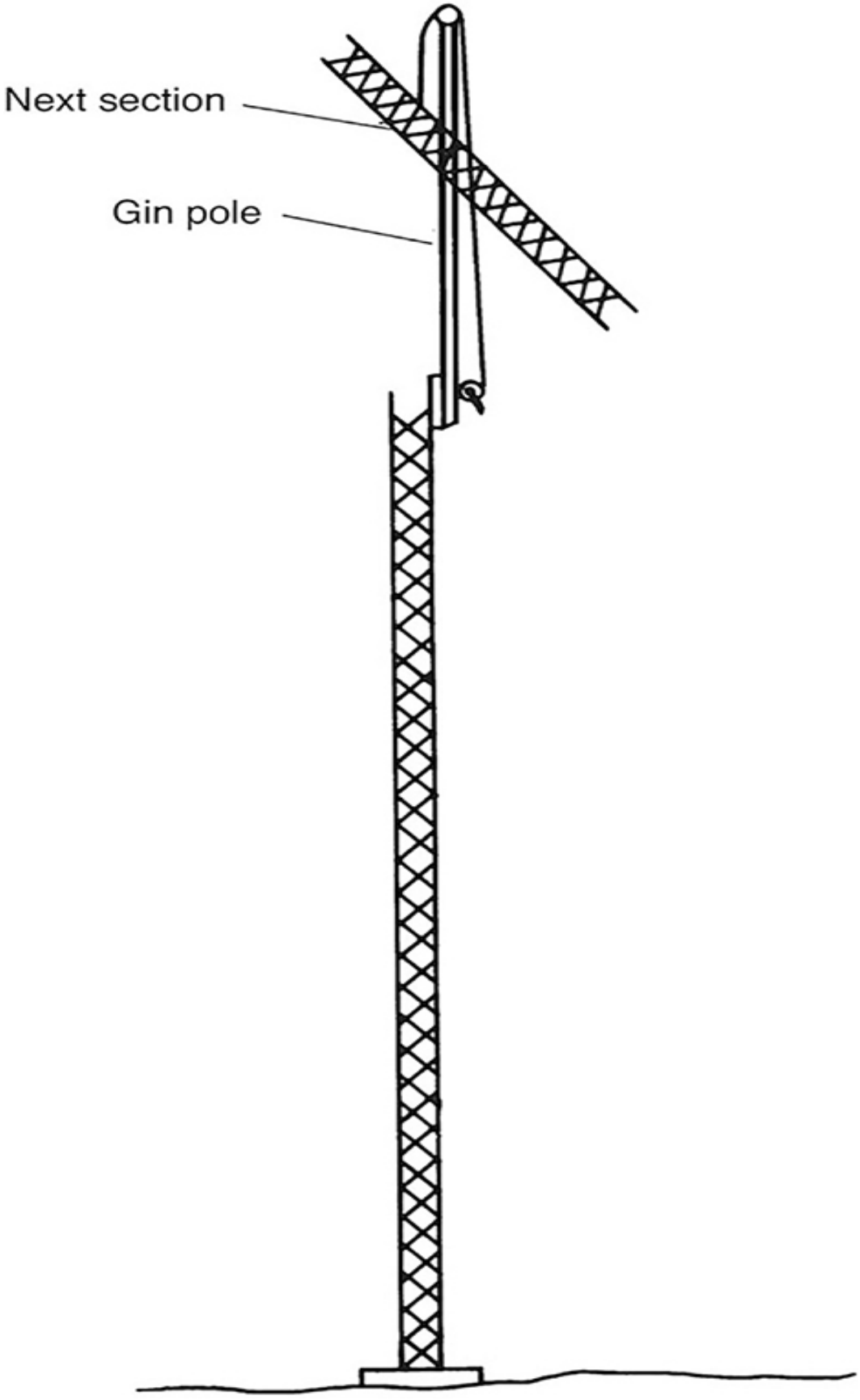

Next section
Gin pole

FIGURE 32.10 Proper tower installation using gin pole.

Caution *Wear a full-body fall-arrest harness when climbing a tower. Avoid free-climbing the tower. Always have two shock-absorbing safety lines connected to the tower above you while climbing, descending, or working in place, and disconnect only one at a time when changing position. Avoid the possibility of "single-point failures." Inspect your harness and associated hardware prior to each use. The lineman's climbing belt is no longer considered adequate because you can slip out of the belt and fall to your death if you are thrown upside down by an antenna or tower section that has gotten out of control. Excellent safety harnesses and shock-absorbing safety lines are well worth the cost.*

The gin pole is clamped near the top of the highest installed section so that it can be used to raise the next section into place. The ground crew hoists each new section up and holds it in position while the installer, who is belted to the tower just below the top of the old section, maneuvers the bottom of the new section directly over the top of the previous one. Together the installer and the ground crew gently lower the upper section onto the lower one so that the corresponding legs of each can be bolted together. The installer then carries the gin pole to the top of the newly installed section and attaches it near the top of this newest section. If this corresponds to a guy station, the ground crew uses the gin pole to supply the installer with a guy wire bracket that has the upper end of three guy wires already attached to it. Following the installation and tensioning of the three or four guy wires, the next section of tower is raised, and the process is repeated. Guy wires are typically installed at 30-, 35-, or 40-ft intervals, depending on the strength of the tower sections chosen. With Rohn 25 and 45 or equivalent tower sections, consider adding temporary guy wires when two sections have been added above the highest existing permanent guy station. Some installers are comfortable with three unsupported sections in low wind conditions.

The gin pole method is preferred by tower manufacturers and is the method recommended by most of them. For towers constructed from 10-ft sections, installation can be efficiently accomplished with two people, one on the tower and one on the ground. Towers with 20-ft sections and larger should be installed by professionals.

The third approach to erecting towers is to rent a cherry picker, crane, or helicopter and use it to pick up the entire tower or a major segment of it and hold it in position while it is bolted to the base or segment below it, and guy wires are attached and secured. The crew should include someone with experience in the use and stabilization of the hoisting machine. Cherry pickers can be rented quite reasonably from local equipment rental stores. They are most effective for erecting towers of 80 ft or less. Your tower site must provide solid and safe access for any machines you plan to bring in.

Caution *Wear a hard hat when installing a tower or its antennas. This is especially true for the ground crew because even a small stainless steel bolt or hand tool can cause great injury when dropped from above. But it is also true for any tower climbers, as it's easy to crack your skull on the underside of a guy bracket or boom-to-mast bracket while ascending or while wrestling an antenna into place.*

At some point you may decide that your best course of action is to have your tower installed by professionals. There are many competent tower installation companies throughout the world, including companies with expertise in tower and antenna installations for radio amateurs.

32.13 Rotators, Masts, and Thrust Bearings

Many antennas placed atop a tower are intended to be rotated because their radiation patterns are highly directional and the desired direction of communication is subject to frequent change. This is true of most Yagi and cubical quad installations. The typical installation of such antennas atop a guyed tower is shown in Fig.

32.11: An antenna rotator mounted to a shelf a few feet below the top of the tower turns a strong mast, which is often prevented from tilting and binding by a thrust bearing at the top of the tower.



FIGURE 32.11 Tower top showing rotator on shelf, mast, thrust bearing, and rotator bypass loop in feedlines.

To allow for rotation, the coaxial cable must be looped. One end attaches to the rotating assembly (usually by being taped to the mast, just above the top of the tower or rotator, whichever is higher) and the other end to the highest practical nonrotating tower surface (usually by being taped to a tower leg). Leave enough slack to allow full rotation of the antenna in both directions without putting any tension on the cable. Dress the loop carefully so that it does not catch on the top of the tower and check its operation by having a helper rotate the antennas while you watch how the cable coils and uncoils. For the minimum required loop length, determine where either end of rotation is and attach the cable below the rotator on the leg of the tower that is most nearly directly opposite the end-of-rotation point. An alternative approach involves coiling a turn or two of the cable around the mast just above the top of the tower so that it coils and uncoils just like a spring as the antennas turn. Most hardline and some kinds of coaxial cable are not meant to be flexed frequently and should not be used to form the rotator loop.

The standard mast outside diameter (OD) for all but the largest amateur Yagis is 2 in. Often it is possible to modify older brackets for use with a 2-in mast by simply drilling two new holes in the bracket and replacing the U-bolts. The best mast material is galvanized seamless steel tubing designed for this purpose. Rohn sells excellent 10-ft masts, and specialty metals suppliers can provide even longer ones. The galvanizing process results in a worst-case OD of slightly more than 2 in, so rotator clamps, thrust bearings, and other support tubes should be sized somewhat larger than 2 in. The mast clamp provided with the MFJ/CDE *Ham-M* and *TailTwister* rotators can handle diameters up to 2-1/8 in.

Some rotator manufacturers make a lower bracket available for mounting the rotator at the very top of a mast, but this type of installation should be used only for small VHF and UHF antennas

with minimal wind load and torque specifications. Even then, only a short (1 or 2 ft long) mast should be used. Most tower manufacturers sell an accessory rotator shelf to fit inside their towers or tower sections. Typically one of these plates is mounted about 2 ft below the top of the tower. A second plate is sometimes used to provide a thrust bearing or clamp support just above the top of the rotator so that if maintenance or repair work on the rotator is required, it can be removed from the tower while the mast and antennas remain clamped in place.

Although some amateurs have successfully used thick-walled aluminum tubing for mast material, this should be limited to Yagis with short booms and elements that apply only limited torque to the mast and rotator. Aluminum of any thickness is no match for steel, and the teeth of a typical rotator mast clamp or the threads of a pinning bolt can eventually chew through even solid aluminum tubing.

For small and medium-size HF Yagis and virtually all VHF and UHF beams, a simple piece of hardwood can serve as a perfectly adequate thrust bearing. Use an inexpensive circle cutter set to the next available diameter greater than 2 in. The tops of most commercial towers and tower sections sold into the amateur radio market have been predrilled or prepunched at the factory with standard hole patterns for thrust bearings and commonly available rotators.

While some amateurs in areas with low wind loading characteristics have had success using larger TV rotators, those devices should not be used to rotate anything larger than small (i.e., short-boom) HF beams—such as a three-element triband Yagi or two-element quad—in a benign environment. Rotators, such as the MFJ/CDE *TR-44*, in this class were originally intended to turn nothing larger than long-boom VHF deep fringe TV antennas; they are sometimes adequate for turning Yagis such as the Cushcraft *A-3S*, Force 12 *C-31*, Mosley *TA-33*, or equivalent. For larger tribanders (Cushcraft *A-4S*, Force 12 *C-32*) and small monobanders, the venerable MFJ/CDE *Ham-M*, or others of equivalent torque and braking capability are appropriate.

The larger MFJ/CDE rotators—notably the *Ham-M* and its big brother, the *TailTwister*, are unique in that the heading of the mast and antennas is locked by a strong brake wedge once the rotation has ended. There are only 60 positions of the wedge, so pointing accuracy is never better than within a 6-degree segment. This is not an issue at HF but possibly could be with very high gain, narrow beamwidth VHF or UHF arrays. Sometimes the wedge gets jammed by transferred wind load or other sideways forces on the rotator, and the only way to get the rotator turning again is to power the motor in the opposite direction for a few seconds before finally heading off in the desired direction. To minimize this problem, a start-of-rotation reversing sequence can be included in the controller firmware.

As boom lengths increase beyond 26 ft or so, larger rotators are required. In addition to the *TailTwister*, excellent units are available from ProSistel,

M^2

(pronounced "M squared"), Orion, and Yaesu. Many amateurs are using heavy-duty rotators from these companies to turn one or more antennas with 40-ft booms on a single high-strength mast. Most of these models have been designed to fit inside any of the top sections in the Rohn 25 product line, and most use mounting hole locations compatible with those that are prepunched in Rohn's rotator mounting plates. For the longest booms, and for rotating the entire tower, modified prop pitch motors are popular.

The most important specification for a rotator is its braking torque—that is, how well it can withstand the force of the wind on the antennas when they are at rest and not being turned by the controller. Some users eschew rotators that employ a mechanical brake because they prefer the soft drift of antennas to the full transmittal of shocks from wind gusts. In those rotators that do not have a separate brake, braking is usually accomplished by a combination of gear ratio and back electromotive force generated by the rotator motor.

Regardless of brand or strength, all rotators require cables to be run from a control box that is usually located indoors, close to the

radio equipment, where it can be conveniently accessed by the operator. Many include RS-232 serial or USB ports so that logging programs and custom software can access and control them from PCs. Controllers like the Green Heron RT-series are available that work with virtually every rotator that has been on the market, including prop pitch motors.

All rotators have end-of-rotation limits, or stops. In general, it is not wise to rotate antennas into either end stop at full speed. Some controllers, such as the Green Heron, include ramp-up and ramp-down circuitry or software to minimize the overall stresses on the tower and rotator when starting and ending rotation. Some rotators allow rotation in excess of 360 degrees and in some cases an additional 180 degrees. This is a great convenience for users such as DXers who may be following propagation with the heading of their Yagis or quads. This requires an extra-long rotator loop in coaxial cables that pass the rotator on their way up the mast.

Each type of antenna rotator imposes different requirements on the rotator cable that runs between the control box and the tower unit. The CDE *Ham-M* standard requires an eight-wire cable, but two of those wires have the high solenoid current of the brake wedge and require larger-diameter wire than the other six wires. Belden Wire offers a jacketed cable with two #16 and six #18 conductors that is designed for moderate runs to a *Ham-M* or *TailTwister* rotator, but any combination of cables of sufficient gauge can be used. The required wire size depends on the total voltage drop of the rotator control signals over the total length of the run. CDE has converted this to a specification on the maximum allowable resistance for each individual wire; this information can be found in the manuals for their rotators. Make sure the cable or individual wires you select are suitable for outdoor and/or burial applications. Protecting your rotator cables by running them inside PVC or other conduit material (with proper drainage for condensation) is desirable. Some wire insulation and cable jackets appear to have a chemical makeup that is more attractive than others to animals. Conduit can be used for cable protection.

To reduce the number of wires running between a CDE rotator and its controller, the motor start capacitor in the controller can be replaced with one of similar capacitance and voltage rating at the bottom of the tower unit. This will eliminate the need for two of the eight wires coming from the controller. Motor start capacitors can have a large temperature coefficient, so one mounted outside at the rotator or tower base may lose much of its rated capacitance during cold winters, leading to some potential for sluggishness compared to its summer performance.

Cabling requirements for the Orion and M^2 rotators and controllers are different. Beam heading information is obtained from a pulse-counting system in those units, whereas CDE rotator control boxes get headings by reading the DC voltage on the wiper of a wire-wound resistor located inside the upper housing of the tower unit. Orion and M^2 don't employ a brake wedge, so there is no brake solenoid to power. A separate cable for the pulse-counting position indication system may be needed to avoid crosstalk.

32.14 Preowned Towers and Accessories

With respect to obtaining and installing used towers of any type, as well as tower accessories, proceed with caution. Tower sections (especially those with hollow cylindrical legs) that have been in service or stored outside for long periods may be weakened by hidden rust. It may be possible with sandblasting, inspecting, and hot-dip regalvanizing to restore tower sections for less than the price of new sections. Discard galvanized guy wire and other critical hardware that is showing rust or black streaks. When replacing tower bolts, use the manufacturer's recommended replacement parts. Bolts of a given size come in many different strength ratings, and those you can find at your local hardware store may be inadequate.

32.15 Additional Information

At the beginning of an antenna project, review the spec sheets and engineering documents available from the tower and antenna manufacturers and the official building code, zoning, and wind speed requirements for your county. A professional engineer or local codes enforcement staffer associated with the tower project should have access to *TIA Standard RS-222-G* and be skilled in applying it to a specific installation. This document, available from the Telecommunications Industry Association, is available in offices of professionals who deal with towers and other antenna support structures. Books devoted to the topic of antenna tower construction are also available.

CHAPTER 33
Grounding for Safety and Performance

In radio transmitting or receiving installations, there are four main types of ground systems to consider and implement:

- Grounding for lightning protection
- Grounding for power distribution safety
- Grounding all station equipment together
- Grounding for maximum antenna efficiency

The first three are absolutely necessary; the fourth may be important, depending on the types of antennas in use.

Note *A good lightning ground is not necessarily a good radio ground, and neither of them necessarily provides a safe power distribution or equipment ground. For a proper communications installation, all four types of ground systems must be considered, designed, and installed. In many installations, components of any of these ground systems may well function as components of one or more of the other three systems.*

33.1 Grounding for Lightning Protection

Lightning is unpredictable but can be understood well enough statistically to provide guidance on protecting structures from damage and users from injury. Electrical potential differences of millions of volts build up between objects on the earth and particles in the earth's atmosphere directly above. At some point, this voltage becomes large enough to jump across the air by ionizing or breaking down the air molecules. A typical breakdown figure is something less than 100,000 V/in of dry air. The exact voltage

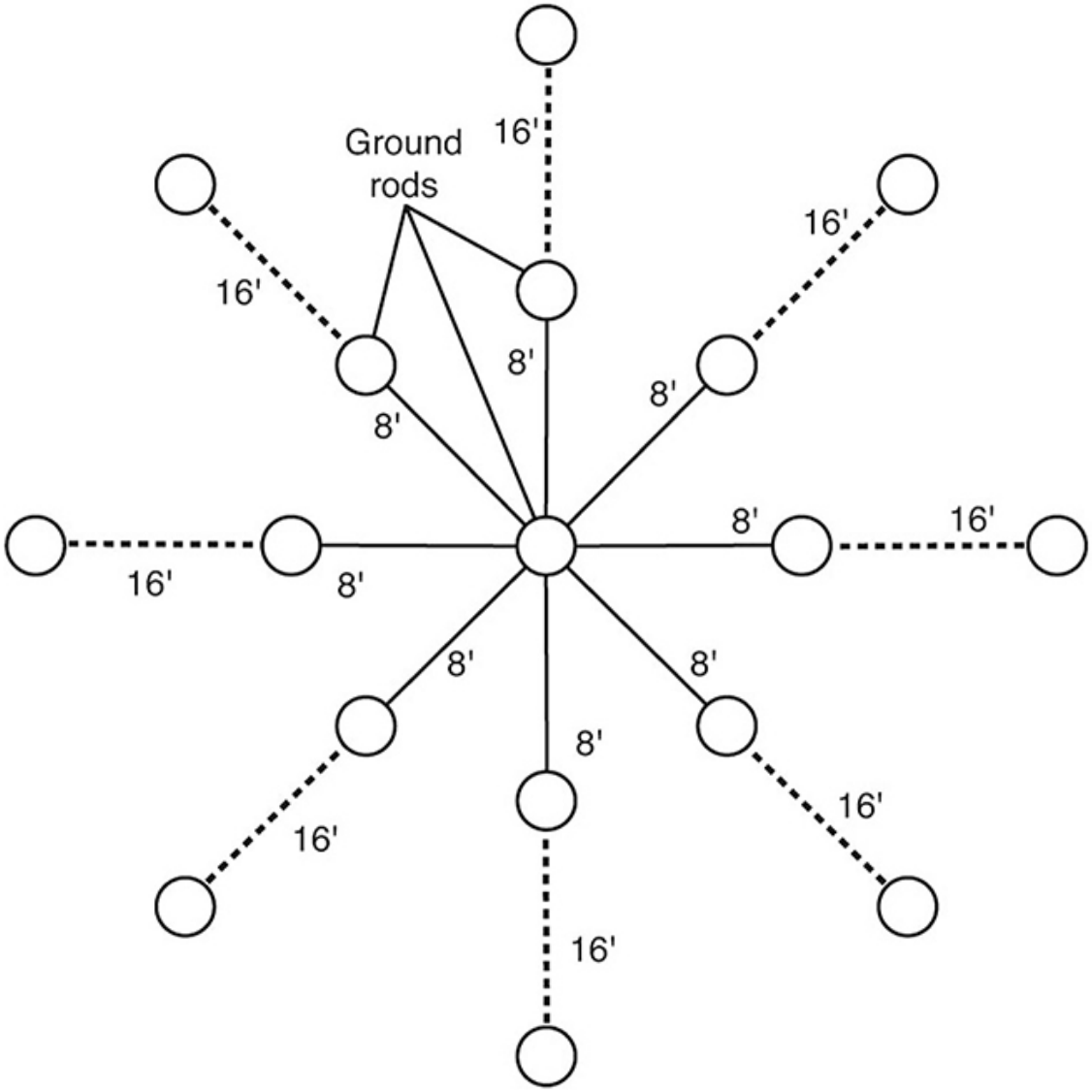
depends on the shapes of the two electrodes, and there is much about high-voltage phenomena that is nonlinear. When there is moisture in the air, the voltage required to ionize the path to the earth drops substantially. The jagged shape of a lightning bolt is the result of the stored electrical energy in a storm system taking a path of least resistance through moving air masses.

Antennas and supports can be targets of frequent lightning strikes. In a thunderstorm, we avoid standing under tall trees, and the same idea holds for avoiding radio towers and utility poles. A typical stroke passes peak currents of thousands of amperes (20,000 A is often cited). The heat generated by such a large current is sufficient to blow apart concrete tower bases and vaporize not just solder connections but entire antennas. In addition to immediate death or permanent disability caused by a lightning strike, there is a comparable risk from exploding objects, extreme radiated heat, vaporized moisture, and flash fires from overheated wires in the walls of wood frame dwellings.

There is no known way to totally prevent a lightning strike. The first purpose of a ground system for lightning protection is to minimize the damage caused to people and property when lightning does strike. The current professional consensus is to attempt to harmlessly dissipate in the loss resistance of the nearby earth as much of the energy in the lightning bolt as possible.

The usual method for dissipating the energy in a lightning surge is to attempt to spread that energy to a large area of ground surrounding the tower and antennas. This is done most often by connecting the base of the tower to a set of lightning radials comprised of straight runs of #6 AWG or larger copper wire or flat copper strap with a cross-sectional resistance no greater than that of #6 wire. Soil conditions permitting, each radial should be connected mechanically or with a brazed joint to multiple 8-ft ground rods of the type found at electrical utility suppliers such as Graybar. As shown in Fig. 33.1, the first ground rod on each radial should be one ground rod length (or 8 ft, in this example) from the base of the tower, and all subsequent rods should be 16 ft (i.e., twice the length of the rod) beyond the previous rod. The minimum

practical number of radials is probably 4, with 8 to 16 being a better choice. Two or three rods per radial probably constitute a minimum system.



F IGURE 33.1 Tower lightning ground system.

Since the objective of lightning radials is to dump currents from the lightning bolt into the soil, there is no reason the radial wires (and the tops of the ground rods) can't be buried many inches under the surface of the ground. If the lightning radials are to be a part of a ground-mounted vertical antenna's radial system, to avoid

ground losses and a reduction in antenna radiation efficiency, they should not be located more than a few inches below the surface. When soil conditions do not allow 8-ft ground rods to be fully pounded in, shorter rods can be used, with some lessening of performance unless the number of radials and rods is increased. As a last resort in very rocky terrain, ground rods can be laid horizontally and buried just under the sod.

Spectral analyses of lightning surges show the bulk of the RF energy to be in the 1- to 3-MHz range. A lightning ground should exhibit a low impedance not just at DC and power line frequencies but up through 3 MHz or more. The radial wires and straps being used to dissipate the energy should avoid sharp bends or other sources of excessive inductance.

Ohmic resistance at a junction in the system should be avoided. The heat generated when lightning surges through ground wires and their connections is well above the amount needed to melt solder, so connections in a lightning ground system should be brazed or mechanically clamped.

The first priority is protection of living beings, followed by protection of property and equipment. Keeping energy from a lightning strike outdoors is the first line of defense. Given the many condition paths over RF cables and power lines into structures and dwellings, a second layer of protection may be needed. Some users disconnect all electronic equipment prior to the arrival of a storm or following each period of use. This means disconnecting power cords, headphones, telephone lines, all audio and control cables, and grounding wires.

The state of the art in electronic equipment protection is a single-point ground (SPG). Signal and control cables coming into or leaving the radio room or building are passed through a metal panel or enclosure near the point of entry (Fig. 33.2). The panel should be ground-mounted or physically as close to earth as possible and solidly attached to its own lightning radial system. Utility grounds should be close to the SPG panel and to each other and electrically bonded together and to the SPG ground. Any other conductors such as cable or satellite TV system cabling that enter the building should

pick up their earth ground at the same point. If the building has a protective lightning rod system, it should be connected here.



FIGURE 33.2 Single-point ground (SPG) system brings grounds for all utilities, antenna feedlines, and control lines together at a common entry point.

When lightning strikes damage electronic equipment, the cause is often a difference in the induced surge voltages on different cables or within different systems. All radio transmission lines, rotator control cables, and remote switch control lines must go through the SPG. The GND lead for a set of control wires should be grounded at the SPG. The same is true for the shield braid of each coaxial cable entering or exiting the building. Open-wire transmission lines should, at a minimum, have spark gaps from both sides of the line to the SPG ground. Remotely operated contactors can be used to short out both sides of the transmission line to SPG ground when that line is not in use.

Driven leads such as the center conductor of coaxial cable or control lines to remote antenna selectors and rotators can be handled using relays to ground the lines when not in use or by passing every hot lead through a lightning protection device made by ICE, Alpha Delta, and others (see Fig. 33.3). Some of these designs employ metal-oxide varistors (MOVs) between each control line and ground. Others use gas discharge tubes for protecting any equipment connected to the hot center conductor of coaxial cables. These devices can fail open, making it impossible to know whether they're still operational or not. Some MOVs fail gradually, becoming less and less useful with each nearby lightning surge. The least expensive devices are shunt protectors that conduct a lightning-induced surge to ground. Another category, generally costing more but preferred by some experts, employs series protection techniques.

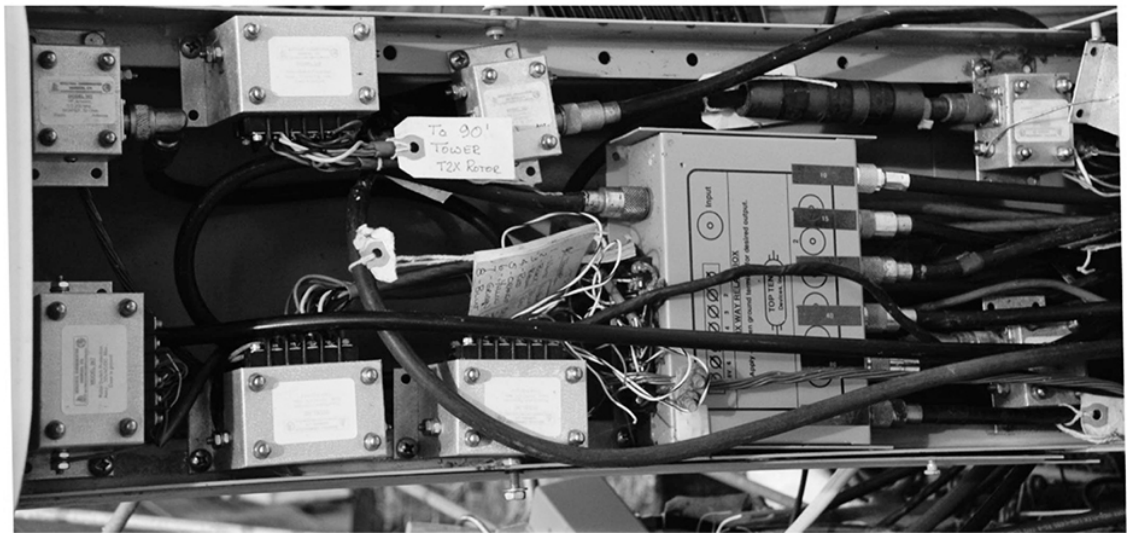


FIGURE 33.3 SPG enclosure contains surge suppressors on all control lines and antenna feedlines.

Coaxial signal cables and rotator or relay control cables that have been routed down the legs of metal towers should have their shields and GND leads connected directly to the tower at both the top and bottom. The bottom connection should be tied to a lightning ground radial field emanating from the base of the tower

in all compass directions. These connections should be mechanical or brazed, since solder will not survive a lightning surge. Cables dropping from masts and other antenna support structures, including trees, should have their shields and GND leads connected to an earth ground at both the top and the bottom of the vertical span. For maximum protection from lightning surges, the key directive is "ground everything well and often!"

Note *Grounding cables at the tower is not a substitute for a good single-point ground at the point where all the exterior cabling enters the building. A good lightning ground at the tower, while highly desirable, has nothing to do with protecting the electronic equipment inside the building from lightning-induced surges on power lines, telephone lines, or even the cabling between the tower and the building.*

Open-wire lines present a special challenge. It is possible to run each side of an open-wire transmission line through its own grounded protective device, but keep in mind that voltages on an open-wire line used in a high-SWR configuration (feeding an 80-m dipole on 40 m, for instance) will be much higher than on a matched OWL or coaxial cable for the same transmitter power levels. Another approach (that can be used in combination with the first) is to pass each side of the open-wire line through a wide-spaced open-frame relay with Form C contacts just before the line enters the building. To be sure the relays fail safe when power is off, the line connected to the antenna should go to the wiper or common contact, and the normally open (NO) contact should connect to the section of OWL that enters the building. The normally closed (NC) contact should be run on as short a path as possible with heavy-gauge wire (#6) to the SPG or lightning ground system, whichever is closer. Even better, the OWL would go to a remotely tuned ATU in the general vicinity of the antenna and protected coaxial and control cables would run from the doghouse to the SPG. NC contactors and a lightning radial field for the ATU would be located at the ATU in a small box or shelter.

Perhaps the best practical approach available to most of us with budget limitations, given the present state of the art, is to employ a three-stage philosophy of equipment protection:

- Ground all transmission and control lines at the top of their respective towers and to lightning ground systems attached to the bases of all towers.
- Use a single-point ground just outside the building that houses the radio room; make sure all incoming utilities are grounded to a suitable lightning ground system at that point and enter the building very close by (i.e., within a few feet).
- At the SPG, mount an enclosure to hold lightning protection devices for all signal and control cables entering and exiting the structure. (See Figs. 33.2 and 33.3.) For an extra measure of protection, include relays or contactors with wide-spaced contacts to disconnect the hot leads of all cables and controls entering the building when they're not in use.

While not a perfect solution, since large surges can still jump these gaps, protection from lesser surges and indirect strikes will be improved.

33.2 Grounding for Power Distribution Safety

The second form of ground every radio station installation must have is a power system ground. The purpose of this ground is to ensure the safe operation of appliances and electronic equipment that uses external power.

At least in the United States, current practice is to wire homes with a three-wire system. One wire at each wall outlet is hot (usually around 120 V, 60 Hz, and usually with black insulation), one wire (the neutral line, usually white) is a return wire for the current used by the appliance or device, and the third wire (usually green) is the safety ground. On devices and appliances with a three-wire

power cord, the green wire is required to be directly connected to the metal chassis of the device.

Based on a typical electrical code, the only place the white and the green wires are to be connected together is back at the main breaker box, where utility power enters the building. Connecting those two wires together in a subpanel or at an outlet strip alongside your operating desk is prohibited.

Power system ground is a DC and low-frequency ground. Impedances can be assumed to be low only for frequencies below 120 Hz or so, and only with respect to the power utility's neutral wire. The earth ground outside your residence does not provide RF grounding. It is important to keep the power system ground tied to other system grounds at one specific entry point to the building so that all grounds inside the building rise and fall with the same surge voltage during a thunderstorm.

33.3 Station Grounding

It does no good to provide a topflight ground system outdoors or throughout the structure's wiring if the interconnections between the various pieces of station equipment and the external ground systems are substandard. Figure 33.4 shows a method for station grounding. On the back of the operating position is a sheet of copper, 7 in wide, running the length of the equipment platform. This form of copper, weighing $1\ \mathrm{lb/ft^2}$, is sometimes used for roof flashing. An alternative is to run a buss bar of 1/2-in or 3/4-in rigid copper plumbing tubing the length of the equipment table. In either approach, each piece of equipment is connected to this ground buss or ground plane through a short length of braid or #12 wire, secured at the buss with hose clamps or sheet metal screws. Small RF accessories (e.g., a low-pass filter) that have no user controls on them are mounted directly to the copper sheet.

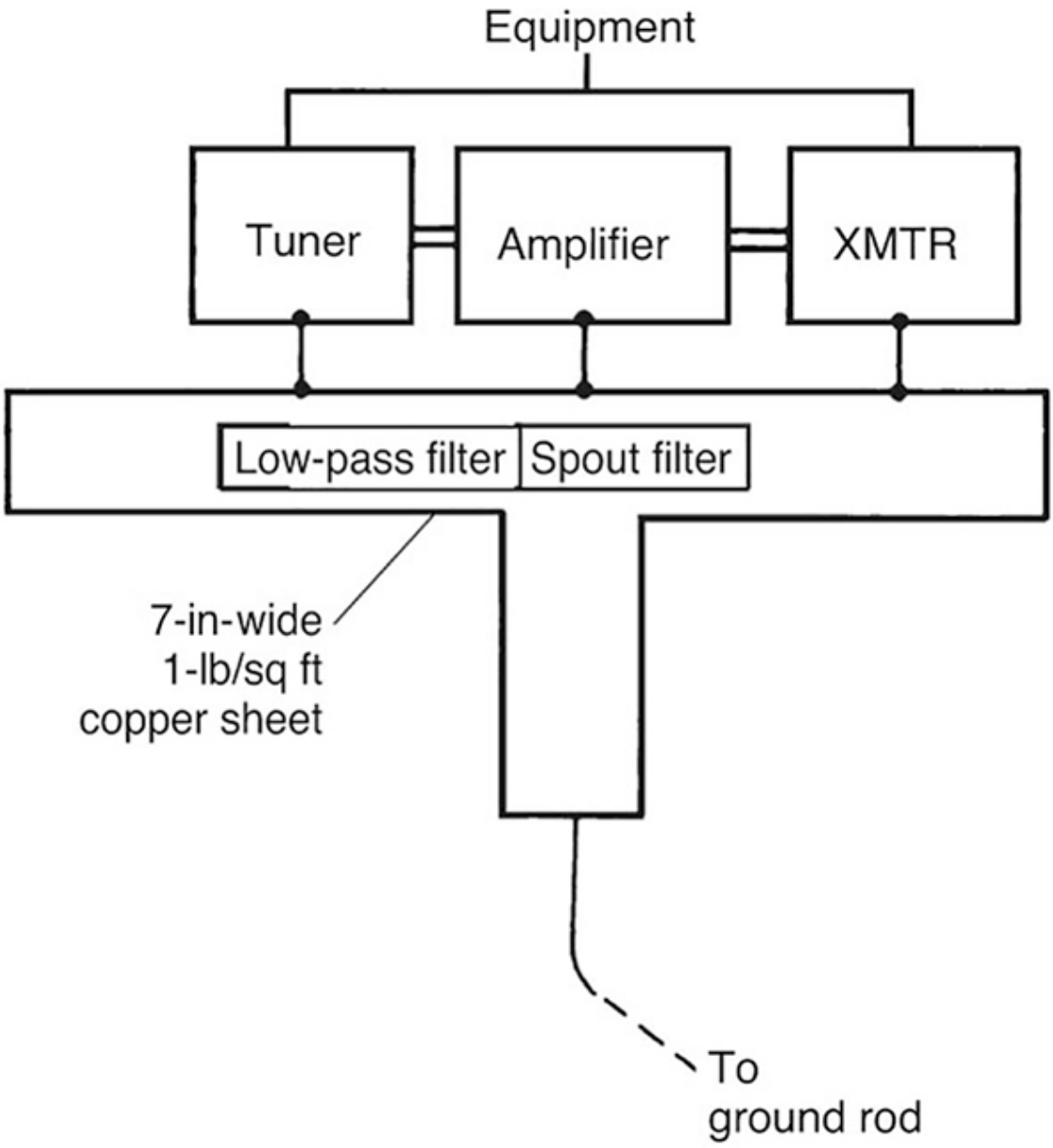


FIGURE 33.4 Ground system inside shack.

These connections are an important part of the other grounding systems discussed, and they must be low impedance all the way from DC to many megahertz. The ground plane or ground buss should be connected to the outdoor ground system with as short a run as possible, so that the least possible resistance and inductance are introduced into the ground path.

Avoid using building electrical ground wiring as an RF antenna ground. Neglecting to install an outdoor ground that will properly do the job, some opt for a single connection to the grounded wire in an electrical outlet. In addition to being potentially dangerous, this is a poor RF ground. It is too long for even the lower HF bands, and it reradiates RF around the building. Transmitters that depend on household electrical wiring as the radio ground can cause radio and TV broadcast interference as well as interference to other electronic devices in their own homes and in nearby buildings.

33.3.1 Ground Wire Tuner

If you use a power tool or a household appliance on the second or third floor of your home, or at the far end of a long three-conductor extension cord, you still have an effective power safety ground at the point of use, despite the relatively long run from your circuit breaker box in the basement. From a lightning or radio standpoint, however, second-and third-story grounds are not grounds at all. At radio frequencies, RF equipment with a long power supply connection may be a substantial fraction of a wavelength distant from true earth ground.

If you have ever been burned by RF energy when you happened to touch the chassis or a metal knob on your radio equipment while transmitting, you have felt firsthand the effect of having an inadequate RF ground in your radio room. Often the user thinks he or she has a ground simply because a wire has been run from the equipment chassis to a ground rod two floors below. That ground wire may be 25 or 30 ft in length, long enough to be a quarter-wave vertical on 20 m.

An alternative that some operators use is the ground wire tuner. These accessories insert an inductor, capacitor, or LC network in series with the ground line. The user adjusts the device for maximum ground current while transmitting on the desired operating frequency. These tuners can bring your radio equipment common chassis connections to a virtual RF ground for purposes of providing proper antenna tuning. A low-impedance RF path to a

single-point ground located outside the building at earth level is still necessary for lightning grounding.

33.4 Grounds for Antenna Efficiency

The success of a radio antenna system often hangs on whether it has a good RF ground. Poor grounds cause many antennas to operate at less than best efficiency. It is possible to waste as much as 50 to 90 percent of your RF power heating the lossy ground under the antenna. This is especially true in the case of ground-mounted monopoles, such as a typical quarter-wave vertical, because the circle of ground lying within roughly
$\lambda/2$
of the base of the vertical is part of the antenna. RF energy delivered to the feedpoint of the vertical monopole and radiated into space must be balanced by return currents delivered to the transmission line via this ground path.

The effective ground resistance of typical soils in the circle beneath a vertical monopole can vary from a low value of, say, 5 Ω, up to more than 100 Ω. The total resistance,
R_L
seen by the transmission line is

$$R_L = R_{\text{gnd}} + R_{\text{rad}} \quad (33.1)$$

where
R_{gnd}
is the effective loss resistance of the ground surrounding antenna and
R_{rad}
is the radiation resistance of the antenna. The total power delivered by the transmission line is

$$P_{\text{out}} = P_{\text{gnd}} + P_{\text{rad}} = \frac{1}{2} I_L^2 \left(R_{\text{gnd}} + R_{\text{rad}} \right) \quad (33.2)$$

The power radiated by the antenna is $P_{\text{rad}} = \frac{1}{2} I_L^2 R_{\text{rad}}$. The radiation efficiency is

$$\eta = \frac{0.5 I_L^2 R_{\text{rad}}}{\frac{1}{2} I_L^2 \left(R_{\text{gnd}} + R_{\text{rad}} \right)} = \frac{R_{\text{rad}}}{R_{\text{gnd}} + R_{\text{rad}}} \quad (33.3)$$

If the effective ground resistance is 30 Ω and the antenna is a full $\lambda/4$ vertical monopole having a radiation resistance of 30 Ω, then half the applied power is wasted in ground losses. Any increase in ground losses (usually as a result of an insufficient radial field) or decrease in radiation resistance (a shorter vertical) can conspire to drop the power radiated even further, perhaps to as low as 5 or 10 percent of the power available from the transmission line.

33.5 Soil Conductivity

Ground resistance is influenced by the type of soil and the water content. Electrical conductivities for different soil types are shown in Table 33.1. Moist soil over a brackish water dome as found in coastal swamps conduct best, whereas the sands of the western deserts make the poorest conductors.

TABLE 33.1 Sample Soil Conductivity Values

Type of Soil	Dielectric Constant	Conductivity (Siemens/meter)	Relative Quality
Saltwater	81	5	Best
Freshwater	80	0.001	Poor

Pastoral hills	14–20	0.03–0.01	Very good
Marshy, wooded	12	0.0075	Average/poor
Sandy	10	0.002	Poor
Cities	3–5	0.001	Very poor

Avoid using outdated techniques for reducing the electrical resistance of soil through treatment with chemicals. In addition to environmental concerns, any benefit from salting will only be temporary.

The soil directly underneath a horizontally polarized antenna influences the effective height of the antenna and its feedpoint impedance. The soil within roughly

$\lambda/4$

to

$\lambda/2$

carries radial return currents back to vertical monopoles. With the possible exception of saltwater, any surface normally found under a vertical monopole is not a good enough conductor to substitute for copper radials. For good vertical monopole efficiency, regardless of soil type, use at least two dozen radials.

Farther from the base of an antenna, soil conductivity affects the first ground reflection of your transmitted signal and the last ground reflection of received signals. For long-distance (DX) communication, optimum takeoff and arrival angles are usually in the 1- to 5-degree range, so bounces may occur up to a mile away.

33.6 Grounding with Radials

The effectiveness of an RF ground system is enhanced by the use of radials either above ground or buried just below the surface. A

vertical monopole antenna is particularly ineffective unless provided with a good RF ground system.

A large number of radials is normally required. AM radio transmitters use 120 or more half-wavelength radials, but that is overkill for amateur systems. Installing more than 30 or 40 $\lambda/4$ radials is generally unnecessary.

Unlike elevated radials, which should be as long as the electrical height of the vertical, radials on the ground can be any convenient length, up to a half-wavelength. A radial in close proximity to earth is nonresonant and precise lengths are not needed. A rule of thumb is to use radials that are a reasonable fraction of a wavelength long with the ends of adjacent radials about 0.05λ apart. At smaller spacings, no significant reduction in ground losses is noted, and when spacings start to exceed that figure, losses start to grow as an increasing percentage of the vertical's return currents find their way into the earth instead of the copper wires.

An ideal layout for a system of radials in a vertical antenna system is depicted in Fig. 33.5. The radials are laid out in a uniform pattern around the antenna element. This coverage provides both the lowest resistance and the best radiation pattern for the antenna. Tie all radials together at a common point at the base of the vertical, and connect that junction of wires to one side of the feedline, usually the shield braid when coaxial cable is used, and to the ground terminal on any remote matching network you may be using at the base of the vertical. If your radials are going to be part of your lightning surge grounding, use mechanical compression or brazed joints rather than solder for the connections.

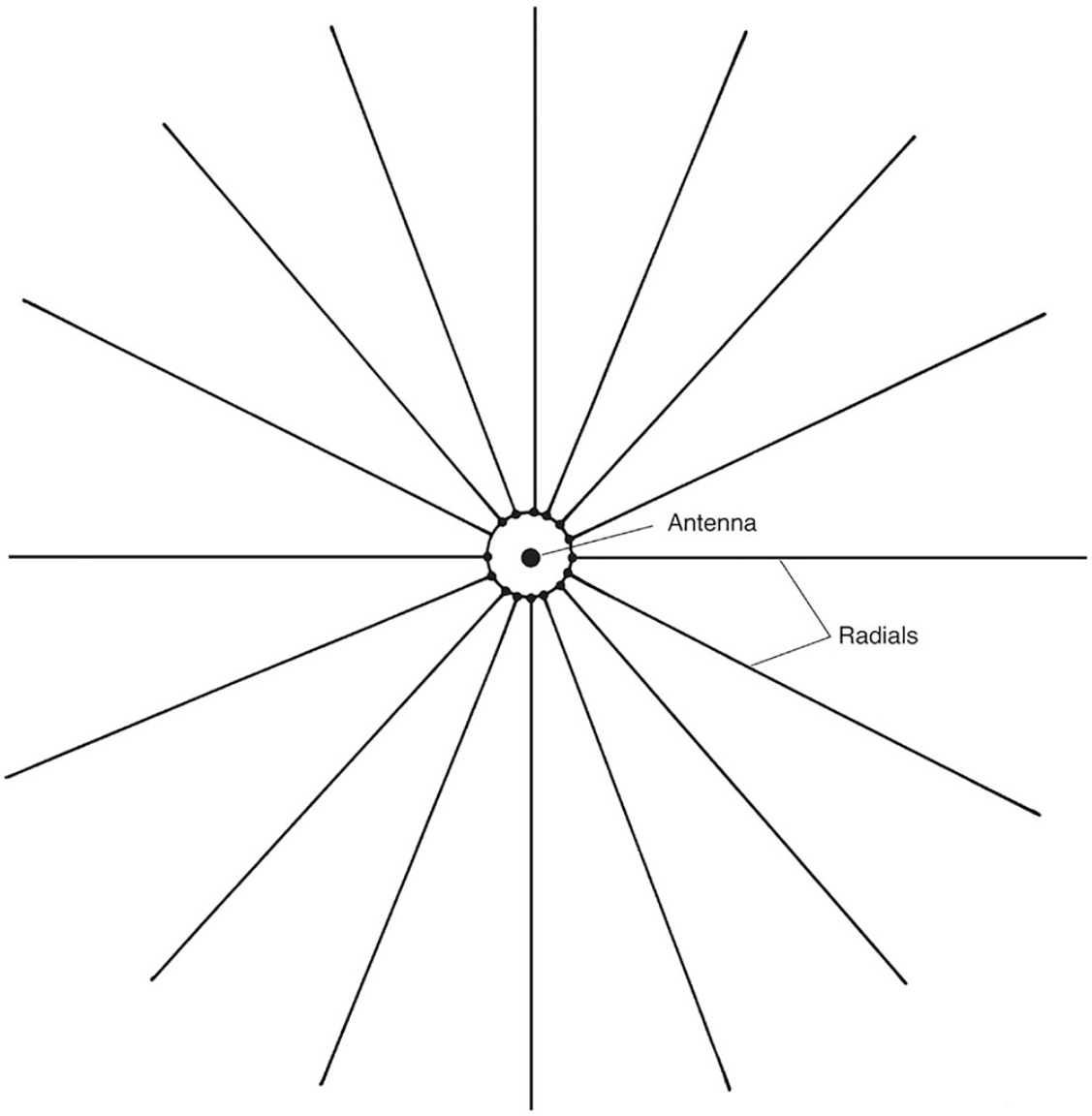


FIGURE 33.5 Radial pattern beneath a grounded monopole (vertical).

For verticals near a property line or a building, make the radials fit the circumstances by adjusting lengths or routing them around obstructions. Return currents from the antenna's fields are radial in orientation, so there is no need to connect radial tips with circumferential wires. Chicken wire mesh can be used but corrosion can degrade its performance. Use just about any kind of copper wire you have on hand for your radials. Strip multiconductor cables into their individual wires, if you wish. The primary determinants of

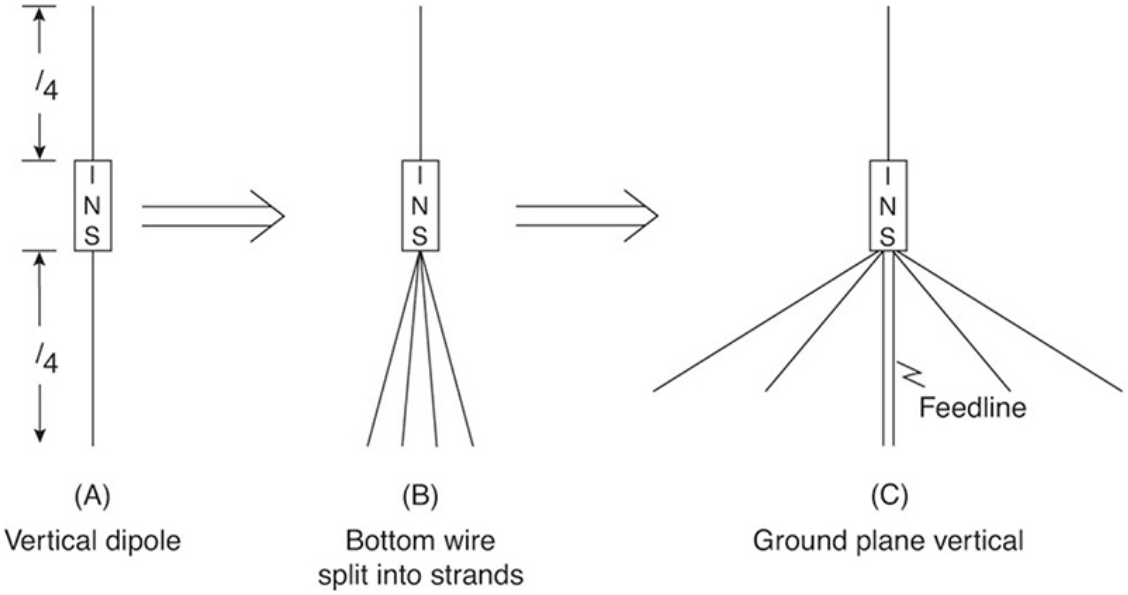
what wire to use are cost and longevity. If you have to purchase radial wire, use ordinary insulated THHN #14 house wire as the insulation will prolong the life of your radials. Even at kilowatt power levels, wire gauge is not critical. The primary reason for specifying #14 wire is so that it will stand up to stresses on it, such as when it is buried or when someone or something trips over it. Many radial fields consist of #18 or finer wire. Avoid stranded wire in outdoor applications.

Ideally, your radials should lie on top of the ground for maximum antenna efficiency, but you can bury them an inch or two below to keep people and animals from tripping on them. Sod staples can be used to hold radials down and out of the reach of lawn mower blades until they sink into the grass and become part of the thatch. Radial attachment plates can simplify the connection of large numbers of radials to the base of the antenna. Others have formed a grounding ring from soft copper tubing purchased from the local hardware store and attached the radials to it with solder joints or noncorroding fasteners. If you have multiple verticals (a four-square, for instance), do not connect the radials of one vertical to those of another. Never directly connect copper to galvanized steel. Use a transition metal in between the two.

33.7 Elevated Grounds

Most users of VHF equipment are familiar with the ground-plane antenna—a vertical with either three or four radials, mounted atop a mast. The same technique can be used at MF and HF, but the mechanical challenges are not trivial. At VHF, the idea behind the ground-plane antenna is simple: a vertical monopole (whether $\lambda/4$ or 5λ/8 or any length in between) requires a return path for the missing half of the charged dipole structure that is necessary for electromagnetic radiation. Starting with a vertical dipole (Fig. 33.6), if we split the wire that is the lower half of the radiating element into four thinner strips, angle them upward by 45 degrees, and space them equally around the azimuth circle, we end up with a

ground-plane vertical. Thus we have created an artificial ground plane high in the air, right where the vertical monopole needs it. Three radials spread 120 degrees apart are usually sufficient.



F IGURE 33.6 Transforming a vertical dipole to a $\lambda/4$ ground-plane vertical.

The ground-plane antenna is fairly easy to implement at frequencies as low as 7 MHz. One popular antenna a few decades ago was the Hy-Gain *14-AVQ* (now the *AV-14AVQ*), a trapped vertical for 40, 20, 15, and 10 m. Propped atop a 20- to 30-ft mast, with a handful of radials sloping down from the base of the antenna, the antenna worked well as a low-angle radiator on 40 m. In some installations, the radials did double duty as guy wires for the mast. In such an installation, each combination radial/guy wire must have an insulator in it at a distance of $\lambda/4$ along the radial from the base feedpoint of the radiating element.

In recent years, increasing numbers of hams have installed elevated radials under 80- and 160-m verticals. A handful of elevated radials (between four and eight) can be as effective as 30

or more radials on the ground as long as the vertical and the elevated radials are well above the earth. Elevated radials should be at least $\lambda/8$ above earth ground, but good results have been reported with lower installed heights. Few amateurs have the resources to put up 160-m verticals whose base is 30 to 70 ft above the earth, so elevated radials are frequently installed as shown in Fig. 33.7.

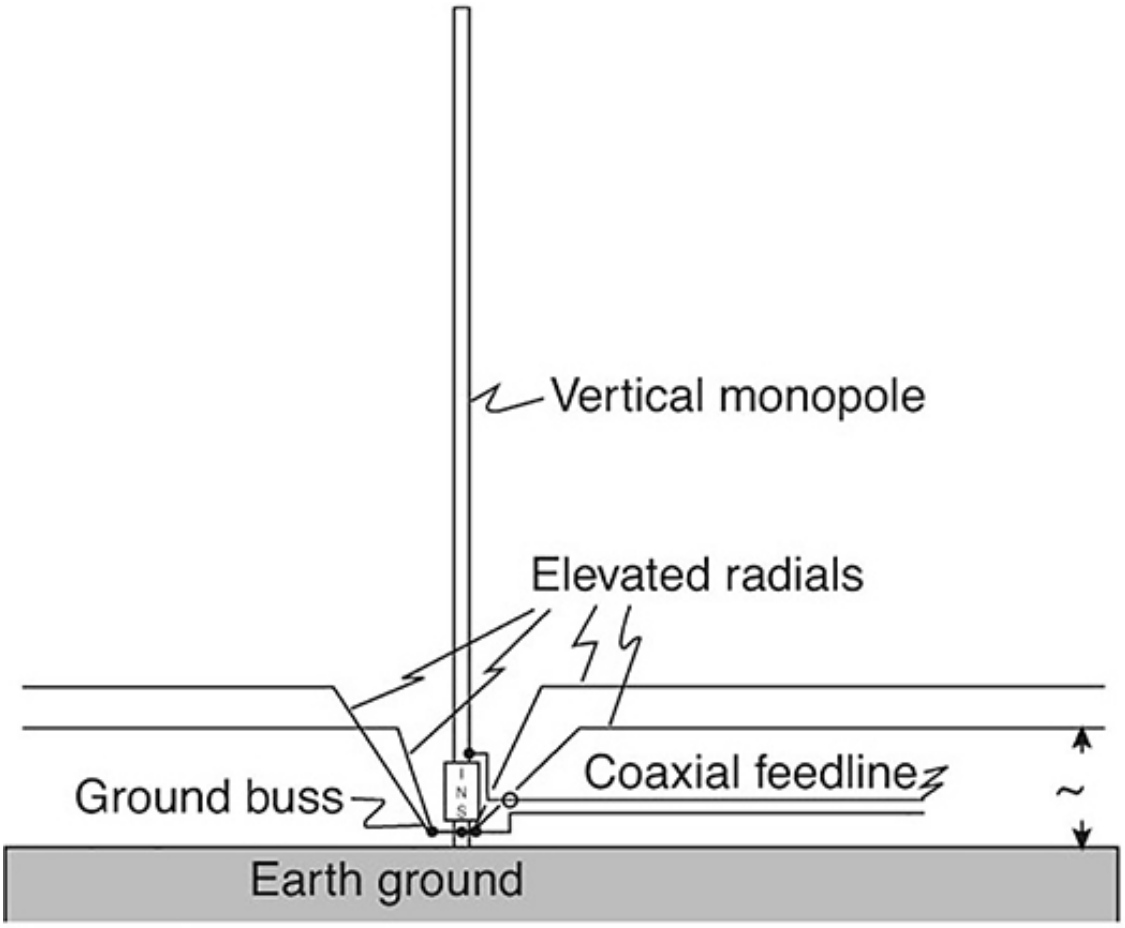


FIGURE 33.7 Elevated radials for a ground-mounted vertical.

Compared to earth ground, copper radials are essentially lossless. When elevated radials are too close to earth, radial return currents for the vertical are shared between the radials and earth, leading to lower radiation efficiency. When the vertical and the elevated radials

are $\lambda/8$ or farther from the earth's surface, virtually all the return currents are collected by the elevated radials.

Unlike radials on the ground, elevated radials exhibit resonance in the same way that an antenna does. Radials for a given band should be cut to the same length and should all make approximately the same angle with the ground, insofar as possible. If the vertical is capable of multiband operation, a set of

$\lambda/4$

radials should be added for each band to be used. The objective is to make the impedance at the feedpoint end of each active radial as low as possible in order to maximize the return currents it supplies the vertical monopole. One possible simplification is to add traps so that a single set of four or eight radials serves multiple bands.

CHAPTER 34
Zoning, Restrictive Covenants, and Neighbors

To a radio enthusiast, an antenna tower may be a thing of beauty, but to other neighbors it may be an eyesore. In populated areas, zoning ordinances and other regulations help to balance the competing needs and desires of people in a community. The difficulty comes when we view the value of our towers and antennas differently than our neighbors and local officials do, or when restrictions prohibit antennas and structures that would be ideal for radio communications. Challenges to erecting towers and antennas can come from governmental bodies, zoning ordinances, building codes, or restrictive covenants and deed restrictions.

34.1 Federal Regulations

In the United States, federal intervention takes two forms:

- Regulations to minimize hazards to aviation from man-made structures that rise far above the average terrain
- PRB-1, issued by the Federal Communications Commission (FCC) in 1985, preempting certain aspects of local regulation of amateur radio structures

Within the United States, the Federal Aviation Agency (FAA) maintains a database of towers and antennas in excess of 200 ft above average terrain or close to public airports. You must notify the FCC and the FAA prior to constructing a tower or other man-made antenna support structure that rises 200 ft or more above the ground around it. Depending on the specifics of your site, your

application may be acceptable, but it will require FAA-approved lighting.

For tower installations in the vicinity of existing airports, the FAA has published rules that set a maximum height for towers and antennas on them as a function of distance to the nearest runway. There are three formulas to be concerned with:

- For airports with one or more runways longer than 3200 ft, the top of your antenna or man-made support cannot exceed 1 ft for every 100 ft between the antenna and the nearest part of any runway at that airport. If you wish to put up a 70-ft tower, for instance, you must be at least 7000 ft (or approximately 1.3 mi) from the nearest part of the runway closest to your tower(s).
- For airports with no runways longer than 3200 ft, the top of your antenna or man-made support cannot exceed 1 ft for every 50 ft between your antenna and the nearest part of any runway at that airport. This effectively doubles your allowable antenna height at a given distance from an airport.
- For heliports, the corresponding limit is 1 ft of antenna height for every 25 ft from your antenna and any portion of the helipad that is used for flight operations.

If the elevation above mean sea level (MSL) of the nearby airport is substantially different from that at the base of your proposed tower, the calculation may be more complicated than these bullet points indicate. When in doubt, submit an application to the FAA and get a formal ruling for your files.

The U.S. FCC has issued a document now familiarly called "PRB-1". PRB-1 conveys limited federal preemption of local municipal restrictions on amateur towers and antennas. It is potentially useful to you, depending on the specific objections being raised by neighbors or local officials, but it is not a substitute for a clearly worded permissive ordinance and/or a building permit. You may encounter a state or locality that has made adjustments to its

zoning ordinance in response to PRB-1, but it's primarily a tool for an attorney should you reach that point in negotiations with your county or town.

34.2 State Regulations

State-level regulation in the United States is rare. Nonetheless, some states have also preempted certain aspects of local zoning with respect to amateur radio structures. An example is Virginia, where paragraph 15.2-2293.1 of the Code of Virginia establishes minimum heights of 200 ft and 75 ft for populous and rural counties, respectively, below which amateurs may not be unreasonably prevented by local jurisdictions from erecting supports. The VA Code does not give the amateur carte blanche as the local zoning body can still require reasonable safety and screening measures.

34.3 Local Zoning

The bulk of zoning restraints on amateur radio is local in scope. Although zoning is defined and regulated at the county level in many states, in other states it falls under the purview of each individual township or city. There are rural and sparsely populated localities with no applicable zoning ordinance. To install a tower and antennas of virtually any height (under 200 ft, of course), a building permit may or may not be required.

You may encounter layers of zoning and restrictive regulations. State or federal wilderness areas can overlay local zoning districts with other, tighter restrictions. A county may add a wrapper around a collection of local ordinances. The local ordinance may have conflicting requirements. A common problem is an overly broad ordinance for cellular and other commercial towers that fails to distinguish between the regulation of for-profit carriers and the unintended inclusion of amateurs pursuing their hobby as a permitted accessory use of residential property.

Before buying a property upon which you intend to erect towers or antennas, research the location. Know what town or municipality the property is located in and obtain a copy of the zoning ordinance and other relevant documents. Meet with a zoning official to confirm that the documents you have are the latest and represent the totality of public regulation pertaining to the property you have in mind. Archive the revision that is current as of the exact date you take legal possession of a new property. Depending on the specifics, if the town revises the zoning later on and makes it more restrictive, you may have recourse based on grandfathering provisions. Avoid depending on oral assurances that there is nothing prohibiting you from putting up a tower. Get it all in writing. Do not expect a real estate agent to be a zoning expert, building codes expert, or legal expert.

If you have located a specific property that seems appropriate for your intended radio uses but the zoning ordinance requires a variance or special use permit for erecting a tower, make the issuance of that permit a contingency in your purchase offer. Once you have a signed purchase contract, immediately set about to clear the contingency. This will most likely involve your appearing before the local planning board or zoning board of appeals to make a presentation in support of your application. Before you close on a property, be sure you have in your possession all the relevant ordinances, rulings, meeting minutes, and permits to allow you to proceed with your tower and antennas.

If your project requires you to appear in front of a town planning board or zoning board of appeals, or if you already have a tower and neighbors are complaining, you need more help than this chapter can give you. Seek additional information from additional books on zoning matters related to towers and antennas. Volunteer counsel provided through the good offices of the American Radio Relay League may be available to you and your own attorney, especially if your situation is likely to establish a legal precedent that will be useful in future litigation of antenna support cases.

34.4 Private Sector Restraints

In many regions of the world, it is becoming harder and harder to find affordable residential properties with enough land to put up a reasonable set of antennas for long-range radio communication. An increasing percentage of homes for sale are located in planned developments or homes with restrictive covenants. These may prohibit antenna towers or externally mounted antennas. The bylaws for a homeowners' association associated with a development may allow for petitioning the association for specific, individual permission to do something generally prohibited. If restrictions, economics, or other compromises preclude an antenna tower, the hidden antennas in Chap. 17 can be used.

Index

Z (impedance), 77
Π network, 498
β (propagation constant), 115
η (intrinsic impedance of space), 13
λ (wavelength), 9
ω (angular frequency), 77
c (speed of light), 11
f (frequency), 9
j (imaginary unit), 112
k_B
(Boltzmann's constant), 456
1.8 MHz (160-meter) antennas, 379
3-dB beamwidth, 98, 467

A

A-index, 51
Absorption
atmospheric, 42
ground, 28, 36
ionospheric, 53, 269, 379
Adcock antenna, 476
Admittance, 134
on Smith chart, 143
Air core, 503
Airborne radio, 33, 38
Amateur radio, 4, 179, 343, 483, 487, 530, 571, 582, 597
American Radio Relay League (ARRL), 45
Amplifier, 84, 99, 128, 195, 225, 257, 293, 336, 381, 452, 495, 519

Anchors, 575
Angle
of incidence, 24, 31, 57, 427
radiation, 57
of reflection, 24
of reflection coefficient, 152
Angular
frequency, 77, 112
velocity, 479
Antenna
160 meter, 379
Adcock, 476
analysis, 507
analyzer, 520
aperture, 443
array, 159, 449
Bandwidth limits, 102
beam, 279
beamwidth, 98
Beverage, 320
bisquare loop, 219
bobtail curtain, 265
chip, 490
construction, 533
cubical quad, 305
dipole, 185, 441
direction finding, 471
electrically small, 102
emergency, 369
EWE, 339
feed, 296
flag, 341
flat-top, 394

gain, 98
grounding, 277, 585
half square, 266
half-wave dipole, 185
halo, 411
helical, 466
hidden, 343
horn, 441
impedance, 94
impedance matching, 415
installation, 533
inverted-L, 392
J-pole, 404
limited-space, 343, 348
log-periodic, 226
longwire, 230, 257, 342
loop, 211
loopstick, 327, 335, 471
loss, 95
marine, 355, 362
masts, 547
microwave, 441
mobile, 355
modeling, 507
monopole, 176
Moxon beam, 300
multiband, 223
noise temperature, 457
patch, 489
pattern, 97
PIFA, 490
planar, 489
portable, 369

Q, 102
quad beam, 412
radio astronomy, 459
receiving, 397
reflector, 443, 470
resistance, 95
rhombic, 264
ring, 461
safety, 535, 585
small loop, 327
snake, 326
support, 374, 535
synthesis, 507
testing, 519
tower, 559
troubleshooting, 519
tuned feeder, 227
tuner, 197, 206, 209, 232, 248, 297, 342, 361, 374, 495
tuning, 333, 361
UHF, 399
vee beam, 262
vertically polarized, 233
VHF, 399
wave, 320
Windom, 231
Yagi-Uda, 279
Aperture
antenna, 443, 487
coupling, 439
effective, 100
efficiency, 446
Arcing, 305, 344, 380, 475
Area, effective, 100

Array
antenna, 159, 449
antenna, feeding, 170
broadside, 162
directional, 275
electronically steered, 452
element, 159
end-fire, 271
factor, 162
flatplate, 450
four-square, 276
longwire, 262
parasitic, 171
receiving, 171, 319
slot, 449
vertical, 269
Watson-Watt Adcock, 478
wire, 257
Wullenweber, 481
Artificial ground, 236, 259, 408, 595
Atmosphere, 21
Atmospheric noise (QRN), 53, 131, 269, 276, 320
Attenuation
atmospheric, 21, 43
constant, 115
ground wave, 29
transmission line, 114, 529
waveguide, 426
ATU, 197, 206, 209, 232, 248, 297, 342, 361, 374, 495
Auroral propagation, 70

B

Back lobe, 280, 446

Backscatter, 70
Balanced transmission line, 191, 195, 197, 500
Balun, 296, 496, 501
Band designation, 17
Bandwidth, 19, 53, 131, 282, 289, 300, 308, 333, 347, 387, 399, 456, 466, 496, 519, 520
  fractional, 201
  half-power, 78
  limitations, 102
  narrow, 361
  percentage, 187
  SWR, 193
  tradeoff, 488
  VSWR, 380
  wide, 223
Beam antenna, 279
Beamwidth
  half-power, 98
  null-to-null, 98
Bent dipole, 208
Beverage antenna, 320
Bidirectional pattern, 263, 275
Bisquare loop, 219
Bobtail curtain, 265
BOG antenna, 326
Boltzmann's constant, 456
Boundary conditions, 425
Bow tie, 203
Bracketed tower, 567
Branch feed, 452
Broadband
  dipole, 201
  impedance matching, 190

Broadside array, 162

C

Cable
    coaxial, 14, 108, 113, 196, 216, 238, 251, 287, 323
    coaxial balun, 132
    coaxial, loop antenna, 339
    coaxial, matching, 216
Cage dipole, 204
Capacitor, 75
Cardioid
    array, 168
    pattern, 169
Cassegrain feed, 448
Characteristic impedance, 105, 111
Charge, 11
Chip antenna, 490
Circuit, 73
    RLC, 77
Cloud burner, 175
Co-channel interference, 99, 167
Coax, 108
Coaxial cable, 14, 108, 113, 196, 216, 238, 251, 287, 323
    balun, 132
    loop antenna, 339
    RG series, 288
    testing, 528
Coaxial vertical, 401
Coil-loaded antenna, 205, 357
Collinear
    dipole, 165
    vertical, 406
Common-mode choke, 356

Complex impedance, 77
Conductance, 134
Conductivity, 36, 245, 325, 389, 515, 592
Cone of silence, 25
Conjugate match, 131, 152
Construction, 533
Continuous wave (CW), 125
Control points, 61, 69
Coplanar waveguide, 109
Corner reflector, 443
Corona effect, 344
Corporate feed, 452
Cosmic radiation, 23, 48
Counterpoise ground, 238, 353
Crank-up tower, 569
Critical angle, 47
Critical frequency, 52
Cryogenic, 456
Cubical quad, 305
    construction, 309
Current, 74
Cutoff frequency, 430
Cycles per second, 9

D

dB (decibels), 43, 148, 158, 530
dBd (decibels relative to half wave dipole), 177
dBi (decibels relative to isotropic), 88
dBm (decibels relative to 1 milliwatt), 101
Dead zone, 32
Dellinger fade, 55
Delta
    loop, 215, 305, 312

match, 416
Destructive interference, 29, 40, 452
Diamond loop, 314
Dielectric
constant, 112, 431, 489, 516, 592
loss, 95
resonator, 490
Diffraction, 24
Dipole, 93, 167, 176, 185, 343, 441, 511
bent, 208
bow tie, 203
broadband, 201
cage, 204
coil-loaded, 205
construction, 185
fan, 204
feedpoint, 94, 187
folded, 106, 202, 545
half-wave, 88
Hertzian, 83
horizontal, 391
impedance, 94
inverted-V, 195
modified, 195
multiband, 209
multiple, 225
off-center-fed, 230
radiation pattern, 191
resonance, 93
rotatable, 286
shortened, 205
sloping, 197
stagger-tuned, 204

trap, 223
tuning, 192
vertical, 233
Direct wave, 29
Direction finding, 327, 471
Directional array, 275
Directional discontinuity ring radiator, 462
Directivity, 270
Director, 173
Discrete (lumped-component) loading, 381
Dish antenna, 470
illumination, 445
Dispersion, 24, 114
Displacement current, 83
Diurnal variation, 54
Diversity, 63
Doppler
RDF antenna, 478
shift, 478
Double extended Zepp, 167
Double vee antenna, 264
Double-ducky direction finder, 483
Doublet, 185
Driven element, 172
Ducting, 37
Dummy load, 432, 530

E

Effective
aperture, 100, 446
area, 100
earth radius, 33
isotropic radiated power (EIRP), 53, 99

Efficiency, 95, 370
Electric
    charge, 74
    current, 74
    field, 11, 15, 78, 80, 86, 327, 424, 431
Electrical length, 94, 114, 133, 240, 259, 269, 284, 306, 361, 383, 393, 418
Electromagnetic
    interference (EMI), 109, 305
    wave, 7, 10, 17, 21, 73, 80, 92, 172, 185, 211, 335, 432
Electronically steered array, 452
Electrostatic field, 261
Element, 159
    parasitic, 172
Elevated ground, 595
Elevated ground-plane antenna, 251
Emergency antenna, 369
End insulator, 543
End-fed
    longwire, 257
    Zepp antenna, 227
End-fire array, 271
EWE antenna, 339
Extended double Zepp, 208

F

Fading, 40, 452
Fan dipole, 204
Far field, 96
Farad, 75
Faraday shield, 339
Federal Aviation Agency (FAA), 597
Federal Communications Commission (FCC), 4

Feed, 296
    illumination, 445
Feedline construction, 538
Feedpoint, 94
Ferrite, 503
    rod, 506
Ferromagnetic material, 82
Field, 79
    electric, 80
    electrostatic, 261
    magnetic, 82
    quasi-static, 11
    static, 11
    strength meter, 521
Finite difference method, 508
Finite element method, 508
Flag antenna, 341
Flat-top antenna, 394
Flatplate array, 450
Folded dipole, 202, 545
Foldover tower, 567
Forward
    scatter, 70
    wave, 149
Four-square array, 276
Fox hunting, 475
Free space, 9, 29, 80, 92, 114, 170, 186, 208, 227, 247, 284, 306, 327, 399, 423, 510, 538
    impedance, 13
    path loss, 16, 101
    wavelength, 135, 272, 490
Frequency, 9
    cutoff, 430

cycles per second, 9
dispersion, 24, 114
diversity, 63
Hertz, 9
of optimum traffic (FOT), 53
units, 9
Fresnel zone, 26
Friis transmission equation, 100

G

G5RV multiband dipole, 229
Gain, 98
Galactic noise, 49, 319
Gamma match, 249, 297, 385, 419, 490
Gigahertz (GHz), 9
Gin pole, 578
Gray line propagation, 67
Great circle path, 61
Ground, 177
artificial, 236, 259, 408, 595
elevated, 595
loss, 29, 36, 240, 294, 356, 370, 383, 394, 587
Ground systems (radials), 231, 241, 257, 277, 380, 397, 585
Ground wave, 28, 57, 63
zone, 56
Ground wire tuner, 591
Ground-mounted vertical antenna, 29, 58, 173, 233, 240, 251, 380, 387, 551, 587
Ground-plane vertical, 234, 400
Grounding, 277, 585
antenna efficiency, 592
power distribution, 590
radials, 593

rod, 246, 553, 574, 586
station, 590
Group velocity, 427
Guy wires, 384, 544, 564, 575

H

Hairpin match, 285, 296, 299, 303
Half-delta sloper, 218
Half-power beamwidth, 98
Half-sloper, 349
Half-square, 266
Half-wave dipole, 88, 185
Halo antenna, 411
Halyards, 549
Hardline, 108, 287
Harmonic, 93, 209, 218, 227, 263, 303, 391, 495
Helical
antenna, 466
line, 108
mobile whip, 353
Helium cooled receiver, 456
Henry, 76
Hertz (Hz), 9
Hertzian dipole, 83
Hidden antenna, 343
Horizon distance, 35
Horizontal
dipole, 58, 163, 173, 174, 187, 195, 229, 252, 379, 391, 501, 514
polarization, 252
Horn antenna, 441
Hula hoop antenna, 463

I

Image antenna, 174, 191, 308
Imaginary
    number, 112
    unit, 31, 77, 112
Impedance, 77, 94
    antenna, 94
    free space, 87
    matching, 131, 195, 415, 495, 502
    matching, gamma, 297
    on Smith chart, 143
    waveguide, 431
Incident wave, 149
Index
    characteristic, 105
    of refraction, 33
Inductance, 76
Inductor, 76
    ferrite, 503
Insertion loss, 158, 521
Installation, 533
Insulators, 538
Interference, 20, 29, 99, 230, 279, 342, 488, 520, 590
    adjacent channel, 167
    co-channel, 167
    constructive, 40
    destructive, 40, 452
    electromagnetic (EMI), 109, 305
    radio frequency (RFI), 475
Interferometer, 468
Intrinsic impedance, 13
Inverted-L antenna, 392
Inverted-vee dipole, 195

Ionosphere, 45
A-index, 51
critical frequency, 52
Dellinger fade, 55
disturbances, 55
frequency of optimum traffic, 53
K-index, 51
layers, 51
lowest useable frequency, 53
maximum useable frequency, 53
propagation, 45
recombination, 23
scatter propagation, 70
sporadic E propagation, 52
storm, 55
sudden disturbance, 55
Sun's effect, 48
variation, 53
virtual height, 53
Ionospheric propagation
critical angle, 47
fading, 63
gray line, 67
multihop, 57
refraction, 56
scatter, 70
skew path, 66
Isotropic source, 14

J

J-pole antenna, 404
Johnson (thermal) noise, 455
Jupiter, 460

K

K factor (effective earth radius), 33
K-index, 51
K9AY loop, 340
Kelvin (temperature), 34
Kennelly-Heaviside layer, 45
Kilohertz (kHz), 9

L

L-section, 497
Large wire loop antenna, 211
Light, speed of, 9
Lightning protection, 585
Limited-space antenna, 343
Line flattener, 364
Line of sight, 17, 269, 559
Linear loading, 296, 308, 381
Link budget, 100
Log-periodic antenna, 226
Longwire
    antenna, 230, 257, 342
    array, 262
Loop
    antenna, 211
    bisquare, 219
    delta, 305, 312
    diamond, 314
    flag, 341
    K9AY, 340
    pennant, 340
    self-resonance, 338
    transformer, 332

Loopstick antenna, 327, 335, 471
LORAN, 319
Loss, 154, 370
    antenna, 95
    ground, 29, 36, 240, 294, 356, 370, 383, 394, 587
    ohmic, 95
    transmission line, 115
Lowest usable frequency (LUF), 53

M

Magnetic
    field, 7, 11, 15, 23, 54, 67, 78, 82, 86, 100, 211, 327, 424, 431, 460
    field intensity, 82
    flux density, 82
Main lobe, 98
Marine antenna, 355, 362
Mast, 547, 580
Matching, 131, 195
    antenna tuner, 495
    conjugate, 152
    hairpin, 285, 296, 303
    network, 94, 95, 102, 131, 152, 171, 187, 195, 238, 250, 266, 285, 293, 356, 380, 415, 496, 520, 593
    quarter-wave, 135
    stub, 152
    T-match, 285, 296, 385
Maximum useable frequency (MUF), 53
Maxwell's equations, 83
Medium, 9
Megahertz (MHz), 9, 10, 17
Meteor scatter, 71
Method of moments, 509

Micromatch, 525
Microstrip, 109
Microwave, 19
    antenna, 441
    network, 156
    waveguide, 423
Mismatch loss, 129
Mobile antenna, 355
Mode, 426
Modeling antennas, 507
Monomatch, 526
Monopole, 176, 240, 380
    ground-plane vertical, 234
    vertical, 379
Monopulse feed, 448
Moxon beam, 300
MUF control points, 58
Multiband
    antenna, 223
    fan dipole, 209
    Yagi, 294
Multifilar winding, 503
Multihop propagation, 58
Multipath, 31, 40, 63, 452, 488

N

Near field, 96
Near-vertical incidence skywave antenna, 220
Neper, 115
Network, 156
Network analyzer, 521
Nitrogen-cooled receiver, 456
Nitrogen-filled cable, 108

Noise
antenna, 457
atmospheric, 131, 269, 276, 320
figure, 456
galactic, 49, 319
sky, 195
temperature, 456
thermal, 455
Nonresonant
longwire, 257
loop, 344
rhombic, 265
Vee beam, 264
Normalized impedance, 140
Null, 89
Null-to-null beamwidth, 98
Numerical
method, 508
modeling, 507
Numerical Electromagnetics Computation (NEC), 509

O

Off-center-fed dipole, 230
Ohm, 13
Ohmic loss, 88, 95, 99, 130, 187, 356, 423, 488
Ohmmeter, 95, 188, 323, 529
Omega match, 250
Omnidirectional, 166
Open-wire line, 106, 170, 190, 203, 229, 350, 373, 396, 406, 496, 542, 589
Oscilloscope, 121, 478

P

Parabolic dish, 443
Parallel
    conductor line, 106
    resonant circuit, 333
Parasitic
    array, 171
    element, 172
Patch antenna, 489
Path loss, 16
Pennant antenna, 340
Perfect conductor, 36, 174, 509
Permeability, 88
Permittivity, 88
Phase velocity, 14, 114, 427
Phase-locked receiver, 319
Phased
    array, 159, 271, 300, 339, 394, 449
    vertical, 391
Phasor, 76
Picket fencing, 32
PIFA, 490
Planar antenna, 489
Planetary terminator, 67
Polarization, 17, 252, 373
    circular, 466
    diversity, 63
    horizontal, 252
    vertical, 233, 252
Portable antenna, 369
Power
    density, 13
    meter, 523
    radio wave, 12

Preamplifier, 276, 320, 326, 413, 461
Propagation
    constant, 115
    dead zone, 32
    ducting, 37
    fading, 40
    gray line, 67
    ground wave, 28
    ionosphere, 45
    line of sight, 17, 269, 559
    modeling, 517
    multipath, 32
    refraction, 37
    space wave, 27
    surface wave, 27
    velocity, 92

Q

Q, 78
Quad
    beam antenna, 412
    loop antenna, 305
Quagi, 413
Quality factor, 78
Quarter-wave
    matching, 135
    shorted stub, 136
    vertical monopole, 176, 393

R

Radar, 5, 98, 443, 449
Radials, 593
    80-meter dipole, 397

Beverage, 322
bobtail curtain, 266
emergency and portable antenna, 370
grounded monopole, 241
half-square, 267
insulated-base vertical, 389
Radiation
angle, 57, 252, 263, 512
efficiency, 102, 370
loss, 423
pattern, 15, 59, 89, 95, 97, 162, 172, 191, 200, 211, 227, 240, 260, 275, 285, 305, 393, 452, 467, 488, 502, 514, 580, 593
resistance, 88, 95, 102, 188, 240, 319, 380, 465, 520, 592
Radio
astronomy, 459
direction finding, 471
frequency interference (RFI), 475
horizon, 33
wave, 7
Random-length longwire, 230
Ray
path, 33
tracing, 26, 517
Receiving
antenna, 16, 319, 397
array, 171, 319
loop antenna, 326
Reciprocity, 100, 131
Recombination, 23
Rectangular waveguide, 425
Reflected
wave, 89, 119, 126, 149
wave (ground), 29

Reflection, 24
coefficient, 119, 148, 515
loss, 152
loss (ground), 58
Reflector
parasitic element, 280
antenna, 443
Refraction, 24, 37
index of, 33
Regulations, 597
Relative permittivity, 81
Resistance, 13, 74
radiation, 95
Resistor, 74, 432, 530
Resonance, 78, 93, 192
Restrictive covenants, 599
Return loss, 148, 149, 158, 515, 521
RF
power measurement, 520, 523
wattmeter, 523
Rhombic, 63, 262, 264
Rigid coaxial line (hardline), 109
Ring antenna, 461
RLC circuit, 77
Rotatable dipole, 286
Rotating pole, 572
Rotators, 580
Rubber ducky VHF antennas, 483

S

S-meter, 471, 475, 476
S-parameters, 156, 521
Safety, 535, 585

lightning, 585
microwave, 441
Satellite communication, 28, 102, 446, 457, 488
Scanner antennas, 413
Scatter propagation, 70
Seasonal variation, 54
Selective fading, 63
Self-resonance, 338
Self-supporting tower, 571
Sense antenna, 344, 473
Series
feed, 248, 358, 385, 394, 419
matching section, 136
resonant circuit, 334
Shadow zone, 25
Shielded
loop antenna, 335
twin-lead, 106
Short skip, 52, 59, 181
Shorted stub, 136, 153, 300
Shortened
coil-loaded dipole, 205
vertical, 381
Shortwave, 4, 48, 61, 270, 284, 295, 319, 342, 476
Shortwave and AM BCB “skip”, 476
Shunt feed, 248, 385
Side scatter, 67, 70
Sidelobe, 98, 161
Signal strength meter, 471
Signal to noise ratio (SNR), 101
Single-sideband (SSB), 252, 363, 460
Sinusoidal
steady state, 116, 125

waveform, 13
Skew path, 66
Skin effect, 95, 112, 188, 423
Skip
distance, 47, 56
propagation, 54, 252, 460, 476
zone, 47, 56, 57, 59
Sky
noise, 195, 327, 457
wave, 57, 63
Skyhook, 414
Sloping dipole, 197
Slot
array, 449
coupling, 439
Small antennas, 488
Small loop antenna, 327
Smith chart, 138, 521
Snake antenna, 326
Software, 507
Soil, 592
conductivity, 36, 245, 325, 389, 515, 592
Solar
flare, 49
flux index (SFI), 51
Spatial diversity, 63
SPC transmatch, 501
Spectrum
access, 19
analyzers, 520
Speed of light (*c*), 11
Sporadic E propagation, 52
Sports fan's loop, 335

Spreaders, 204, 300, 309, 413
Stacking, 303
Stagger-tuned dipole, 204
    Standing wave, 89, 432
    ratio (SWR), 126, 149, 495, 519
Station grounding, 590
Stepped-diameter, 288
Strain relief, 551
Stratosphere, 22, 28
Stripline, 109
Stub matching, 133, 152, 195, 251, 418
Subrefraction, 34
Sudden ionospheric disturbance (SID), 55
Sun
    effect on ionosphere, 48
    signals from, 459
Sunspot, 49
    number (SSN), 50
Superposition, 73
Superrefraction, 37
Support, 374, 535
Surface wave, 36
Susceptance, 134
Switched-pattern antenna, 483
SWR, 519
    analyzer, 520

T

T antenna, 394
T-match, 285, 296, 298, 385
Telescoping mast, 239, 375
Temporary antenna, 369
Terminated longwire antenna, 320

Terrain, 517
Testing, 519
Thermal noise, 455
Thrust bearing, 580
Tilt angle, 36, 265
Time difference of arrival, 482
Time domain reflectometry (TDR), 121, 521
Time-harmonic, 13
Top loading, 231, 384, 388, 393
Top-loaded mobile antenna, 360
Topband (160 m), 379, 391
Toroid, 502, 503
Tower, 559
    bases, 573
    bracketed, 567
    crank-up, 569
    foldover, 567
    guyed, 564
    installation, 578
    loading, 383
    self-supporting, 571
    wind load, 560
Transformer
    loop, 332
Transient response, 116
Transmatch, 500
Transmission line, 105
    attenuation, 114, 529
    balanced, 191, 195
    characteristic impedance, 111
    coaxial, 108
    coplanar waveguide, 109
    impedance matching, 131

loss, 115
microstrip, 109
mismatch loss, 129
phase velocity, 114
quarter-wave matching, 135
reflection coefficient, 119
response, 116
sinusoidal steady state, 125
Smith chart, 138
steady-state response, 116
stripline, 109
stub matching, 133, 152
SWR, 126
transient response, 116
types, 106
unbalanced, 195, 197, 251, 284, 296, 387, 500
VSWR, 126
Transverse
electric (TE) mode, 425
electromagnetic (TEM) wave, 12, 424, 425
magnetic (TM) mode, 425
Trap dipole, 223
Traveling wave antenna, 261, 264
Triangulation, 471
Trifilar winding, 323, 503
Troposphere, 22, 37
Troubleshooting, 519
Tuned feeder antenna, 227
Tuner, 495
Tuning, 333, 361
Twin-lead, 106, 202, 209, 396, 416, 500, 545

U

Ufer ground, 574
UHF antennas, 399
Umbra region, 25
Unbalanced transmission line, 195, 197, 251, 284, 296, 387, 500

V

Varactor tuning, 334
Vector network analyzer (VNA), 121, 156, 171, 195, 323, 521
Vee beam, 262
Velocity, 92
    factor, 14, 114, 132, 149, 217, 267, 406, 516, 520
    group, 427
    phase, 427
    in waveguide, 427
Vertical
    antenna, 176, 218, 233, 247, 345, 379, 387, 406, 537, 593
    array, 172, 269
    dipole, 233
    ground-plane, 400
    phased, 391
    polarization, 252
    polarized antennas, 233
    shortened, 381
    support, 535
VHF/UHF
    antennas, 399
    scanner antennas, 413
Virtual height, 53
VLF/LF loop, 338
Voltage standing wave ratio (VSWR), 126, 190, 223, 305, 509, 519

W

Watson-Watt Adcock array, 478

Wattmeter, 523
Wave
    antenna, 320
    standing, 89
    surface, 36
Waveguide, 423
    attenuation, 426
    coplanar, 109
    cutoff frequency, 430
    impedance, 431
    microwave, 423
    mode, 426
    modes, 424
    rectangular, 425
    termination, 432
    velocity, 427
Wavelength, 9
    in transmission line, 114
Wind load, 560
Windom antenna, 231
Wire antenna
    construction, 538
    multiband and tunable, 223
Wire arrays, 257
Wire Yagi, 293
Wireless devices, 487
Wolf number, 50
Wonder Bar antenna, 204
Wullenweber array, 481

Y

Yagi-Uda antenna, 279
    construction, 285

director, 173
feeding, 296
grounds, 409
multiband, 294
Quagi, 413
reflector, 173
stacking, 301
stepped diameter, 288
wire, 293

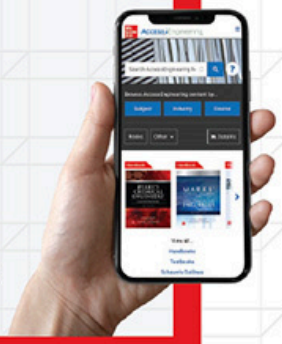
Z

Zepp antenna, 208, 227
ZL-special beam antenna, 267
Zoning and regulatory, 559, 597

ACCESS Engineering

AccessEngineering® from McGraw Hill is an award-winning engineering reference and teaching platform that delivers world-renowned, interdisciplinary engineering content integrated with analytical teaching and learning tools.

AccessEngineering prepares students to solve real-world problems, makes curriculum planning and delivery easy for faculty, and helps professionals find relevant information faster.



With AccessEngineering, you can:

- Search the latest editions of renowned engineering handbooks and hundreds of other expert references
- Access leading upper-level engineering textbooks
- Better understand material properties with **DataVis,** the interactive data visualization tool
- Accurately solve complex engineering equations with downloadable Excel spreadsheets
- Watch engineering solutions in action with hundreds of faculty-created videos
- Analyze key data with thousands of interactive graphs and downloadable tables
- Review comprehensive step-by-step explanations of engineering problems using solutions walkthroughs
- Find answers quickly with powerful, faceted search capabilities
- Develop entrepreneurial business skills to capitalize on innovations
- Browse content by subject, industry, course, or codes & standards to find the information you need
- Organize project information with personalization tools

Check with your institution's library or administrator to see if you have access to AccessEngineering.

Visit AccessEngineering at accessengineeringlibrary.com

Mc Graw Hill