Next-Generation Extreme Ultraviolet Lithographic Projection Systems

Next-Generation Extreme Ultraviolet Lithographic Projection Systems

Proefschrift

ter verkrijging van de graad van doctor

aan de Technische Universiteit Delft,

op gezag van de Rector Magnificus prof. dr. ir. J.T. Fokkema,

voorzitter van het College voor Promoties,

in het openbaar te verdedigen op maandag 10 februari 2003 om 16:00 uur

door Matthieu Frédéric BAL

natuurkundig ingenieur

geboren te Utrecht

Dit proefschrift is goedgekeurd door de promotor:
Prof.dr.ir. J.J.M. Braat.

Samenstelling promotiecommissie:

Rector Magnificus	voorzitter
Prof. dr. ir. J.J.M. Braat	Technische Universiteit Delft, promotor
Prof. dr. L.H.J.F. Beckmann	Universiteit Twente
Prof. dr. I. Livshits	University of St. Petersburg, Rusland
Prof. dr. D.L. Shealy	University of Alabama, VS
Prof. dr. H.P. Urbach	Technische Universiteit Delft
Dr. F. Bociort	Technische Universiteit Delft
Dr. H.-J. Mann	Carl Zeiss, Duitsland

ISBN	90-9016530-4
NUR	950 Technische wetenschappen algemeen

Abstract

Promising successors of the current transmissive deep ultraviolet lithographic systems are the reflective extreme ultraviolet lithographic systems, which image feature sizes below 50 nm. The small feature sizes demanded and the restriction on the number of reflectors necessitates high-order aspherical surfaces to enlarge the set of parameters describing the projection system. In the paraxial approximation, we can impose several conditions to the projection system and drastically decrease the number of variables. The set of parameters can be imagined to define a vector in a solution space. In this solution space, the prohibition of obstruction limits systems to small domains. The paraxial obstruction-free domains appear to include the corresponding finite obstruction-free domains. This suggests that all obstruction-free domains of interest can be found with methods based on the paraxial approximation. We perform exhaustive searches in the paraxial approximation with the absence of obstruction as the principal criterion. The resulting unobstructed systems are starting points for further optimisation.

A new classification method of reflective systems is based on the relative arrangement of mirrors and on the ray paths through the system. We present the systems found with the exhaustive searches for projection systems with four, six and eight reflectors. The classes with four- and six-mirror systems are mostly known from the patent literature. An exception with a relatively large numerical aperture and good optical performance is a four-mirror system in class 6-. Some systems show surprising mirror arrangements, although other factors than the absence of obstruction make some of these classes less attractive. Some of the eight-mirror systems with good optical performance are new and promising.

A typical multilayer consists of a hundred alternating layers of molybdenum and silicon. The thickness of these multilayers is approximately twenty-five wavelengths. The average reflection occurs at approximately four wavelengths below the top of the multilayer. The multilayers entail important consequences for the imaging properties, such as resolution, depth of focus and tolerances. Principally the average or effective reflection depth determines the imaging properties of the multilayer, e.g. the average lateral displacement and average path length change of the optical rays on reflection.

In optical design programs, phase changes that occur on reflection due to the presence of multilayers are summed at the exit pupil and combined with the optical path length of the ray, while reflectance changes are accounted for by an intensity multiplication factor. These phase and intensity variations are derived from thin film calculations, which suppose a coherent plane wavefront incident on a plane multilayer and sum the reflected fields into a single outgoing field. Another similar condition is that all rays at each position on the reflecting multilayer must have a unique direction. Otherwise, the optical performance drastically deteriorates. We use this latter condition to calculate the spatially varying optimum thickness (*grading*) of a multilayer.

Samenvatting

De reflectieve extreem-ultraviolet lithografiesystemen zijn veelbelovende opvolgers van de huidige refractieve diep-ultraviolet systemen en kunnen structuren kleiner dan 50 nm afbeelden. De gevraagde resolutie en de beperking van het aantal spiegels maakt het gebruik van aspherische oppervlakken noodzakelijk om het aantal correctie-variabelen in een projectiesysteem te vergroten. In de paraxiale benadering kunnen we meerdere condities opleggen aan het projectiesysteem en het aantal variabelen wordt daardoor drastisch verminderd. Een reeks van variabelen kan beschouwd worden als een vector in een oplossingsruimte. In deze oplossingsruimte beperkt het uitsluiten van bundelobstructie de existentiegebieden van de systemen tot kleine domeinen. De paraxiale domeinen zonder obstructie blijken de domeinen zonder obstructie berekend met eindige stralen te omvatten. Dit suggereert, dat alle interessante obstructie-vrije domeinen met de paraxiale benadering gevonden kunnen worden. Met de paraxiale benadering doen we uitgebreide zoektochten met de afwezigheid van obstructie als belangrijkste criterium. De resulterende systemen geven beginwaarden voor systemen die daarna geoptimaliseerd worden.

Een nieuwe onderverdeling van spiegelende systemen is gebaseerd op de relatieve positie van de spiegels en op de stralengang. Wij tonen de gevonden groepen van systemen met vier, zes en acht spiegels. De meeste groepen met vier en zes spiegels zijn bekend uit de patentpublicaties. Sommige systemen hebben verrassende spiegelopstellingen, echter andere eisen dan de afwezigheid van obstructie maken deze groepen soms minder interessant. Sommige nieuwe groepen van acht-spiegel systemen hebben goede afbeeldingskwaliteiten en zijn veelbelovend.

Het gangbare reflecterende oppervlak in deze systemen bestaat uit ongeveer honderd lagen van afwisselend molybdeen en silicium aangebracht op een asferisch substraat. Deze opstapeling is ongeveer vijfentwintig golflengten dik, waarin de gemiddelde reflectie ongeveer vier golflengten onder het oppervlakte plaatsvindt. Deze lagen hebben belangrijke gevolgen voor de afbeeldingseigenschappen, zoals de resolutie, de focusdiepte en de systeemtoleranties. De gemiddelde of effectieve reflectiediepte bepaalt voornamelijk de afbeeldingseigenschappen, zoals de gemiddelde zijdelingse verschuiving en de gemiddelde weglengtetoename van een lichtstraal bij een reflectie.

In optische ontwerpprogramma's worden de faseveranderingen, die ontstaan door de aanwezigheid van de lagen, in de uittreepupil opgeteld bij de optische weglengte van elke straal, terwijl de reflectiecoefficienten worden vermenigvuldigd. Deze fase- en intensiteitsveranderingen worden verkregen met dunne lagen berekeningen, die een coherente vlakke golf veronderstellen die aan een vlak oppervlak spiegelt. Een soortgelijke voorwaarde is dat alle gereflecteerde stralen op een bepaalde plek op een oppervlak allemaal dezelfde richting en fase hebben; anders verslechtert de afbeeldingskwaliteit van het systeem sterk. Met deze laatste voorwaarde berekenen we de plaatsafhankelijke optimale dikte van de lagen.

Contents

1 Introduction

Lithography is the art of printing from a flat limestone or metal plate by a method based on the attraction of grease and the repulsion of water by grease. A design or image is drawn on the surface with a greasy material, for instance using a crayon or an inkpencil. When the stone is made wet with a sponge and a roller with ink is passed over it, the ink is deposited on the greasy drawing but not on the wet stone. Alois Senefelder invented the art at the very end of the 18^{th} century in Austria. By the middle of the 19^{th} century the process had been refined to the point that it was possible to print 10.000 prints per hour.

The literal translation of photo-litho-graphy from Greek is light-stone-writing. Photolithography is similar to lithographic printing. A layer of photoresist is exposed with the image of an object pattern, is developed and, as a final result after other processing steps, we obtain a pattern of a specific deposited material. Ultimately, a large collection of layers of different materials on a substrate forms an integrated circuit. An integrated circuit is an assembly of electronic components. These components can be either active, for instance transistors and diodes, or passive, for instance capacitors and resistors.

The tendency is to decrease the feature sizes of the patterns on the integrated circuits by, among others, decreasing the wavelength of the illumination. The current systems use wavelengths in the deep ultraviolet with wavelengths of 248 nm, 193 nm, and in the very near future, 157 nm. One of the most-promising next-generation lithographic techniques is extreme ultraviolet lithography, which should be capable of imaging patterns with feature sizes of 50 nm and smaller.

Structure of the thesis

In this thesis, we focus on the design of projection systems that can be utilized in an extreme ultraviolet lithographic tool. The second chapter introduces lithography, the current and the next generation of lithographic systems with special emphasis on the extreme ultraviolet lithographic systems. The different components of an extreme ultraviolet lithographic system are discussed, in particular the requirements the imaging system has to fulfill.

The modeling of optical systems is the subject of the third chapter. This chapter includes an introduction to paraxial and finite ray tracing. The representation of optical surfaces and the aberrations in the exit pupil are discussed. Then, classification methods of reflective imaging systems are treated.

In the first stage of designing a projection system, we use the paraxial approximation. In the paraxial model of an imaging system, most requirements for the final system can already be imposed, leading to a further decrease of the number of parameters. We discuss this approach in the fourth chapter. The very restrictive demand that the system is unobstructed is used in exhaustive searches to find initial configurations that are later optimized.

The fifth chapter introduces different local and global optimization algorithms. We discuss in particular the optimization of extreme ultraviolet projection systems and their typical difficulties.

The sixth chapter presents new four-, six- and eight-mirror systems. Also, a new system to be used in a maskless extreme ultraviolet system is shown. All these systems were found with the exhaustive search method.

The multilayers and their implications for the imaging system are discussed in the seventh chapter. First an introduction to thin-layer calculations is given, after which the concept of an effective reflection depth is discussed and methods are presented to calculate the local optimum thickness (*grading*) of multilayers.

A discussion of the subjects dealt with in this thesis is given in chapter eight accompanied by some conclusions.

2 Lithography

Silicon is a natural semiconductor. By exposing the semiconductor to chemicals or by bombarding it with charged atoms some of the original atoms are replaced by impurities. This process, called doping, modifies the electrical conductivity of the material and enables the construction of miniature electronic circuits. The selection of the areas to be treated is done with the aid of a process called lithography. In this thesis, we will focus on a particular method called photolithography. Silicon is inexpensive and abundant; therefore, almost all of today's computer chips are built on a thin slice of silicon called the wafer. Recently, certain plastic materials also receive attention, since they show conductive and light-emitting properties and can be used to manufacture pliable plastic chips, flexible displays, and thin-film batteries. A great variety of integrated circuits is produced with silicon, for instance Dynamic Random Access Memories (DRAM), MicroProcessor Units (MPU), Micro Electro-Mechanical Systems (MEMS) and Application-Specific Integrated Circuits (ASIC). All these integrated circuits consist of layers with intricate patterns of different materials.

2.1 Photolithography

Wafer steppers and / or scanners are crucial to manufacture integrated circuits. The geometric features of a layer are copied by a photolithographic system from a mask to a wafer coated with photoresist, see Fig. (2.1). The process to make the integrated circuits starts with the production of the wafer. Polycrystalline silicon is melted and grown into silicon crystals. Thin wafers are sliced from the silicon and polished.

Photolithography comprises the following process steps[1]. In the first step, a whole layer of a material is deposited on the wafer. Subsequently the wafer is coated with a light-sensitive chemical layer called the photoresist. A pattern on a plate, named the mask or reticle, is imaged with a typical absolute reduction of 4:1 or 5:1 by the projection optics to expose the photoresist. The exposure to light changes the chemical properties of the photoresist. The development and baking process washes away the exposed regions of the photoresist; other treatments like etching and implantation modify the material at the locations where the photoresist has disappeared. The remaining resist is removed and a final bake is often used to drive out volatile organic materials and water. These processes of photolithography are repeated in succession to create an assembly of layers with patterns of different materials on the wafer. The integrated circuits on the wafer are tested after the deposition of all layers and sawn from the wafer. The individual integrated circuits are put into protective packages, which include connections.

Three optical components can be distinguished within a photolithographic system:

1. the condensor system, to uniformly illuminate the mask,
2. the projection system, to image the mask onto the photoresist on the wafer,
3. the alignment optics, to precisely position the various components like wafer, mask, projection lenses.

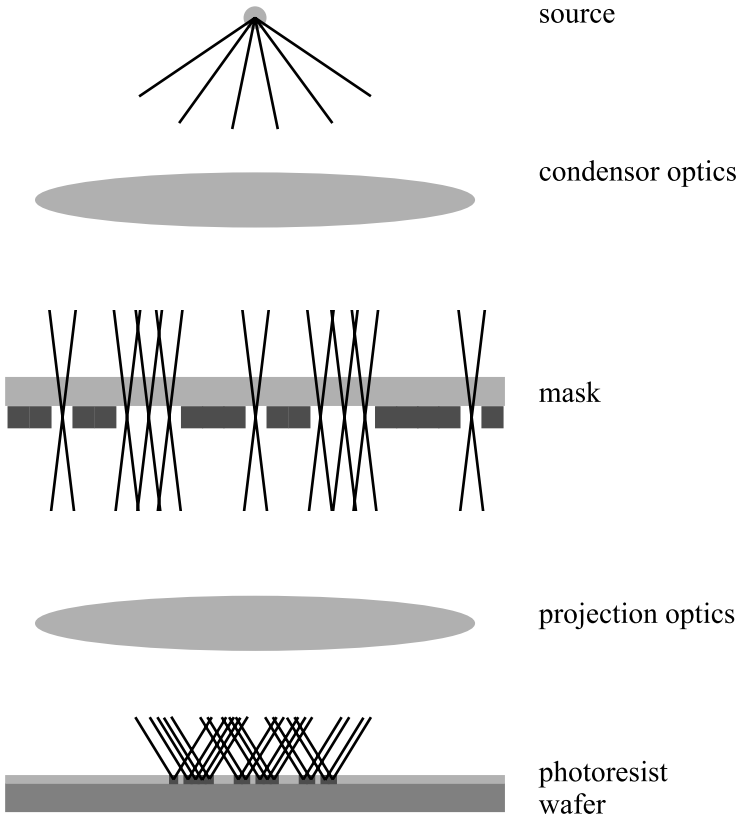

Fig. (2.1) The different components in a photolithographic system are schematically shown. The light travels from the source to the photoresist. Both the mask and wafer are scanned to produce a multitude of identical images on the wafer that are defined by the mask. Note that diffraction effects heavily determine the image formation, so that the ray picture according to the Figure is not complete.

2.1.1 The development of photolithography

In the late 1960's, the emerging technique of optical lithography was based on both contact and projection printing. Integrated circuits with feature sizes of 5 μm were produced. Ever since, optical lithography has been the prime volume manufacturing method applied in the micro-electronics industry. The urgency to produce integrated circuits with decreasing feature sizes forced the industry to rapidly change over to new lithography technologies.

The very first optical projection systems were simple 16 mm movie camera lenses. Nowadays, the lenses weigh nearly a ton and the machines cost several millions of dollars. In 1965, former Intel executive Gordon Moore stated his well-known expectation that the number of transistors on a computer or memory chip doubles every three years. A prediction updated to every two years in 1995. Originally intended as a rule of thumb, his

observation became known as Moore's Law and proves to be a consistent trend indicator. The number of transistors or memory elements on a wafer primarily increases on account of smaller feature sizes and improved product and process design. Moore also observed that the performance in terms of clock frequency of the MicroProcessor Unit (MPU) doubles every one and a half to two years. The clock frequency is expressed in Millions of Instructions Per Second (MIPS), also abbreviated as MHz. The improvements and miniaturization of integrated circuits also reduces the power consumption and extends battery life in portable applications like laptops or cell-phones by decreasing the leakage current. The increasing costs of new techniques are justified for high-volume products like memories and microprocessors, since the costs are spread over millions of produced parts.

Lithography roadmap

The International Technology Roadmap for Semiconductors (ITRS) publishes the objectives of the semiconductor industry in the next fifteen years[2]. It is a cooperative effort of global industry manufacturers and suppliers, government organizations, consortia, and universities and is organized by International Sematech. The SEmiconductor MAnufacturing TECHnology (Sematech) started as a research and development consortium of American companies in the semiconductor industry. From this consortium, the International Sematech arose with the aim to accelerate the development of semiconductor manufacturing processes, materials, and equipment.

Table (2.1) The improvement of the performance of Dynamic Random Access Memory (DRAM) and MicroProcessor Units (MPU) illustrated by means of typical dimensions. The pitch of DRAM is often referred to as a technology node.

Year of production	2001	2002	2003	2004	2005	2006	2007	2010	2013	2016
DRAM half pitch (nm)	130	115	100	90	80	70	65	45	32	22
MPU half pitch (nm)	150	130	107	90	80	70	65	45	32	22
MPU printed gate length (nm)	90	75	65	53	45	40	35	25	18	13
MPU physical gate length (nm)	65	53	45	37	32	28	25	18	13	9

Historically, the introduction of a new Dynamic Random Access Memory (DRAM) generation marks a *technology node*. For this type of integrated circuits, primarily the width of the smallest line (called pitch) indicates the capacity of integrated circuits. In logic, the length of the transistor gate limits the speed of the transistor and has to be controlled precisely to avoid electrical breakdown. The shape of the sidewall profile of the transistor gate is important, to acquire an acceptable sidewall oxide coverage. Another challenge is stopping the etch process at a very thin gate dielectric, without cutting a trench into the underlying silicon. Logic technology, for instance microprocessor units, developed slower than DRAM technology. Nowadays, the performance and costs of logic technology are quickly improving and catching up with DRAM technology. Table (2.1) shows the predicted improvements of both DRAM and MicroProcessor Units in recent and coming years.

Often the end of photolithography is prophesized, but optical methods have always been extended beyond predicted limits until now. The lithographic roadmap consists of only photolithographic methods until the end of 2010, see Table (2.2). If these predictions work out, the DRAM chips can store approximately 200 billion bits in 2014. That is roughly 23 gigabytes, nowadays the size of an average hard disk. The microprocessors will also be a hundred to a thousand times more powerful.

Table (2.2) The lithography roadmap shows the expected lithographic method to enable volume production at certain technology nodes[3].

Year	Node	Lithographic technique
1981	2000 nm	i/g-line steppers
1984	1500 nm	i/g-line steppers
1987	1000 nm	i/g-line steppers
1990	800 nm	i/g-line steppers
1993	500 nm	i/g-line steppers
1995	350 nm	i-line → 248 nm DUV
1997	250 nm	248 nm DUV
1999	180 nm	248 nm DUV
2001	130 nm	248 nm DUV
2003	90 nm	193 nm DUV
2005	65 nm	193 nm → 157 nm
2007	45 nm	157 nm → EUV
2009	32 nm and below	EUV

Critical dimension

The minimal or critical dimension (CD) of the patterns imaged on the photoresist depends on the imaging characteristics of the projection system. In a diffraction-limited optical system with a circular aperture and coherent illumination, the image of a single point is named the Point Spread Function (PSF). The point spread function of an optical system with circular aperture results in the Airy disc in the focal region. The distance between the peak of the Airy distribution and its first zero is defined as the Rayleigh resolution or the resolved distance δ_R

$$\delta_R = 0.61 \frac{\lambda}{NA}, \tag{2.1}$$

where λ is the wavelength and NA the numerical aperture at the image side of the optical system. In lithography, the feature size δ is defined as

$$\delta = k_1 \frac{\lambda}{NA}, \tag{2.2}$$

where the parameter k_1 is a process-related factor and accounts for factors such as camera performance, resist contrast, etch characteristic, and the use of reticle enhancement technologies. A k_1 of 0.61 can be found in Fig. (2.2 a) to be the diffraction-limited resolution of an optical system for the imaging of two point objects. The diffraction determined depth of focus is defined as the distance along the optical axis between the peak and the first minimum of the Airy distribution. In a good approximation, this depth of focus χ_D is written as (see Fig. (2.2 b))

$$\chi_D = \frac{\lambda}{2(NA)^2}. \tag{2.3}$$

In lithography, an effective depth of focus χ is defined according to

$$\chi = k_2 \frac{\lambda}{(NA)^2}, \tag{2.4}$$

where k_2 is another process-related factor. The depth of focus is preferably larger than 1 μm, limiting the resolution improvement achievable by enlarging the numerical aperture. In practice, the acceptable values for k_1 and k_2 are determined experimentally and are those values which yield the desired control of critical dimensions within a tolerable process window.

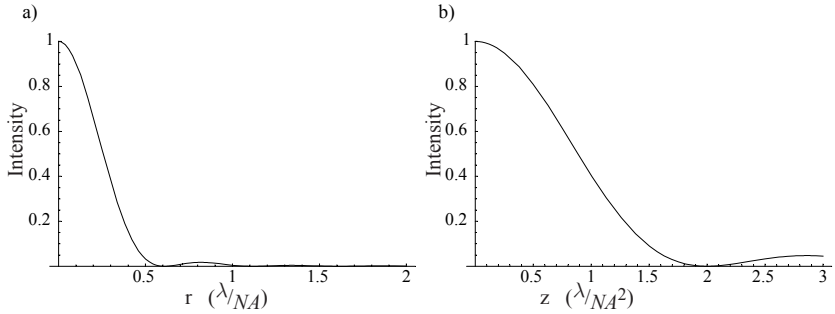

Fig. (2.2) The two graphs show the cross-sections of the Airy distribution in the focal region. The left graph is a plot of the intensity as a function of the lateral distance r from the optical axis. This radial coordinate is expressed in units λ/NA. The right graph is a plot of the intensity as a function of the defocus z along the optical axis. The axial coordinate is expressed in units $\lambda/(NA)^2$.

The expression for the feature size in Eq. (2.2) reveals three different methods to decrease it:

1. an increase of the numerical aperture. Current refractive systems have numerical apertures as large as 0.8, which could even be increased by immerging the focus region at the image in a liquid with a refractive index larger than one.
2. a decrease of the k_1 factor. The k_1 can be decreased by several methods, including the decrease of the aberrations, the use of optical proximity correction patterns, mixed phase-amplitude structures in the reticle, the illumination properties of the condensor, and resist properties. Present etch process techniques result in a gate dimension that is smaller than the illuminated region of the resist. These techniques tend to be expensive; consequently other resolution improvement methods are often preferred and employed.
3. a decrease of the wavelength.

The finite limit of the maximum numerical aperture and minimum k_1 factor achievable in optical systems constrain the maximum possible resolution at a particular wavelength[4], see Fig. (2.3).

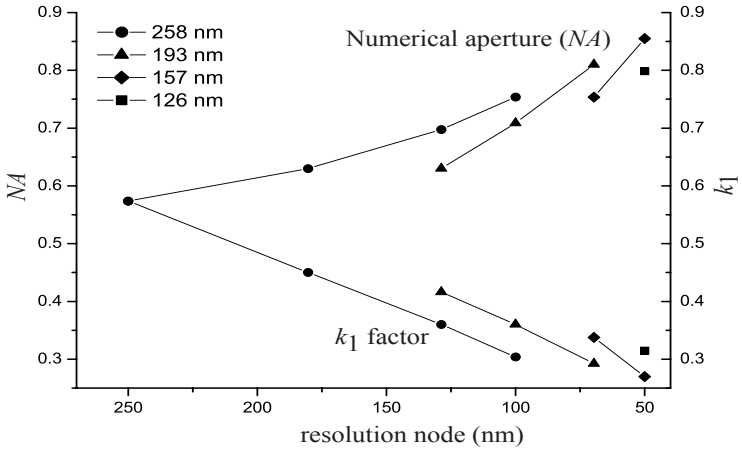

Fig. (2.3) The resolution can be improved by increasing the numerical aperture (the upper half of the plot) or decreasing the value of the k_1 factor (the lower half of the plot). The different line sections use different wavelengths, in a range from 248 nm down to 126 nm.

2.1.2 Current and near-future lithographic systems

Currently, Deep UltraViolet (DUV) wavelengths are used in photolithographic systems. The lithographic systems working with a 248 nm wavelength have a mature infrastructure and were utilized as from the 250 nm technology node. The first deep ultraviolet systems had a numerical aperture of 0.50. A next generation of deep ultraviolet lithographic systems has a numerical aperture larger than 0.80 and is due in 2003.

The coming generation of lithographic systems will work at the shorter wavelength of 193 nm. The earlier lithographic systems with a 193 nm wavelength will have a numerical aperture of 0.75 to enable the 90 nm technology node. The introduction into high-volume manufacturing should be in the second half of 2002. In early 2003, deep ultraviolet lithographic systems working with a wavelength of 193 nm and a numerical aperture of 0.85 could follow. These systems will probably be able to work at the 65 nm technology node, with phase shift masks. For mask cost reasons, the 157 nm wavelength will probably be preferred to image features with a size of 65 nm.

The lithographic systems with a 157 nm deep ultraviolet wavelength are now still in development. The introduction is planned around 2004. Typically, the projection systems working with 248 nm and 193 nm are all-refractive systems. The systems working with 157 nm on the other hand are mostly catadioptric, that is a combination of reflective and refractive. The difficulties or challenges of the 157 nm DUV systems include the volume supply of large diameter, high-purity CaF_2 crystal material needed for the lenses and the correction of the birefringence in the optical system. The strong absorption of short wavelengths limits the materials practicable in deep ultraviolet lithography.

At a wavelength of 193 nm, the base material is quartz but the optical projection system needs some elements made of CaF_2 to correct the chromatic aberrations. Systems with a

wavelength of 157 nm need entirely crystalline refractive optics. The incorporation of crystalline optics gives rise to a second order wavevector-dependent birefringence and results in unexpected difficulties[5]. The so-called spatially induced birefringence of cubic CaF_2 is proportional to λ^{-2}. Consequently, the induced birefringence and the difficulty to correct the birefringence increase drastically in future systems with a wavelength of 157 nm.

2.1.3 Next-Generation Lithography (NGL)

The minimum feature size achievable with optical lithography can not be significantly smaller than the wavelength of the illumination. The nodes, minimal dimensions or feature sizes that can be printed with the current deep ultraviolet lithographic systems will reach their boundaries or optical limits in the near future. A number of alternatives and extensions to optical lithography were conceived and are currently the subject of research or transferred to the development stage. At various times, the next-generation lithography has included proximity X-ray, ion projection, extreme ultraviolet, and electron projection lithography. The International Sematech nowadays concentrates on extreme ultraviolet and on electron projection lithography. These two techniques are considered to be the most promising, although every technique has its supporters and skeptics. The following subsections briefly describe these potential successors of current lithographic methods.

Electron Beam Direct Write (EBDW)

Direct write electron lithography is already well established. In electron-based lithography, electrons, instead of photons, illuminate a wafer. The wavelength of these electrons is many factors smaller than the achievable feature sizes. The main disadvantage is the low throughput: an electron beam only writes a pixel at a time. Wafer throughput with e-beam lithography is too slow for use in large-scale semiconductor wafer production. The lithographic technique is for instance used for mask writing (e.g. for X-ray lithography), device prototyping, and low-volume productions.

Electron Projection Lithography (EPL)

To circumvent the main disadvantage of the low throughput of electron-beam direct-write lithography, electron projection lithography proposes to use masks. Electron projection lithography is very analogous to conventional photolithography, with an electron source, a transmissive mask, and a reduction lens.

The absorption of the illumination results in heating. One type of masks consists of a thin layer of low-atomic-number material. The non-transparent regions have a high-atomic-number material that scatters the electrons to a larger extent. The scattering at mainly the low-atomic-number material broadens the energy distribution of the electrons at the wafer, leading to unsharpness. An alternative is the stencil mask that has completely transparent regions. A complication of stencil masks is that designs with isolated regions are problematic. For instance, a donut shape demands at least two exposures.

The Coulomb forces of the electrical field of an electron affect other electrons, especially in the back focal region of the projection system, where they are very close together. This

mutual repulsion limits the resolution, the maximum intensity of the beam, and the maximum throughput in wafers per hour.

Several groups are working on electron projection lithography:

- The SCALPEL system (Scattering with Angular Limitation Projection Electron-Beam Lithography) was initiated at Lucent's Bell Laboratories. In 2001 Applied Materials and ASML dissolved eLith LLC, a joint venture formed to commercialize the SCALPEL technology[6].
- The IBM and Nikon companies work on electron projection lithography in a project named Projection Exposure with Variable Axis Immersion Lenses (PREVAIL)[7]. IBM developed the accompanying magnetic lens with a 4:1 reduction that they call Curvilinear Variable Axis Lens (CVAL).
- A Japanese venture capital company works on Low Energy Electron Proximity Lithography (LEEPL)[8, 9]. The electrons have an energy of 2 keV, lower than the 10 keV used in other electron-based lithography techniques.

Multi-Aperture Pixel-by-Pixel Enhancement of Resolution (MAPPER)

MAPPER Lithography aims to combine optical lithography and electron beam lithography[10, 11]. The first part of the proposed lithographic tool is based on present deep ultraviolet steppers, with optionally a mirror array that replaces the mask. The difference is that the mask is illuminated by a microlens array, which is imaged on a converter plate. Each light beam from the microlens array triggers the emission of electrons by one of the photocathodes on the converter plate. Each element of the array of photocathodes is imaged 1:1 with a magnetic field and a small accelerating electrostatic lens on the wafer.

Proximity X-ray Lithography (PXL)

X-rays have a wavelength in the order of 0.01-100 nm and therefore the ability to define very high resolution images. In proximity X-ray lithography, the mask is brought very close to the wafer[12]. X-rays then illuminate the mask wafer combination. The mask must be as transparent as possible. Very thin membranes of low-atomic-number materials are used. The absorption of X-rays is proportional to the atomic mass of the material. The patterns on the mask consist of highly absorbing or scattering high-atomic-number materials. The 1:1 reproduction of the mask on the wafer puts severe restrictions on the geometric deviations of the mask. In practice, the requirements turn out to be too severe. Therefore, this lithographic technique is not considered a serious candidate for next-generation lithography anymore, but it could be used in niche and military applications.

Extreme UltraViolet (EUV)

Extreme ultraviolet lithography uses radiation with a wavelength of 10 to 14 nm, which is in the soft X-ray region[13]. Virtually all materials, even gases, absorb the extreme ultraviolet wavelengths. The near-normal incidence reflectivity of individual materials is also low. Only at grazing angles of incidence, a usable reflectivity is obtained. However, a system with only glancing incidence angles has a limited available aperture and resolution. A breakthrough was the development of multilayers in this wavelength region: thin-film

coatings known as distributed Bragg reflectors, with a 70% maximum reflectivity at the extreme ultraviolet wavelength region.

In 1997, Intel, AMD, and Motorola established the Extreme UltraViolet Limited Liability Company (EUV LLC) to develop Extreme UltraViolet Lithography (EUVL)[14]. They funded a prototype machine, called the Engineering Test Stand (ETS), jointly built by Lawrence Berkeley National Laboratory (LBNL), Lawrence Livermore National Laboratory (LLNL), and Sandia National Laboratories (SNL). These three laboratories form the Virtual National Laboratory (VNL). Completed in April 2001, the engineering test stand is used to test and refine the technology. The engineering test stand is a complete four-mirror stepper capable to image 70 nm features. The projection system in the engineering test stand has a numerical aperture (NA) of 0.10, a 4:1 reduction, a k_1 and a k_2 of 0.52, a wavelength of 13.4 nm, and a depth of focus of 0.7 μm.

In Europe, ASML, Carl Zeiss, Oxford Instruments, and their partners started the Extreme Ultraviolet Concept Lithography Development Systems (EUCLIDES) program[15-18]. They now design and realize an exposure tool called the alpha tool, which should be completed in 2003[19]. The optical system will comprise a six-mirror Projection Optics (PO) system. Carl Zeiss and ASML also create and test a two-mirror high numerical aperture projection system capable of lithography with feature sizes down to 30 nm and below. The extreme ultraviolet beta tool, which is scheduled for delivery in 2005, will be used to create components with 45 nm feature sizes. Intel ordered this beta tool from ASML. A gamma tool, that is a production and commercial version of the machine, is expected to be available in 2007.

In Japan, Canon Inc., Nikon Corp., and other Japanese companies formed the Association of Super-Advanced Electronics Technologies (ASET)[20]. They completed the development of a new alpha prototype high-numerical aperture tool based on extreme ultraviolet technology and gear for 35-nm wafer processing[21]. A new program is started to develop blank photomasks, metrology systems, resists technologies, and laser sources for extreme ultraviolet lithography. Canon and Nikon are also separately developing their own extreme ultraviolet system and collectively developing the metrology tools for extreme ultraviolet lithography.

2.2 Extreme ultraviolet photolithography

In an extreme ultraviolet lithographic system, the same major subsystems can be distinguished as in a conventional photolithographic system, see Fig. (2.1). A light source with condensor optics illuminates the mask. The mask holds a magnified image of the circuit and is projected with a reduction onto the wafer. The following subsections discuss these subsystems. The differences with traditional optical lithographic systems include the vacuum condition for the whole system and the all-reflective optics and mask. The vacuum is needed to avoid absorption and scattering of the radiation and degradation of the multilayers and mask.

2.2.1 Source

Ultimately, the output of an extreme ultraviolet source should be in the range of 50 to 150 W in order to produce the desired throughput of eighty wafers per hour[22]. The intensity

must be that large to allow for the substantial losses in the optical path. At the wafer, an extreme ultraviolet system with a typical number of nine reflectors has a relative intensity of at most only $0.7^9 \approx 0.04$. Other important characteristics are the pulse-to-pulse repeatability, the spatial stability control, and the repetition rates. The spectral content and the spatial coherence of the source need to be well known due to, among others, the use of multilayers and the sensitivity of the resist to other wavelengths than the design wavelength. Different sources were proposed and investigated, the most promising candidates include:

- laser-produced plasma (LPP) sources and discharge lamps sources. A hot plasma of gases as lithium, xenon, and oxygen releases its energy in the extreme ultraviolet portion of the spectrum. If the plasma is small, it can be considered as an isotropically radiating point source. The gas is excited either with a laser, in the laser-produced plasma source, or with a high voltage and large currents, in the discharge lamp. In the engineering test stand (ETS), a pulsed solid-state laser beam (YAG) focuses onto a xenon cluster target[23, 24].
- synchrotron source. Although their size and installation requirements are quite considerable, synchrotrons do have satisfactory lifetimes, enough output power, and do not contribute to wafer contamination.

2.2.2 Illuminator

The main task of the illuminator is an efficient coupling of the light emitted by the extreme ultraviolet source to the projection optics interface in the reticle plane. The design of the collector mirror depends on the source used. The illuminator should give a homogeneous illumination of the useful arc-shaped field in the reticle plane and a uniform fill of the entrance pupil of the projection optics. The illumination beam and its reflection from the reticle should be as perpendicular to the reticle surface as possible to minimize the effects of misplacements of the reticle. We return to this telecentricity constraint in the subsection on **telecentricity** (page 26).

2.2.3 Mask

In contrast with masks used in conventional photolithography, the mask in extreme ultraviolet lithography must be reflective. The extreme ultraviolet masks are coated with a similar multilayer as on the reflectors in the projection system. An additional absorbing layer patterns the desired integrated circuit features.

An alternative for a mask is an array of very small mirrors (10^5-10^8 mirrors). The micromirrors modulate the light direction, as they rotate around one or two axis when an electrostatic force is applied[25]. The disappearance of the need for masks offers inherent cost and cycle-time reduction. A problem is the huge amount of data, strongly increasing with the resolution, which must be transmitted to the mirror arrays to achieve an acceptable throughput. The throughput also demands a very high frame rate (e.g. several MHz) of the source, which probably precludes the use of a pulsed extreme ultraviolet source such as the laser-produced plasma source. A fabrication methodology of mirror arrays for Extreme UltraViolet (EUV) maskless lithography is discussed in the literature[26]. We present some projection systems for maskless extreme ultraviolet lithography in sections 6.4.

2.2.4 Projection system

In extreme ultraviolet projection optics, the mirrors, object, and image share an axis of symmetry; they are coaxial. This simplifies the system assembly, alignment, and reduces the number of possible aberrations. The use of mirrors can lead to obstruction of the beam, see the subsection on **obstruction** (page 25). To enable an undisturbed propagation of the beam and a large illuminated area, the systems have a ring-shaped field. The wavefronts converging to the field points within these narrow annular strips are highly corrected and should be virtually perfectly spherical up to a fraction of the wavelength of the radiation. The shape of the ringfield is shown in Fig. (2.4).

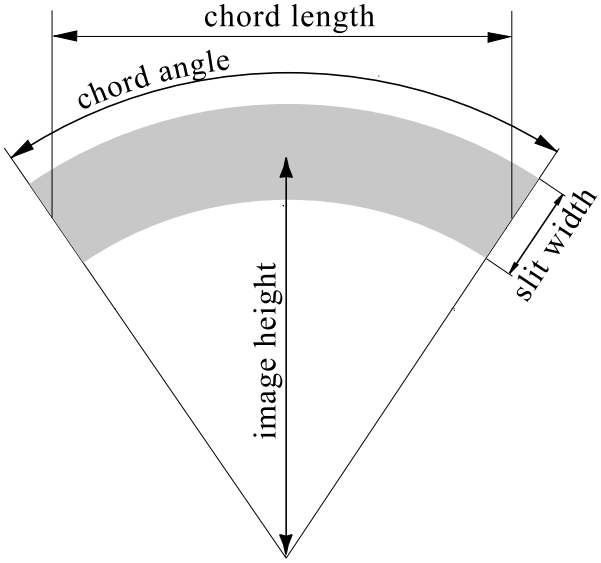

Fig. (2.4) The illuminated region on the wafer is a section of a ring-field. Typical values are 26 mm for the image height, 2 mm for the slit width, 60° or less for the chord angle and 26 mm for the chord length in a six-mirror system.

An example of a typical extreme ultraviolet projection system is shown in Fig. (2.5). The system considerations and requirements include:

- a high throughput of wafers. The resist needs a specific dose, so a high throughput of wafers demands a powerful source. A large slit width in combination with a high scanning or stepping velocity helps to reduce the intensity as much as possible. Otherwise, the multilayers and mask degrade too fast by heating.
- an accessible aperture stop surface,
- large free working distances near the object and image, see Fig. (2.10),
- enough beam clearance, as the mirrors need a small (mechanical) margin around the illuminated part,
- a reflective mask geometry with almost perpendicular illumination of the mask,
- small asphericities and angles of incidence,
- a magnification typically fixed at $\pm^1/_4$ or $\pm^1/_5$, to allow exchange of the mask with other lithographic systems,
- a high resolution,

- extremely small distortion,
- a perpendicular illumination of the image.

In the following subsections, we treat these and other requirements for the reflective ring-field projection systems, which are sometimes conflicting.

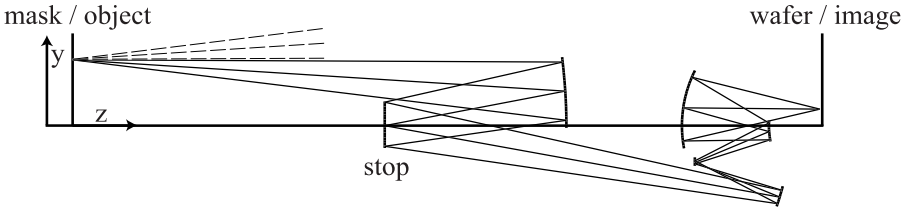

Fig. (2.5) In an extreme ultraviolet arrangement, the mask is imaged on the wafer with a reflective projection system. The requirements include a sufficiently large free working space, quasi telecentricity at the mask side, perfect telecentricity at the wafer side, and a fixed magnification. The condensor system illuminates the mask with a beam represented by the dotted line in the Figure.

Number of mirrors

Because extreme ultraviolet lithography uses short wavelength radiation and the image is required to be diffraction-limited, the demands on the mirrors are very severe. The number of mirrors is minimized because every additional multilayer-coated surface reduces the energy throughput. The reflection of an extreme ultraviolet reflector, coated with a multilayer, does not exceed a typical value of 70%. The design advantages of more mirrors are more optimization variables and a larger achievable numerical aperture. Many design variables are required to control the aberrations in order to obtain a diffraction-limited projection system. The required imaging quality corresponds to a root-mean-squared (rms) wavefront error smaller than $\lambda/50$ in image space at a large numerical aperture. These advantages have to be weighted against the disadvantages of a decreased transmission and an increased complexity of the extreme ultraviolet system.

The use of aspheric mirrors compensates the small number of surfaces and coefficients available to improve the imaging quality. The additional variables are the aspheric coefficients that represent the departure of each mirror from a sphere. The aspheric departure with respect to the base surface should preferably be small. In general, aspheric coefficients of a specific order can control or correct aberration terms of the same order[27]. Typically, field curvature and astigmatism are the more exacting aberrations to control within the ring-field.

The number of mirrors determines the maximum achievable numerical aperture. For instance, the numerical aperture of a four-mirror system may typically amount to 0.15, of a six-mirror system to 0.25 or 0.30 and of an eight-mirror system to 0.4. Increasing the numerical aperture of optical projection systems introduces larger aberrations, larger incidence angles on the multilayers, an increased chance of obscuration of the beam, and a smaller depth of focus.

Obstruction

Obstruction occurs when rays that should propagate freely between two consecutive surfaces are accidentally blocked by a third surface. Obstruction deteriorates the optical resolution and should therefore be avoided. To evaluate whether a system with N mirrors is obstructed or not, we examine the combinations of beam-fragments between successive surfaces i and i+1 (including object and image planes) and mirrors j for all possible pairs of i and j. A combination of a beam segment and a mirror is shown in Fig. (2.6).

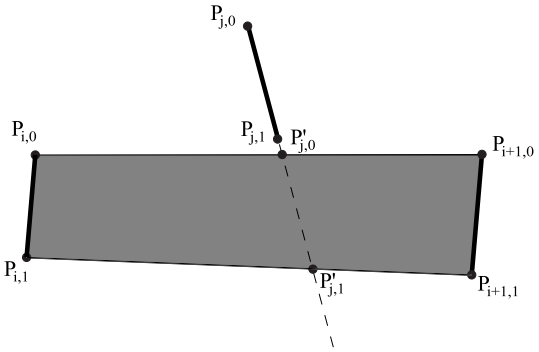

Fig. (2.6) When mirrors unintentionally intersect and block rays, the system is obstructed. In the Figure, the points $P_{i,0}$, $P_{i,1}$, $P_{i+1,0}$, $P_{i+1,1}$, $P_{j,0}$ and $P_{j,1}$ are the intersections of the extreme meridional rays with the surfaces. The thick lines are the reflective surfaces and in gray the beam propagates from mirror i to mirror i+1.

The lowermost and uppermost rays starting from the object determine the size of the beam. These two rays (called upper and lower meridional rays) intersect mirror j in two points $P_{j,0}$ and $P_{j,1}$ and delimit thus the useful part of the mirror. The heights of these points are denoted by $y_{j,0}$ and $y_{j,1}$ and their location along the optical axis by $z_{j,0}$ and $z_{j,1}$ respectively. The co-ordinates of the intersections $P_{i,0}$, $P_{i,1}$, $P_{i+1,0}$, and $P_{i+1,1}$ of the two extreme meridional rays with mirror i and mirror i+1 are denoted similarly. We extend surface j to find the intersection points $P'_{j,0}$ and $P'_{j,1}$ with the co-ordinates $y'_{j,0}$, $z'_{j,0}$ and $y'_{j,1}$, $z'_{j,1}$ of the two extreme meridional rays propagating between mirror i and i+1 with mirror j. These four co-ordinates can be obtained by solving the equations

$$y_{i,k} + \frac{y_{i+1,k} - y_{i,k}}{z_{i+1,k} - z_{i,k}}\left(z'_{j,k} - z_{i,k}\right) = y_{j,0} + \frac{y_{j,1} - y_{j,0}}{z_{j,1} - z_{j,0}}\left(z'_{j,k} - z_{j,0}\right) = y'_{j,k}\,, \tag{2.5}$$

with k=0 or 1. We discriminate between three different positions of mirror j relative to the position of the beam:

- $y_{j,0}$ and $y_{j,1}$ are both larger than $y'_{j,0}$ and $y'_{j,1}$, therefore the beam passes beneath the mirror,
- $y_{j,0}$ and $y_{j,1}$ are both smaller than $y'_{j,0}$ and $y'_{j,1}$, therefore the beam passes over the mirror,
- the beam is obstructed, since at least one of the points $P'_{j,0}$ and $P'_{j,1}$ belongs to the used part of mirror j. Obstruction occurs when the logical expressions

$$\bigvee_{k=0}^{1}\left(y'_{j,k} \in \left[y_{j,0}, y_{j,1}\right]\right) = \bigvee_{k=0}^{1}\left(z'_{j,k} \in \left[z_{j,0}, z_{j,1}\right]\right) \tag{2.6}$$

are fulfilled, where $\bigvee_{m}$ denotes the logical summation (or) extended over all values of integer m. The obstruction of the beam is either real or fictitious, depending on whether the mirror j intersects the beam within or outside the segment between the two mirrors

i and *i*+1, see Fig. (2.6). The real obstructions can be distinguished from the harmless fictitious obstructions with the logical expressions

$$\bigvee_{k=0}^{1}(z'_{j,k} \in [z_{i,k}, z_{i+1,k}]) = \bigvee_{k=0}^{1}(y'_{j,k} \in [y_{i,k}, y_{i+1,k}]), \tag{2.7}$$

which are true in case of real obstructions.

This special case of the evaluation of the occurrence of obstruction with one mirror and one beam segment leads to a general logical expression for the occurrence of obstruction *O* in a mirror system

$$O = \bigvee_{i=0}^{N}\bigvee_{\substack{j=1\\ j\neq i\\ j\neq i+1}}^{N}\bigvee_{k=0}^{1}(y'_{j,k} \in [y_{j,0}, y_{j,1}]) \wedge (y'_{j,k} \in [y_{i,k}, y_{i+1,k}]), \tag{2.8}$$

where $\wedge$ represents the logical multiplication (and). In Fig. (2.7) different situations of obstruction are shown.

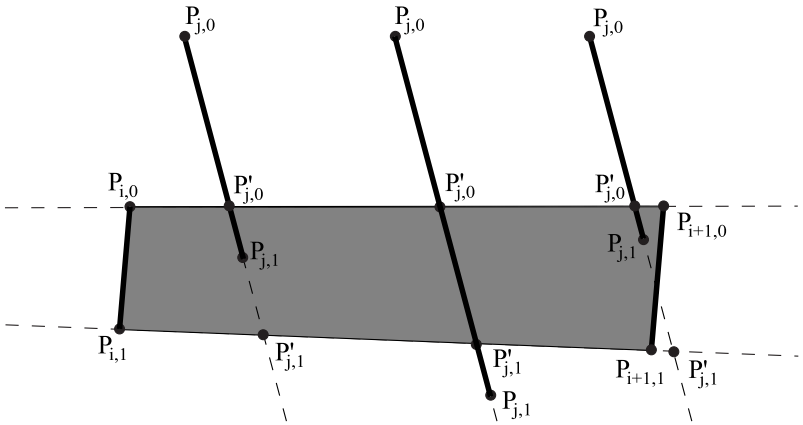

Fig. (2.7) A mirror can cause obstruction in different ways. On the left, the common situation is shown where only a part of the mirror with number *j* obstructs the beam. In the middle situation, the whole beam is obstructed. On the right, a part of the mirror is obstructed, while the other intersection point $P'_{j,1}$ is fictitious.

Telecentricity

In the case an optical system is telecentric in the object and / or image space, the principal ray is parallel to the optical axis in the object and / or image space. Telecentricity results in a constant magnification, even when the object and / or image distance varies. A non-constant magnification produces a lack of proportionality in the images; a defect named distortion. The distortion within the annular field in the image plane should be very small, not exceeding a few nanometers. To minimize the distortion induced by a defocus, the system must be as close to telecentricity as possible. Otherwise slight axial shifts of the mask or wafer plane from their ideal position cause unacceptable transverse image displacements. The effective depth of focus χ is also optimum for a telecentric beam, see Fig. (2.8).

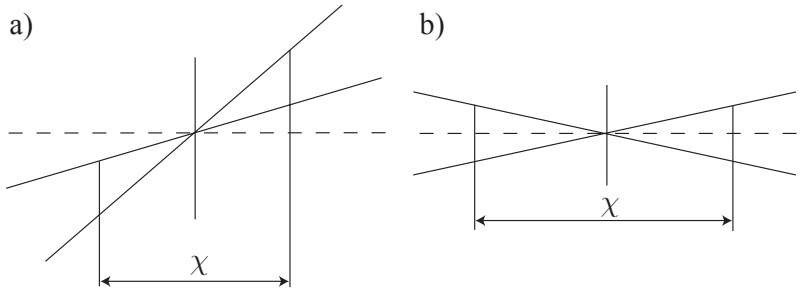

Fig. (2.8) The dashed line indicates the required axial feature position, corresponding to constant magnification. A defocus in a non-telecentric system induces distortion, see a). The magnification in a telecentric system remains constant with defocus, see b). The advantages of a telecentric system compared to a non-telecentric system also include the optimum depth of focus χ.

However, since the mask is reflective and illuminated by the condensor system, the projection system can only be *quasi-telecentric* at the object side. Either the upper or the lower meridional ray can be almost perpendicular to the mask surface, see Fig. (2.9). If a transmissive or stencil mask is used, the projection system can be telecentric on both the mask and wafer side.

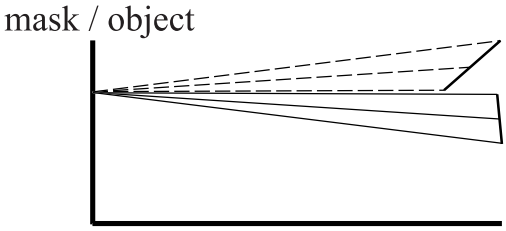

Fig. (2.9) The mask reflects a beam from the condensor system into the projection system. The beams can not be telecentric, since the two beams are not allowed to overlap. Otherwise the first mirror in the projection system obstructs the beam from the condensor system, or the last mirror in the condensor system obstructs the beam going into the projection system. The angles of incidence on the mask should be as small as possible that is *quasi-telecentric*, to maximize the tolerable defocus and to enhance the reflectivity of the multilayer on the mask.

Clearance

Both the wafer and the mask need enough mechanical clearance or workspace to scan or step, see Fig. (2.10). The best location of the object and image is both outside and on different sides of the projection system, for practical convenience. Using an odd number of mirrors results in a severe restriction in wafer or mask motion to avoid obstruction of the optical system. Therefore, we only investigate projection systems with an even number of mirrors. Nevertheless, the approach described in Chapter 4 works as well for odd-numbered mirror systems and examples of this kind of systems exist in the (patent) literature. Clearance can be a problem at the wafer side since the solid angle of the imaging bundles is maximal at this location.

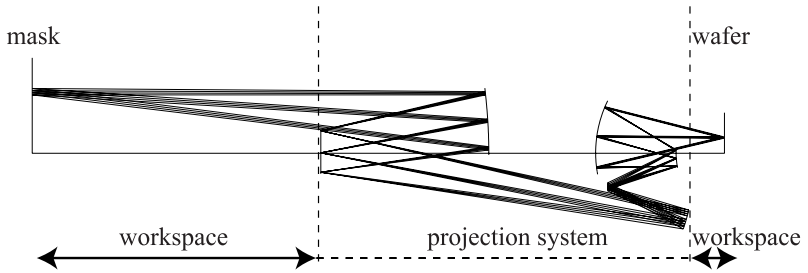

Fig. (2.10) In an extreme ultraviolet lithographic system, the mask is imaged on the wafer with a reflective projection system. The requirements include a sufficiently large free working space, quasi telecentricity at the mask side, perfect telecentricity at the wafer side, and a fixed magnification.

Existing designs

In the past years, several extreme ultraviolet projection systems have been designed, which are mainly four-, six-, and eight-mirror systems. The patent publications show designs of several different persons and companies[28-31]. The characteristics of some of these systems are summarized in Table (2.3); the operating wavelength of virtually all mirror systems is 13.4 nm.

Table (2.3) Systems from the patent literature, classified according to a convention described in paragraph 3.4.2.

Inventor	Class	Ring width	*NA*	Rms wavefront error in units λ
Bal[32]	4 mirrors 6-	1 mm	0.15	0.024 rms
Braat[33-36]	6 mirror 41+	1.8 mm	0.20	0.03 rms
	6 mirror 26-	2 mm	0.23	0.03 rms
	5 mirror 22+	1.5 mm	0.20	0.03 rms
Bruning[37, 38]	3 mirrors 5-	1 mm	0.10	0.030 rms
Dinger[39]	4 mirrors 6-	1 mm	0.10	0.029 rms
	4 mirrors 2-	1.25 mm	0.10	0.003 rms
Cohen, ETS[40]	4 mirrors 10-	1.5 mm	0.10	0.03 rms
Hudyma[41, 42]	4 mirrors 10-	1.5 mm	0.10	0.021 rms
	5 mirrors 21-	1.5 mm	0.18	0.015 rms
	6 mirrors 41+	2 mm	0.25	0.024 rms
Ichihara[43]	6 mirrors 41+	10 mm	0.02	
Jewell[44, 45]	4 mirrors 10-	1.5 mm	0.10	0.037 rms
Mann[46, 47]	6 mirrors 41+	2.0 mm	0.20	0.016 rms
	8 mirrors 150-		0.40	
	8 mirrors 165+		0.40	
Shafer[48-50]	8 mirrors 165+	1 mm	0.50	λ=126 nm
	6 mirrors 41+	1 mm	0.45	λ=126 nm
	4 mirrors 10-	1.5 mm	0.10	0.024 rms
Sweat[51]	4 mirrors 10-	2 mm	0.10	Strehl > 0.986
Williamson[52, 53]	4 mirrors 10-	2 mm	0.10	< 0.07 rms
	6 mirrors 45+	2 mm	0.25	0.01 nm

2.2.5 Multilayers

The multilayer coatings are designed to optimally reflect rays of light incident at a predetermined range of incidence angles. The larger the average angle of incidence, the larger the decrease in reflectivity for angles which deviate from the average angle. The phase deviations also become larger. The consequences of the presence of multilayers in an optical system are discussed in more detail in Chapter 7.

Two combinations of materials present in the multilayers were considered: a molybdenum / beryllium combination and a molybdenum / silicon combination[54, 55]. Beryllium is a toxic material. The molybdenum / silicon combination is more often used. The multilayer coatings are produced by electron-beam evaporation in combination with ion-bombardment of the layers and by direct current magnetron sputtering. The rotating substrates are mounted face down on spinner assemblies and sweep over sputter sources, see Fig. (2.11). A bilayer is deposited in each complete revolution of the spinner assembly. The substrates are rapidly spun around their own axes of symmetry to provide symmetry of the coatings. The layer thicknesses are determined by the time the substrate is exposed to a sputter source. The desired uniform or graded thickness distribution on a given optic is achieved by modulating the velocity of the substrate while it passes through the sputter flux.

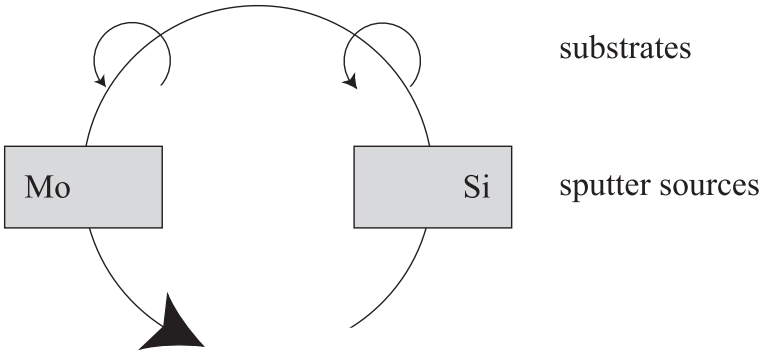

Fig. (2.11) The individual layers of the multilayers in the engineering test stand are deposited with two rectangular sputter sources. The substrates rotate quickly around their axes to achieve a better coating uniformity and rotate above the sources. With each rotation cycle, a pair of layers is deposited.

Figure errors

The total wavefront error of an optical system consists of:

- the residual errors in the design,
- the discrepancy between the specification in the design and the real system,
- the difference between reality and the model, due to approximations.

The critical specifications include the refractive indices of the materials and the shape and location of the transition between two materials. The figure error of a surface refers to the discrepancy between the modeled and actual surface shape. For extreme ultraviolet optics, the surface figure-errors divide in three groups:

- Low Spatial Frequency Roughness (LSFR) errors have a spatial wavelength larger than approximately 1 mm. Such errors are typically measured deterministically with

interferometers and expressed in a root-mean-squared wavefront error. In the ETS projection optics, the total surface error of each mirror is less than 0.25 nm root-mean-square. The Maréchal condition states that the wavefront error at the exit pupil must be smaller than $\lambda/14$ for diffraction-limited performance, corresponding to a Strehl ratio larger than 0.8. The extreme ultraviolet projection systems aim at a Strehl ratio above 0.9, which is a root-mean-squared wavefront error below $\lambda/20$.

- Mid-Spatial Frequency Roughness (MSFR) errors have a spatial wavelength between 1 μm and 1 mm. This scattering, often named flare, causes a reduction in the contrast of images because it scatters light from bright into dark intended regions at the image plane. Because the amount of scattering scales with λ^{-2}, the difficulty to control flare increases as the wavelengths in photolithography reduces[13, 14]. In extreme ultraviolet optics, the mid-spatial frequency errors must be approximately below 0.2 nm root-mean-square.
- High-Spatial Frequency Roughness (HSFR) errors have a spatial wavelength less than 1 μm. Typically these errors are statistically determined and induce large scatter angles. The light that is scattered at large angles ends up outside the image and reduces the intensity. The high spatial frequency roughness must be approximately below 0.1 nm in extreme ultraviolet optics.

2.3 Conclusion

The production and improvement of microelectronics is a major driver for the world economy. Lithographic methods constantly become more performing in imaging ever-smaller features on an integrated circuit. The current lithographic systems use wavelengths in the deep ultraviolet wavelengths. A future generation of lithographic systems will use much shorter wavelengths, to make the production of features smaller than approximately 70 nm possible. A photolithographic method working at extreme ultraviolet wavelengths of approximately 13 nm is the dominating candidate. The projector is the essential component of a photolithographic system, which images structures on a photoresist layer on the wafer. The design of projection systems that can be utilized in an extreme ultraviolet lithographic system, including the reflective multilayers, is the subject of this thesis.

2.4 References

1. H. Levinson, *Principles of lithography,* 1 ed. (The International Society for Optical Engineering, Bellingham, Washington, 2001).
2. "The National Technology Roadmap for Semiconductors" (Semiconductor Industry Association, 2002), retrieved http://www.sematech.org.
3. P. J. Silverman, "The Intel Lithography Roadmap," Intel Technology Journal **6**(2), 55-61 (2002).
4. C. Wagner, W. M. Kaiser, J. Mulkens, and D. G. Flagello, "Advanced Technology for extending Optical Lithography," in *Optical Microlithography XIII*, C. J. Progler, ed., Proc. SPIE **4000**, 344-357 (2000).
5. J. Burnett, Z. H. Levine, and E. L. Shirley, "Intrinsic birefringence in calcium fluoride and barium fluoride," Phys. Rev. B **64**(24), 1102 (2001).

6. D. Lammers, "EUV gains as ASML/Applied venture ends e-beam lithography work" (Semiconductor Business News, 2001), retrieved http://www.siliconstrategies.com/printableArticle?doc_id=OEG20010105S0065.
7. R. S. Dhaliwal, W. A. Enichen, S. D. Golladay, M. S. Gordon, R. A. Kendall, J. E. Liebermann, H. C. Pfeiffer, D. J. Pinckney, C. F. Robinson, J. D. Rockrohr, W. Stickel, and E. V. Tressler, "PREVAIL - Electron projection technology aproach for next-generation lithography," IBM J. Res. & Dev. **45**(5), 615-638 (2001).
8. M. Yoshizawa and T. A. Savas, "A feasibility study of 50 nm resolution with low energy electron beam proximity projection lithography," Jpn. J. Appl. Phys. 2 **41**(1AB), L87-L88 (2002).
9. T. Utsumi, "Low-Energy E-Beam Proximity Lithography (LEEPL): Is the Simplest the Best?," Jap. J. Appl. Phys. **38**(12B), 7046-7051 (1999).
10. B. Kampherbeek, M. Wieland, and P. Kruit, "Field emission photocathode array for lithography system and lithography system provided with such an array," WO patent 02/23580A1 (Mar. 21, 2002).
11. P. Kruit, "Lithography system," WO patent 98/54620 (Dec. 3, 1998).
12. S. Ohki and S. Ishihara, "An overview of X-ray lithography," Microelectron. Eng. **30**, 171-178 (1996).
13. C. W. Gwyn, R. Stulen, D. Sweeney, and D. Attwood, "Extreme ultraviolet lithography," J. Vac. Sci. Technol. B **16**(6), 3142-3149 (1998).
14. J. E. Bjorkholm, "EUV Lithography - The successor to optical lithography?," Intel Technology Journal **2**(3), 1-8 (1998).
15. J. P. H. Benschop, U. Dinger, and D. C. Ockwell, "EUCLIDES: First Phase Completed!," in *Emerging Lithographic Technologies IV*, Proc. SPIE **3997**, 34-47 (2000).
16. J. P. H. Benschop, A. J. J. van Dijsseldonk, W. M. Kaiser, and D. C. Ockwell, "EUCLIDES: European EUVL program," J. Vac. Sci. Technol. B **17**(6), 2978-2981 (1999).
17. J. P. H. Benschop, A. J. J. van Dijsseldonk, W. M. Kaiser, and D. C. Ockwell, "EUCLIDES: European EUV lithography milestones," Solid State Technol. **42**(9), 43- (1999).
18. H. Handschuh, J. Froschke, M. Julich, M. Mayer, M. Weiser, and G. Seitz, "Extreme ultraviolet lithography at Carl Zeiss: Manufacturing and metrology of aspheric surfaces with angstrom accuracy," J. Vac. Sci. Technol. B **17**(6), 2975-2977 (1999).
19. H. Meiling, J. P. H. Benschop, U. Dinger, and P. Kurz, "Progress of the EUVL alpha tool," in *Emerging Lithographic Technologies V*, E. A. Dobisz, ed., Proc. SPIE **4343**, 38-50 (2001).
20. A. Ishitani, "Next generation lithography and ASET's activities," Microelectron. Eng. **42**, 5-6 (1998).
21. M. LaPedus, "Japan's ASET group discloses EUV tool for 35-nm processes" (SPIE Microlithography conference in Santa Clara, CA, 2002), retrieved http://www.siliconstrategies.com/story/OEG20020306S0009.
22. V. Banine, J. P. H. Benschop, M. Leenders, and R. Moors, "The Relationship between an EUV Source and the Performance of an EUV Lithographic System," in *Emerging Lithographic Technologies IV*, E. A. Dobisz, ed., Proc. SPIE **3997**, 126-135 (2000).
23. H. N. Chapman, A. K. Ray-Chaudhuri, D. A. Tichenor, W. C. Replogle, R. H. Stulen, G. D. Kubiak, P. D. Rockett, L. E. Klebanoff, D. O'Connell, A. H. Leung,

K. L. Jefferson, J. B. Wronosky, J. S. Taylor, L. C. Hale, K. Blaedel, E. A. Spiller, G. E. Sommargren, J. A. Folta, D. W. Sweeney, E. M. Gullikson, P. Naulleau, K. A. Goldberg, J. Bokor, D. T. Attwood, U. Mickan, R. Hanzen, E. Panning, P. Y. Yan, C. W. Gwyn, and S. H. Lee, "First lithographic results from the extreme ultraviolet Engineering Test Stand," J. Vac. Sci. Technol. B **19**(6), 2389-2395 (2001).

24. G. Schriever, K. Bergmann, and R. Lebert, "Narrowband laser produced extreme ultraviolet sources adapted to silicon/molybdenum multilayer optics," J. Appl. Phys. **83**(9), 4566-4571 (1998).
25. J. Lin, V. Sharma, F. Michael, and A. Kirk, "Design and characterization of two-axis rotational micromirrors using multi-user microelectromechanical system processes," J. Microlith., Microfab., and Microsyst. **1**(1), 70-77 (2002).
26. Y. Shroff, Y. J. Chen, and W. Oldham, "Fabrication of parallel-plate nanomirror arrays for extreme ultraviolet maskless lithography," J. Vac. Sci. Technol. B **19**(6), 2412-2415 (2001).
27. W. J. Smith, *Modern lens design*, Optical and Electro-Optical Engineering Series (Mc Graw Hill, Boston, 1992).
28. R. Hudyma and D. Shafer, "High numerical aperture ring field projection system for extreme ultraviolet lithography," US patent 6,188,513 (Feb. 2, 2001).
29. R. Hudyma, "High numerical aperture ring field projection system for extreme ultraviolet lithography," US 6,033,079 1999).
30. M. Suzuki, N. Mochizuki, S. Minami, S. Ogura, Y. Fukuda, Y. Watanabe, Y. Kawai, and T. Kariya, "X-ray reduction projection exposure system of reflection type," US 5,153,898 1988).
31. V. Viswanathan and B. Newnam, "Reflective optical imaging system for extreme ultraviolet wavelengths," US patent 5,212,588 (May 18, 1993).
32. M. Bal, F. Bociort, and J. Braat, "Lithographic apparatus and device manufacturing method," European patent EP 1 20 95 03 A2 (May 29, 2002).
33. J. Braat, "Mirror projection system for a scanning lithographic projection apparatus, and lithographic apparatus comprising such a system," US 6,299,318 1998).
34. J. Braat, "Lithographic apparatus comprising a dedicated mirror projection system," US 6,396,067 1999).
35. J. Braat, "Mirror projection system for a scanning lithographic projection apparatus, and lithographic apparatus comprising such a system," US 6,255,661 1999).
36. J. Braat and J. Verhoeven, "Method of imaging a mask pattern on a substrate by means of euv radiation, and apparatus and mask for performing the method," US 6,280,906 1999).
37. J. Bruning, A. Phillips, D. Shafer, and A. White, "X-ray projection lithography camera," US patent 5,220,590 (Jun. 15, 1993).
38. J. Bruning, A. Phillips, D. Shafer, and A. White, "Lens system for X-ray projection lithography camera," US patent 5,353,322 (Oct. 4, 1994).
39. U. Dinger, "Ringfeld-4-Spiegelsysteme mit konvexem Primarspiegel fur die EUV-Lithography," European patent EP 0 962 830 A1 (Dec. 8, 1999).
40. S. Cohen, H. Jeong, and D. Shafer, "Four-mirror extreme ultraviolet (EUV) lithography projection system," US patent 6,142,641 (Nov. 7, 2000).
41. R. Hudyma, "High numerical aperture ring field projection system for extreme ultraviolet lithography," US patent 6,318,869 (Nov. 20, 2001).
42. R. Hudyma, "Reflective optical imaging systems with balanced distortion," US patent 6,226,346 (May 1, 2001).

43. Y. Ichihara, "Reflecting optical imaging apparatus using spherical reflectors and producing an intermediate image," US patent 5,071,240 (Dec. 10, 1991).
44. T. Jewell and J. Rodgers, "Apparatus for semiconductor lithography," US patent 5,063,586 (Nov. 5, 1991).
45. T. Jewell and K. Tompson, "X-ray ringfield lithography," US 5,315,629 1991).
46. H.-J. Mann and R. Hudyma, "Projection system for EUV lithography," WO patent 02/48796A2 (Jun. 20, 2002).
47. H.-J. Mann, W. Ulrich, and G. Seitz, "8-mirrored microlithographic projector lens," World Intellectual Property Organisation WO patent 02/33467A1 (Apr. 25, 2002).
48. D. Shafer, "Projection lithography system and method using all-reflective optical elements," US 5,686,728 1996).
49. D. Shafer, "Reflective projection system comprising four spherical mirrors," US 5,410,434 1993).
50. D. Shafer, "Reflective optical imaging system," US patent 6,014,252 (Jan. 11, 2000).
51. W. Sweat, "Ringfield lithographic camera," US patent 5,805,365 (Sept. 8, 1998).
52. D. Williamson, "Four mirror EUV projection optics," US 5,956,192 1997).
53. D. Williamson, "High numerical aperture ring field optical reduction system," US patent 5,815,310 (Sep. 29, 1998).
54. E. Louis, H. J. Voorma, N. B. Koster, F. Bijkerk, Y. Y. Platonov, S. Y. Zuev, S. S. Andreev, E. A. Shamov, and N. N. Salashchenko, "Multilayer Coated Reflective Optics for Extreme Uv Lithography," Microelectron. Eng. **27**(1-4), 235-238 (1995).
55. C. Montcalm, E. Spiller, F. J. Weber, M. Wedowski, J. A. Folta, and E. M. Gullikson, "Multilayer coating and test of the optics for two new 10X Microstepper extreme-ultraviolet lithography cameras," J. Vac. Sci. Technol. B **19**(4), 1219-1228 (2001).

3 Modeling of optical systems

Light is an electromagnetic (EM) wave, which satisfies the Maxwell's equations. For most optical design problems, we can neglect the vector nature of the field by assuming a scalar wave. Moreover, in homogeneous media and for monochromatic radiation, the Maxwell equations reduce to the so-called time-independent Helmholtz equation

$$\left(\nabla^2 + k^2\right)U(x,y,z) = 0\,, \tag{3.1}$$

where $k = \omega/c = 2\pi/\lambda$ is the (circular) wave number and U is the amplitude of the chromatic wave. Wave theory applied to imaging problems shows that the image will have a finite smallest size, called the diffraction limit. The propagation of waves is commonly described by means of wavefronts. A wavefront is a bounded surface with a constant phase. Fourier analysis of the wavefront is used to account for diffraction effects.

Except in focal regions and near aperture boundaries, the wave-like nature of light can be further ignored in most optical problems using the limiting case of $\lambda \to 0$. The propagation direction of the light energy now is represented by a ray that, by definition, is perpendicular to the wavefront. The behavior of a ray at a transition between two materials is described by Snel's law, which can be deduced from Fermat's principle. Fermat's principle states that the optical path length along a physically possible ray between two points has a stationary value. Snel's law states that a ray incident on a surface i at an angle θ_i with the normal of the surface refracts at an angle θ'_i with the normal to the surface

$$n_i \sin(\theta_i) = n_{i+1} \sin(\theta'_i)\,. \tag{3.2}$$

Reflection can be considered a special case of refraction, where the two refractive indices are equal but opposite in sign. Image evaluation based on rays is called geometrical optics. Ray tracing is a successive combination of translating rays from one surface to the next, finding the intersection point of the ray with a transition between two materials, and calculating the angle after refraction. Geometrically, it is possible that all rays from an object point converge to a single image point, leading to a so-called stigmatic or perfect image point.

3.1 Paraxial optics

Paraxial optics deals with the propagation of light through a system with refracting or reflecting surfaces that, in our case show rotational symmetry around the common optical axis of all surfaces. In the paraxial approximation, the real surfaces are replaced by approximating surfaces whose second order curvature is identical. Any aspheric coefficients are ignored, by which a reflective system with N mirrors is described by N curvatures, N-1 distances between mirrors, an object distance, and an image distance. Strictly spoken, the paraxial approximation only applies to rays that are displaced infinitesimally from the optical axis of a system.

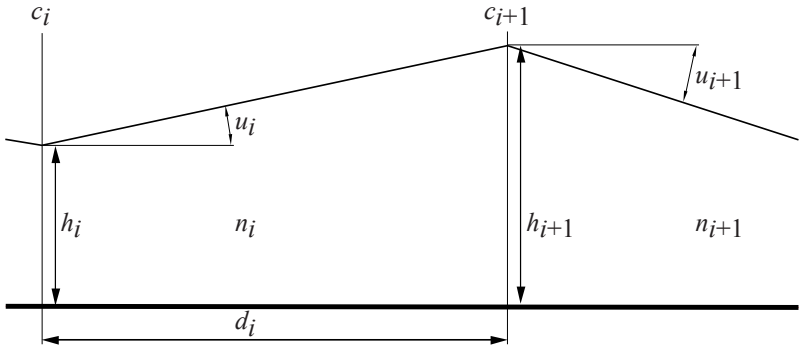

Fig. (3.1) The propagation of a ray that travels from surface i to surface i+1, calculated with paraxial optics. The optical properties of a system are defined by attaching to each surface its curvature c, the distance d to the subsequent surface, and the refractive index n_i of the medium between surface i and i+1.

Let **M** be the matrix that relates rays with object height h and angle u_0 from the object plane to rays in the image plane with height y_{N+1} and angle u_{N+1} as follows[1]

$$\begin{pmatrix} y_{N+1} \\ u_{N+1} \end{pmatrix} = \mathbf{M} \cdot \begin{pmatrix} h \\ u_0 \end{pmatrix} + O(3), \tag{3.3}$$

where $O(3)$ represents all higher order terms neglected by the paraxial approximation. The matrix **M** can then be written as

$$\mathbf{M} = \mathbf{T}_N \cdot \mathbf{R}_N \cdot \mathbf{T}_{N-1} \cdot \mathbf{R}_{N-1} \cdot \ldots \cdot \mathbf{T}_2 \cdot \mathbf{R}_2 \cdot \mathbf{T}_1 \cdot \mathbf{R}_1 \cdot \mathbf{T}_0 . \tag{3.4}$$

The transfer matrices $\mathbf{T}_i$ are of the form

$$\mathbf{T}_i = \begin{bmatrix} 1 & d_i \\ 0 & 1 \end{bmatrix}. \tag{3.5}$$

In the case of reflection, the refractive index changes sign

$$n_i = -n_{i-1} . \tag{3.6}$$

Therefore, the refraction matrix simplifies to the reflection matrix $\mathbf{R}_i$ if the ray is reflected on surface i

$$\mathbf{R}_i = \begin{bmatrix} 1 & 0 \\ c_i \dfrac{n_{i-1} - n_i}{n_i} & \dfrac{n_{i-1}}{n_i} \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 \\ -2c_i & -1 \end{bmatrix}. \tag{3.7}$$

In the above formulas, the refractive index of the medium after surface i is denoted by n_i, the curvature of surface i by c_i, and the distance from surface i to the successive surface by d_i, see Fig. (3.1). Note that in the case of reflection both the transfer and reflection matrices depend on one single variable, either a distance or a curvature.

From each object point, a pencil of rays propagates through the optical system. The principal or chief ray passes through the center of the stop surface and in the absence of aberrations through the center of the entrance pupil and the center of the exit pupil, leading to the following matrix equation

$$\mathbf{M} \cdot \begin{pmatrix} h \\ \overline{u}_0 \end{pmatrix} = \mathbf{T}_N \cdot \mathbf{R}_N \cdot \mathbf{T}_{N-1} \cdot \mathbf{R}_{N-1} \cdot \ldots \cdot \mathbf{T}_s \cdot \mathbf{R}_s \cdot \begin{pmatrix} 0 \\ \overline{u}_s \end{pmatrix} = \begin{pmatrix} \overline{y}_N \\ \overline{u}_N \end{pmatrix} = \begin{pmatrix} M_t\, h \\ \overline{u}_N \end{pmatrix}, \tag{3.8}$$

where M_t is the transverse paraxial magnification. The bar $\bar{}$ above a variable refers to the principal ray. The two marginal rays pass through the rim of the diaphragm or stop surface,

which is surface number s in the system. Since we assume rotational symmetric systems, we only have to consider one of the marginal rays

$$\mathbf{M} \cdot \begin{pmatrix} 0 \\ NO \end{pmatrix} = \mathbf{T}_N \cdot \mathbf{R}_N \cdot \mathbf{T}_{N-1} \cdot \mathbf{R}_{N-1} \cdot \ldots \cdot \mathbf{T}_s \cdot \mathbf{R}_s \cdot \begin{pmatrix} y_s \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ NA \end{pmatrix} = \begin{pmatrix} 0 \\ M_a NO \end{pmatrix}, \quad (3.9)$$

where NA is the numerical aperture at the image side, NO is the numerical aperture at the object side, and M_a is the angular magnification.

The incidence plane is defined by the incident and refracted ray. When the incidence plane includes the optical axis, the ray is said to be a meridional ray, otherwise a skew ray. The meridional rays limit the beam going through the system. Paraxially, the addition of the ray angles of the principal and a marginal ray leads to the angles of the upper and lower meridional rays, see Fig. (3.2).

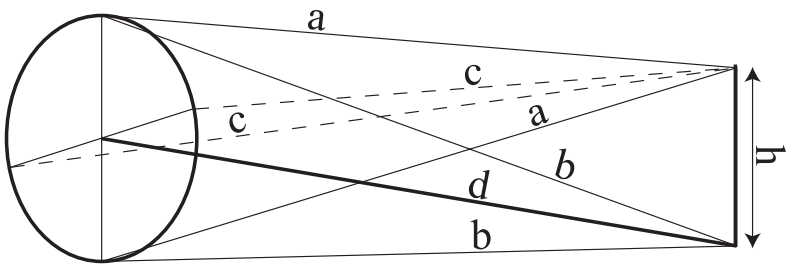

Fig. (3.2) The meridional plane comprises the object, the optical axis and all meridional rays. In the Figure, the entrance pupil is on the left side and the object at a distance h from the optical axis. The different line segments represent: a) the upper and lower meridional rays of an off-axis pencil, b) the upper and lower marginal rays of the axial pencil, c) the left and right sagittal rays, and d) the optical axis.

A perfect lens forms a sharp, undistorted image of a planar object. Although paraxial rays can be used to predict the location and magnification of images, paraxial ray trajectories do not obey Fermat's principle, except in a region very close to the optical axis. A perfect optical system cannot form a perfect image at more than one magnification and the trajectories of rays through it cannot be found using paraxial ray tracing.

3.2 Invariants

The principle of Fermat states that a light ray travels an optical path length that is stationary with respect to variations of that path. From this principle the law of refraction of Snel, see Eq. (3.2), could be deduced.

For the aberration-free imaging of objects with finite size outside the paraxial domain, an important condition can be derived which is known as the *sine condition* that relates the angles of finite aperture rays in object and image space according to

$$n_1 y_1 \sin(u_1) = n_0 y_0 \sin(u_0), \quad (3.10)$$

where u_i is the finite angle of the marginal ray and y_i the generally small distance of the object and image point to the optical axis. A so-called aplanatic optical system respects the sine condition and approaches the ideal of a perfect lens. Within the paraxial approximation an optical Helmholtz or Lagrange invariant can be deduced from the sine condition

$$H = n_i y_i u_i \,, \tag{3.11}$$

where u_i now is restricted to the paraxial domain.

3.3 Representation of optical surfaces and aberrations

For each object point a bundle of rays emanates from the entrance pupil and is traced through the system to the image plane. In the presence of aberrations, the intersections points of these rays with the image plane form a pattern (named spot diagram or transverse ray aberrations) that is characteristic of the type of aberrations present. An ideal system will focus all rays from any given field point in the image plane to a single point on the image surface and will therefore have a zero spot size.

The shape of the wavefront emerging from a system that produces sharp point images for a given field point is that of a sphere centered on a point on the image surface. The distance to the center of the exit pupil is the radius of this reference sphere. The optical path difference (OPD) for a single ray from a given object field point is the distance along the ray from the reference sphere to the wavefront multiplied with the refractive index in the image space, see Fig. (3.3). The imaging quality of a system can be expressed in terms of these deviations of the emerging wavefront (named the wavefront or wave aberrations) from this reference sphere. Via integration along a chosen path, the wavefront aberrations can be calculated from the transverse aberrations.

Other methods to characterize the imaging performance of an optical system include the modulation transfer function (MTF), which is the modulus of the optical transfer function (OTF). The optical transfer function quantifies the potential of the system to reproduce frequency components of the object intensity distribution in the image plane.

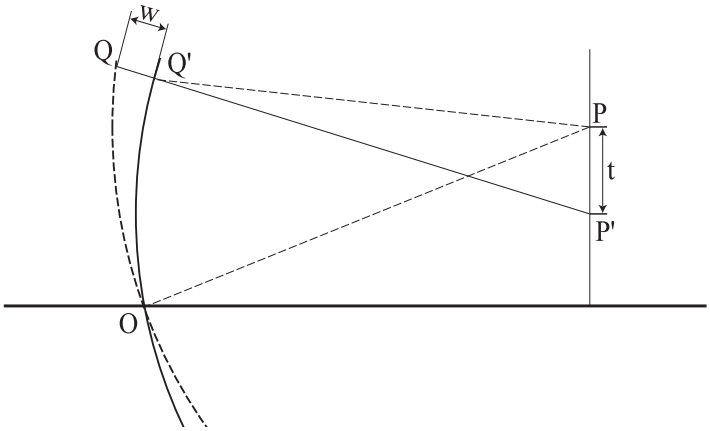

Fig. (3.3) The transverse and the wavefront aberrations both express the imaging quality of a system. The reference sphere is represented on the left by the dashed part of a circle; the distance OP is its radius of curvature. The distance PP' between the intersection points of two rays with the image plane is *t*, representing the transverse aberration of a ray. The optical distance QQ' between the intersections of a ray with the wavefront and the reference sphere is the wavefront aberration.

Generally, when the system is far away from the diffraction limit, the imaging characteristics of a system are expressed in terms of transverse ray aberrations. Near the

diffraction limit, the wavefront aberrations are used. With the laws of geometrical optics, the deviations of the wavefront from the ideal sphere in the exit pupil are calculated. Often these deviations are expressed as a function of Seidel coefficients or as a function of an orthogonal set of functions, e.g. Zernike polynomials.

In the next sections, we treat several methods and specific polynomial sets to represent a surface. The surface to be represented can be the wavefront deviation expressed as a function of the pupil coordinates, as well as a reflective or refractive surface expressed as a function of the position on the surface. The surfaces we use are rotationally symmetric and are a function of a single variable, the lateral radial distance. This is in contrast with the wavefront surfaces, which are functions of two independent variables, in e.g. the exit pupil. The shape of the surfaces present in an extreme ultraviolet projection system is extremely accurately adjusted to achieve an optimal imaging quality. At the same time, the number of surfaces is kept as low as possible, because of throughput reasons. Several numerical methods have been developed and proposed to efficiently compute form factors, like those to define high-order aspherical surfaces[2].

3.3.1 Seidel aberrations

The analytic computation of aberration expansion coefficients for rotationally symmetric systems has a long tradition. Expressions for the third order coefficients for centered rotationally symmetric systems were developed by Seidel in the 19th century and they express the monochromatic geometrical aberrations of rotationally symmetric optical systems. The Seidel coefficients are computed from paraxial ray data and are named spherical aberration, coma, astigmatism, field curvature, and distortion. The contributions of each surface to the Seidel aberrations are independent. The higher order aberrations of a surface are, however, affected by aberrations originating at preceding surfaces.

We first define the two quantities

$$P_i = \frac{n_{i-1} - n_i}{n_i n_{i-1}} y_i \left(\frac{\theta_i}{n_{i-1}} + u_{image} \right), \tag{3.12}$$

$$\overline{P}_i = \frac{n_{i-1} - n_i}{n_i n_{i-1}} \overline{y}_i \left(\frac{\overline{\theta}_i}{n_{i-1}} + \overline{u}_{image} \right). \tag{3.13}$$

The five lowest order Seidel aberration coefficients, numbered S_1 to S_5 are given by

Spherical aberration

$$S_1 = \sum_{i=1}^{N} P_i \theta_i^2 + (8 s_{i,1} + k_i c_i^3)(n_{i-1} - n_i) r_i^4, \tag{3.14}$$

Coma

$$S_2 = \sum_{i=1}^{N} P_i \theta_i \overline{\theta}_i + (8 s_{i,1} + k_i c_i^3)(n_{i-1} - n_i) r_i^3 \overline{r}_i, \tag{3.15}$$

Astigmatism

$$S_3 = \sum_{i=1}^{N} P_i \overline{\theta}_i^2 + (8 s_{i,1} + k_i c_i^3)(n_{i-1} - n_i) r_i^2 \overline{r}_i^2, \tag{3.16}$$

Distortion

$$S_4 = \sum_{i=1}^{N} \overline{P}_i \theta_i \overline{\theta}_i + H \frac{n_{i-1} - n_i}{n_i n_{i-1}} \overline{\theta}_i \overline{u}_{image} + \overline{u}_{image-1} + (8 s_{i,1} + k_i c_i^3)(n_{i-1} - n_i) r_i \overline{r}_i^3 , \quad (3.17)$$

Petzval sum or curvature of field

$$S_5 = \sum_{i=1}^{N} c_i \frac{n_{i-1} - n_i}{n_i n_{i-1}} H^2 , \quad (3.18)$$

where θ_i are the paraxial angles of incidence, $s_{i,1}$ the fourth order aspherical coefficient of surface i, see Eq. (3.19). The distance between the intersection with surface i and the optical axis is r_i. A proper linear combination of the Seidel coefficients yields the final fourth-order wavefront aberration of a ray.

The Buchdahl coefficients are the higher-order equivalents of the Seidel coefficients and describe the transverse aberrations referring to rotational invariant coordinates in the entrance pupil. More recently, Andersen developed a more advanced iterative method to calculate arbitrary order aberration coefficients[3, 4].

3.3.2 Conic plus even-order polynomial representation

Generally, in optical design software rotationally symmetric aspherical surfaces consist of a base surface represented by a conicoid, augmented with even-order monomial

$$z_i(r_i) = \frac{c_i r_i^2}{1 + \sqrt{(1 - (1 + \kappa_i) c_i^2 r_i^2)}} + \sum_{j=2} s_{i,j} r_i^{2j} , \quad (3.19)$$

where z_i is the sag of the surface parallel to the z-axis. The curvature at the pole of the surface is c_i, r_i is the distance perpendicular to the optical axis ($r^2 = x^2 + y^2$) and $s_{i,j}$ are aspherical coefficients of surface i. The conic constant κ changes the basic form of the surface:

$\kappa > 0$	ellipsoid of revolution generated by an elliptical cross-section with its major axis perpendicular to the optical axis (oblate spheroid),
$\kappa = 0$	sphere,
$-1 < \kappa < 0$	ellipsoid with the major axis coincident with the optical axis (prolate spheroid),
$\kappa = -1$	paraboloid,
$\kappa < -1$	hyperboloid.

The conicoid base function and the *aspheric* monomials do not constitute a set of orthogonal functions. Even stronger: the conic constant is redundant with respect to the set of aspherical constants used. Nonetheless, the conic constants remain practical variables to vary during optimization, because they are capable of keeping the absolute values of the aspheric coefficients small. The non-orthogonality of the optimization variables is an awkward feature: a change of one value entails an optimization or recalculation of other variables of the same surface. The values are also prone to round-off and truncation errors, as they are not scaled.

3.3.3 Representation using orthogonal functions

An arbitrary function can be written as an expansion of the form

$$z(t) = \sum_{i=0} g_i T_i(t)\,, \tag{3.20}$$

where z is for instance the surface sag, T_i is a suitable set of basis functions and g_i the set of coefficients of the expansion. Two possibly complex functions are orthogonal on the interval $[t_1,t_2]$ if their inner product is zero

$$\int_{t_1}^{t_2} T_i(t) T_j^*(t)\, w(t)\, dt = \varepsilon_i \delta_{i,j}\,, \tag{3.21}$$

where $w(t)$ is a weighting function, ε_i are real constants and $\delta_{i,j}$ the Kronecker delta function, which has the properties

$$\delta_{i,j} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}. \tag{3.22}$$

The value of each coefficient of an orthogonal expansion is independent of the values of the other coefficients. A function is normalized if the inner product with itself is unity, in this case ε_i=1 for every i. An orthonormal set consists of a set of functions that are all mutually orthogonal and also are all individually normalized. The set is complete if any function $z(t)$ can be represented on the interval $[t_1,t_2]$ by a superposition of the polynomials. Examples of orthonormal sets are the Fourier basic functions on the interval [-∞,∞], the Legendre polynomials on the interval [-1,1], the Laguerre polynomials on the interval [0,∞], the Hermite polynomials on the interval [-∞,∞], and the Chebyshev polynomial on the interval [-1,1], all associated with a certain choice of the weighting function $w(t)$.

When using an orthogonal set to describe the shape of a surface, it is practical to check if the surface is not overdetermined by too many coefficients. When the low-order aberrations contribute largely to the total aberrations, a surface hardly profits from higher order coefficients. To evaluate the effect of decreasing the number of coefficients in the description of an aspherical surface on the system aberrations, the higher coefficients can simply be ignored when an orthogonal set is used. Although, strictly speaking, the set of polynomials is only orthogonal with respect to the surface sag, the small higher order aberrations often show a linear dependence on varying z-value of the same order. Observation has confirmed this relation between the order of aberrations and the order of the polynomials used in the surface descriptions.

Optimization routines normally compute and use an orthogonal set of variables and their derivatives with respect to the aberrations. Roundoff and truncation errors are avoided with orthonormal polynomials, since the higher order coefficients are not as small as the coefficients of the standard monomials.

Although the individual polynomials of an orthogonal set are orthogonal over the specified interval, the polynomials are not orthogonal over portions of the interval. This can be cumbersome if the polynomials are fit to a limited set of data points or when only a part of a surface is used. Sasian proposes to use rescaled aspheric surface descriptions in annular field surfaces, which leads to shortened design times[5]. When we adopt the orthogonal polynomials to describe the sag of a surface, we rescale the used part of a surface to the orthogonal interval.

Zernike polynomials

The Zernike polynomials are a complete orthogonal set on the unit-circle defined by two variables, the normalized radius *r* and the angle ψ. The Zernike polynomials separate into a radial and an angular function of the form

$$Z(r,\psi) = R(r)G(\psi) = R_n^m(r)e^{im\psi}, \tag{3.23}$$

where *k* and *m* are integers, fulfilling the following conditions

$$0 \le m \le n, \tag{3.24}$$

$$n - m = 2a \quad \wedge \quad a \in \mathbb{N}. \tag{3.25}$$

The real and complex parts in Eq. (3.23) each form a polynomial. The radial part of the Zernike polynomials is given by

$$R_n^m(r) = \sum_{s=0}^{\frac{n-m}{2}} (-1)^s \frac{(n-s)!}{s!(\frac{n+m}{2}-s)!(\frac{n-m}{2}-s)!} r^{n-2s}. \tag{3.26}$$

The lower orders Zernike polynomials can for instance be found in *Principles of Optics*[6]. The deviation of the wavefront from an ideal sphere is often represented in terms of Zernike polynomials, especially in optical interferometric testing. The Zernike coefficients are ordinarily found using a least-squares fit to a grid of exact ray data. Aspheric surfaces are sometimes also specified with the aid of an expansion in Zernike polynomials.

The Seidel aberrations are not uniquely related to the Zernike polynomials. Only when the high-order aberrations are negligible, the first few Zernike coefficients can be directly related to the Seidel aberrations.

Chebyshev polynomials

The Chebyshev polynomials of degree *n* are given by[7]

$$T_n(t) = \cos(n\ \cos^{-1}(t)). \tag{3.27}$$

The first polynomials and the general recursion formula are given by

$$T_0(t) = 1, \tag{3.28}$$

$$T_1(t) = t, \tag{3.29}$$

$$T_{n+1}(t) = 2\,x\,T_n(t) - T_{n-1}(t) \quad n \ge 1. \tag{3.30}$$

In the interval [-1,1] the Chebyshev polynomials are orthonormal. Another practical property is that the difference between an ideal surface with an infinite number of coefficients and a surface with only a few coefficients is distributed smoothly over the interval [-1,1].

The rotationally symmetric surfaces are defined as a function of the lateral distance, the radius. The optically used part of the mirror corresponds to an interval $[r_1,r_2]$ of the radius, with the absolute minimum r_1 and the absolute maximum r_2 of the radius. The [-1,1] interval of the variable *t* of the Chebyshev polynomials is converted to this interval of the radius. The two transform factors α_0 and α_1 are given by

$$r^2 = \alpha_0 + \alpha_1 t, \tag{3.31}$$

$$r_1^2 = \alpha_0 - \alpha_1, \tag{3.32}$$

$$r_2^2 = \alpha_0 + \alpha_1 . \tag{3.33}$$

The sag of the surface as a function of the Chebyshev polynomials T_i and coefficients τ_i becomes

$$z(t) = z(\frac{r^2 - \alpha_0}{\alpha_1}) = \sum_{i=0} \tau_i T_i(t) . \tag{3.34}$$

A routine was written to transform a surface defined as a function of a standard monomial with an arbitrary conic constant, see Eq. (3.19), into a standard monomial with a conic constant equal to –1. This is not an exact transformation, as the contributions of twentieth and larger orders of the polynomials are neglected. Another routine exactly converts the standard parabolic monomial expansion to Chebyshev polynomials. A user defined function returns the sag of the surfaces defined by Chebyshev polynomials to the optical design program. The comparison of the systems before and after conversion needs some attention. The optical design program sometimes relocates the exit pupil, which results in different total aberrations. The change to a conic constant of –1 using standard monomials sometimes causes an increase of the aberration, especially in highly corrected systems.

A comparison of the use of standard monomials and Chebyshev polynomials to express the shape of a surface is shown in Fig. (3.4). When the number of coefficients is decreased, the aberration increases with both the standard expansion and the Chebyshev polynomials. The curve for the standard monomials however has a valley, which points out the non-orthogonal behavior of the standard expansion. Expressing the shape of a surface in translated Chebyshev polynomials instead of standard monomials saves approximately one coefficient, at the cost of more time needed by the optical design program to evaluate a system.

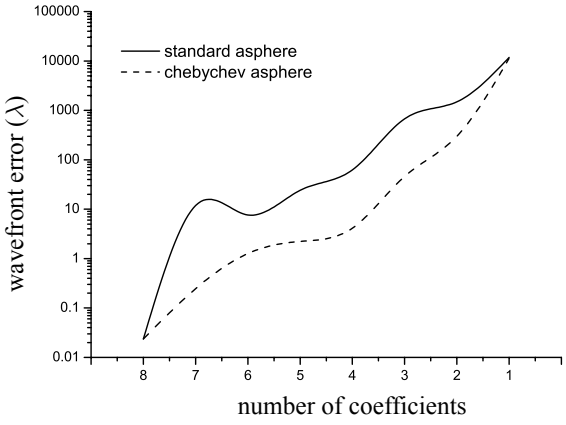

Fig. (3.4) The wavefront error as a function of the number of coefficients used in both the standard monomial and the Chebyshev polynomials. The system under test is a four-mirror system.

The expression for the shape of the surface in standard monomials and in Chebyshev polynomials has been compared for a large collection of surfaces. In Fig. (3.5), the number of coefficients to describe the shape of the surfaces is decreasing along the horizontal axis. The graph shows the average ratio of the root-mean-squared wavefront aberration of surfaces expanded with standard monomials and with Chebyshev polynomials. The systems

start and end with the same wavefront error as expected. Without coefficients, both surfaces are parabolas. The standard aspheric surfaces of the systems were converted to Chebyshev polynomials using eight coefficients, which makes the surfaces identical when described by eight standard aspherical monomials or by eight Chebyshev polynomials. Not all surfaces make (effective) use of all coefficients, causing the *plateau* between seven and five coefficients.

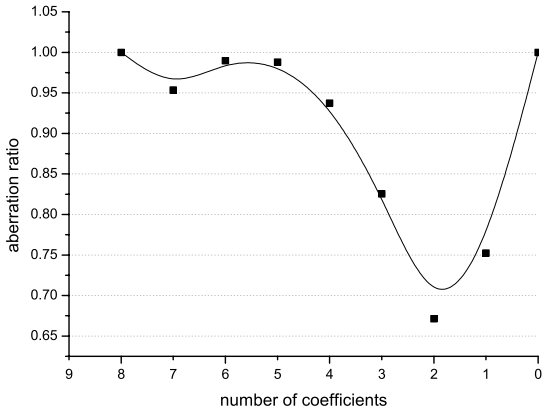

Fig. (3.5) A comparison between the expression of the shape of a surface in standard monomials and in Chebyshev polynomials. On the horizontal axis, the number of coefficients is plotted that is used in the polynomial expansion to describe a surface. The ratio between the root-mean-squared (rms) wavefront aberration of a surface expressed in Chebyshev polynomials and the same surface expressed in standard monomials is plotted along the vertical axis.

3.3.4 Distortion

The spot size or the root-mean-squared (rms) wavefront do not depend on the image distortion of the system. To control distortion, a set of additional terms are added to the error function of an optimization (see Section 5.2). A remaining requirement for ideal imagery is that the image must be geometrically similar to the object. Positive distortion is called pincushion distortion, while negative distortion is referred to as barrel distortion. To minimize the creation of distortion by mispositioning of the mask or wafer, the beams are as telecentric as possible, that is perpendicular to mask and wafer surface, see the subsection on **telecentricity** (page 26). In extreme ultraviolet projection systems, only an extremely small distortion is acceptable, in the order of a few nanometers.

3.4 Classification of reflective projection systems

The optical power of a refractive or reflective surface is given by

$$p_i = c_i\,(n_i - n_{i-1})\,. \tag{3.35}$$

A positive reflecting power is concave and typically converges the incident light, the opposite holds for a convex surface. The constraint to have an unobstructed beam

propagation through the reflective system results in small allowed domains, as we show in Section 4.2.1. These domains are bounded by obstructions. An ideal classification method for these reflective systems groups the systems in the same unobstructed domain and separates systems in different domains. A typical optimization routine should be able to modify a system into any other system in the same solution space, but is unable to modify a system into systems from other unobstructed domains.

3.4.1 Classification by the sign of the power of the surfaces

A method to classify systems is to evaluate the sign of the power of each mirror, see Eq. (3.35). When a positive power of a mirror results in a P and a negative power in a N, a *n*-mirror system ends up with a *n*-letter combination. The consideration that testing convex (negative power) aspherical mirrors is more difficult than testing concave (positive power) aspherical mirrors, gives preference to a system with only positive surfaces.

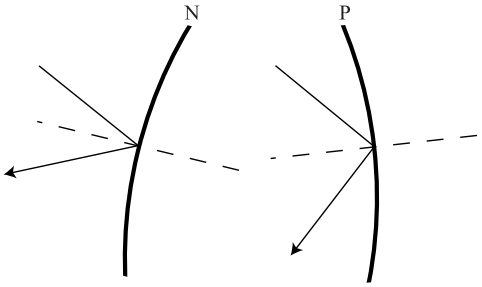

Fig. (3.6) The classification by the sign of the power of the mirrors would separate the two situations shown above as belonging to different classes. During optimization however, a transition from one class to another is conceivable.

Unfortunately, the task of each mirror in a reflective projection system is not always well characterized by the sign of the power. Moreover, the sign of the power of a surface can change during optimization, see Fig. (3.6). Due to this effect quite similar systems end up in different classes. A two-dimensional analysis of a six-mirror system is shown in Fig. (3.7). The white region is obstructed and thus a forbidden region. The different gray levels correspond to different classes.

Fig. (3.7) A two-dimensional analysis of a six-mirror system in our class 37+. Horizontally, the curvature of the second surface varies, the curvature of the fourth surface alters vertically. The gray area is the unobstructed domain, wherein the different grayscales correspond to a specific class according to the classification by the sign of the powers of the mirrors.

3.4.2 Classification by the sign of the angles of incidence

The idea behind the classification by the sign of the incidence angles is that a zero incidence angle inevitably leads to obstruction by the adjacent mirrors. The systems with either a positive or a negative incidence angle have therefore distinctly separated different solution domains. To determine the class number of a system, the signs of the incidence angles $\overline{\theta}_i$ of the principal ray are evaluated consecutively at all the reflective surfaces. Our arbitrary convention to compute the class number is

$$C = \sum_{i=1}^{N} a_i 2^{(N-i)}, \tag{3.36}$$

where a_i=0 if the incidence angle $\overline{\theta}_i$ at mirror i is positive and a_i=1 if it is negative. We start from a positive object height and a reduction projection system. The addition of a plus or minus sign to the class number indicates the positive or negative overall magnification of the system.

We recall that the sign convention means that, seen along the principal ray from object to image, the contribution of each surface is a binary 1 if the ray is reflected to the right and otherwise a 0, see Fig. (3.8). The class number is the decimal value of the binary number obtained in this way. We find it useful to divide the unobstructed reflective projection systems in angle of incidence classes.

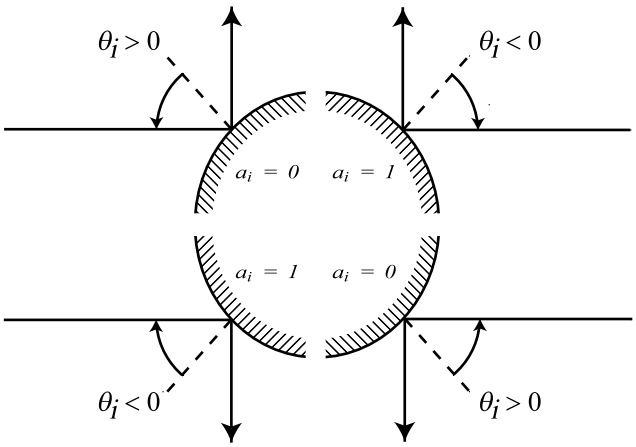

Fig. (3.8) The contribution a_i of each incidence angle to the class number. If the incidence angle θ_i is positive, a_i=0, otherwise a_i=1.

Separated solution spaces of a class

The classification by the angle of incidence is inspired by the existence of small obstruction-free domains in the solution space, see subsection 4.2.1. In most cases, a given class is associated with a single domain in the parameter space and we supposed that systems from the same class belong to the same solution domain. In principle, disjoint domains could exist in the same class; a class can theoretically be split up in two or more unobstructed domains. These separations are caused by obstruction by non-neighboring mirrors that are not included in the analysis of the classification. Fig. (3.9) gives an

example of separate solution domains with eight-mirror systems, which belong to the same class.

When the number of mirrors increases, the risk of obstruction caused by non-neighboring surfaces grows. We encountered a similar situation of separated solution spaces in a six-mirror system. In that case, the largest solution domain is useful and the systems of the disjoint smaller domain are always pushed away from their unobstructed domain by optimization: the obstruction constraint becomes active. Consequently, the potential of these systems to achieve good imaging quality is limited.

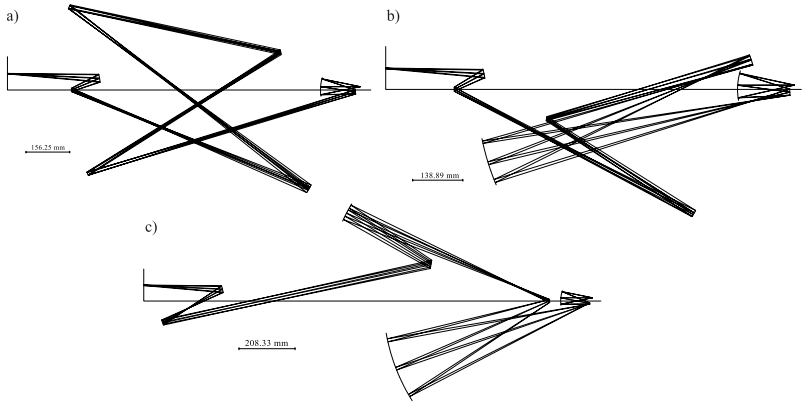

Fig. (3.9) Three eight-mirror systems belonging to the class 153+. The systems in *a* and *b* are in the same obstruction-free domain, whereas the system in *c* is a separate obstruction-free domain. The systems *a* or *b* can not be changed into system *c* without having obstruction.

An extension of the classification method proposed in the previous Section separates systems in different solution spaces. The number of intersections a beam propagating between two reflectors has with other beam segments can not be changed without causing obstruction. When we mark the number of intersections after a reflector, we separate the different solution domains within a class. The systems a) and b) in Fig. (3.9) are for instance $10^20^211^20^201$ and in c) 10011001, where the cases of zero intersections are omitted for clearness.

3.5 Intermediate images

The relation between the transverse magnification M_t and the angular paraxial magnification M_a results from the Lagrange invariant, see Eq. (3.11)

$$M_t M_a = \frac{n_0}{n_N} . \tag{3.37}$$

In this relation, the refractive index n switches sign after each reflection. We can conclude that a system with an even number of mirrors and a negative lateral and angular magnification has an even number of intermediate images. An odd number of intermediate images are found with systems with an even number of mirrors and a positive magnification. In symbolic notations

$$(M_t > 0) \oplus \mathrm{even}(N) = \mathrm{even}(N_i) , \tag{3.38}$$

where the function even() is defined as a logical evaluation and $\oplus$ is the logical *exclusive or* operation. The number of intermediate images is N_i, the number of mirrors in the system N.

The object and image can be considered a special class of intermediate images. In an intermediate image, the intersection heights of the marginal rays are zero and the height of the principal ray is non-zero. Vise versa, the height of the principal ray is zero and the heights of the marginal rays are non-zero at the stop surface, the entrance pupil, the exit pupil and other intersections of the principal ray with the optical axis.

The number of times a principal ray crosses the optical axis, N_c, is closely related to the number of intermediate images N_i

$$N_c = N_i + 1\,. \tag{3.39}$$

In words, the intersections of the principal ray with the optical axis alternate with the intersections of the marginal rays with the optical axis. Between two intersections of a principal ray with the optical axis, the marginal rays cross the optical axis once. Between two intersections of the marginal rays with the optical axis, the principal ray crosses the optical axis once. This relation between the number of intersections is another consequence of the paraxial Lagrange invariant, as the heights of a marginal ray at the two intersection points of the principal ray with the optical axis must be opposite in sign. Just as the heights of the principal ray must be opposite in sign at two subsequent intersections of the marginal rays with the optical axis.

3.6 Scaling

In the paraxial model of reflective systems, the multiplication of specific sets of variables with the same factor leaves certain characteristic properties of the system unchanged. For instance, the obstruction property has a well-defined mathematical structure, characterized by the presence of two such invariants:

1. When all quantities having the dimension length are multiplied with the same factor f_1, the system simply changes its scale. This scale method is often used to adjust the overall length of a system to the desired dimensions. Obviously, for the obstruction function O given by Eq. **(2.8)** we obtain

$$O(h_0, NA, c_1 \ldots c_N, d_0 \ldots d_N) = O(f_1 h_0, NA, \frac{c_1}{f_1} \ldots \frac{c_N}{f_1}, f_1 d_0 \ldots f_1 d_N)\,, \tag{3.40}$$

   where NA denotes the numerical aperture, h_0 the object height, and d_i the distance between surfaces i and i+1 in a N-mirror system.
2. The second invariant is a consequence of the linearity of the paraxial model. When all distances perpendicular to the optical axis can be linearly scaled with a factor f_2

$$O(h_0, NA, c_1 \ldots c_N, d_0 \ldots d_N) = O(f_2\, h_0, f_2\, NA, c_1 \ldots c_N, d_0 \ldots d_N)\,. \tag{3.41}$$

   The obstruction borders are a result of the equality between two (paraxially) computed ray heights

$$y_{j,k} = y'_{j,k'}\,, \tag{3.42}$$

   where j is a surface number and k and k' are 0 or 1 and distinguish between the upper and lower meridional ray, see equation **(2.6)** in the subsection on **obstruction** (page 25). An example of these equalities defining the borderlines of an unobstructed domain is shown in Fig. (4.6). When the ray heights change with the same factor, equation is

unaffected. Consequently, when the values of the numerical aperture NA and the object height h_0 are both multiplied with a factor f_2, the obstruction borders remain unchanged. The paraxial obstruction borders depend therefore only on the numerical aperture to object height ratio.

This scale method changes the angles in the system. The angles have an important influence on the aberrations of a system and on the difference between the paraxial and finite rays. When we decrease the angles in the system, the differences between the paraxial approximation and the finite ray trajectories decrease.

These two invariants can be combined to

$$O(h_0, NA, c_1 \ldots c_N, d_0 \ldots d_N) = O(\frac{f_1}{f_2} h_0, \frac{NA}{f_2}, \frac{c_1}{f_1} \ldots \frac{c_N}{f_1}, f_1\, d_0 \ldots f_1\, d_N)\,. \tag{3.43}$$

Consequently, by choosing in Eq. (3.43)

$$f_1 = \frac{NA\, h_0'}{NA'\, h_0}, \tag{3.44}$$

and

$$f_2 = \frac{NA}{NA'} \tag{3.45}$$

the results of an exhaustive search with a certain numerical aperture NA' and object height h_0' can determine the obstruction properties of other numerical apertures NA and object heights h_0.

3.7 Conclusion

An optical system can be analyzed at various levels of approximation. When the wave-like nature of light can be ignored, i.e. when the surfaces and distances are much larger than the wavelength of the light, geometrical optics can be used. At each transition between different media, the law of Snel predicts the propagation of a ray. For rotationally symmetric systems wherein the rays are close to the optical axis, the well-known paraxial approximation is used; this drastically simplifies the analysis of an optical system.

Various representation methods for surfaces exist, depending, among others, on the number of parameters. The wavefront aberration of a wave reflected by a surface is often described in terms of Seidel aberration coefficients or with the aid of Zernike polynomials. A conic plus an even-order monomial is often used to represent a surface in an optical system. Alternatives include the use of scaled orthonormal sets, e.g. Chebyshev polynomials and their use have proven to be beneficial in optical system optimization.

Systems in a single unobstructed domain can be classified by the angle of incidence. A typical optimization procedure is assumed to be able to transform a system to any other system in the same unobstructed domain, while systems in other unobstructed domains are unreachable. The number of intermediate images and the related number of crossings of the principal ray with the optical axis are other characteristics of an all-reflective optical system; these characteristics lead to an extra subdivision possibility to classify unobstructed mirror systems.

3.8 References

1. F. Pedrotti and L. Pedrotti, *Introduction to optics* (Prentice Hall, 1993).
2. S. A. Lerner and J. M. Sasian, "Optical design with parametrically defined aspheric surfaces," Appl. Opt. **39**(28), 5205-5213 (2000).
3. T. B. Andersen, "Automatic computation of optical aberration coefficients," Appl. Opt. **19**(22), 3800-3816 (1980).
4. T. B. Andersen, "Optical aberration coefficients: FORTRAN subroutines for symmetrical systems," Appl. Opt. **20**(18), 3263-3268 (1981).
5. J. M. Sasian, "Annular surfaces in annular field systems," Opt. Eng. **36**(12), 3401-3403 (1997).
6. M. Born and E. Wolf, "Principles of Optics," 7 ed., Cambridge University Press 64-74 (1999).
7. W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical recipes in C,* second ed. (Cambridge University Press, 1992).

4 Paraxial predesign

Extreme ultraviolet lithography uses reflective ring-field projection systems. The stringent optical performance conditions demanded in the projection system include a root-mean-squared wavefront error smaller than $\lambda/50$ in image space. To correct third- and higher order aberrations, many optimization variables in the form of aspheric surface coefficients are required. It may seem paradoxical, that we want as few variables as possible at the beginning of a system design. Due to the large number of variables and constraints involved, choosing unobstructed starting configurations for subsequent optimization is a nontrivial issue in extreme ultraviolet system design. This first design stage remains often based on existing experience and patent literature, see e.g. the subsection on **existing designs** (page 28). Compared to other types of optical systems, the amount of experience in the field of extreme ultraviolet projection systems is at present still limited.

The paraxial approximation is a simplification of the finite ray-tracing model. By comparing the paraxial and finite ray paths on system drawings, we notice that paraxial ray tracing agrees very well with finite ray tracing for the purpose of e.g. obstruction analysis, especially when the aberrations are small (see Fig. (4.1)). Since the aspheric surface coefficients do not appear in paraxial ray-tracing formulas, the reduction of the number of variables and the mathematical simplification is however substantial.

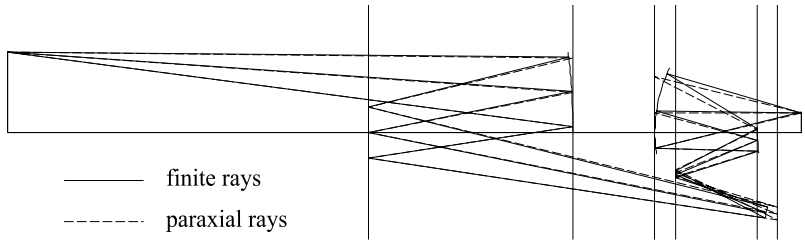

Fig. (4.1) The paraxial rays agree well enough with the finite rays, for a first evaluation of a system, e.g. the presence of obstruction, the workspace, and the telecentricity at the mask and wafer. This example shows a six-mirror system in the class 37+.

The use of mirrors may lead to obstruction. The absence of obstructions is an essential and, as we will show, a very restrictive demand. Geometrical obstruction confines the possible system configurations to small domains of the parameter space. In this chapter, we show the severity of the obstruction-free system requirement by using an analysis method. We also demonstrate how the condition of absence of obstruction can be used to locate all useful starting configurations of systems. The approach generates unobstructed starting configurations in a systematic way, based on paraxial analysis.

4.1 Paraxial model

The extreme ultraviolet projection optics has a large number of design variables: the positions of the object and image plane, the separations between mirrors, the curvatures and the aspherical coefficients describing the surfaces. All these variables influence the numerous image aberrations that must be reduced to very low values for good performance. Initially, the ray propagation through a system is found by applying the basic reflection law

at each ray-surface intersection in an optical system, see Eq. (3.2). The multilayer-covered surface is replaced by a single transition between two media.

For determining an unobstructed *starting point* for a design, we use the limiting case of the paraxial approximation. Paraxial analysis has already been applied successfully to extreme ultraviolet design by Lerner et al[1]. When using the paraxial approximation we ignore the aspheric coefficients and consequently drastically reduce the dimensionality of the search space. After this first design stage, known iterative computer software programs are employed for moving the values of the variables from their respective *starting points* toward their respective values at which optical aberrations are minimized[2, 3]. At the same time, all constraints mentioned in a previous section have to be fulfilled.

4.1.1 Constraints

In the paraxial approximation, a system with N mirrors is described by N curvatures, N-1 distances between mirrors, an object distance, and an image distance. Some constraints can be imposed to the system and, consequently, an independent variable is eliminated by each imposed constraint. Constraints that reduce the number of variables are the magnification, the requirement that object and image are conjugated planes, the 'quasi'-telecentricity on the mask side, the rigorous telecentricity on the wafer side, and the Petzval condition. In all-reflective systems the Petzval sum condition

$$H^2 \sum_{i=1}^{N} c_i \frac{n_{i-1} - n_i}{n_{i-1}\, n_i} = 0 \tag{4.1}$$

simplifies to

$$\sum_{i=1}^{N} (-1)^i \, c_i = 0\,, \tag{4.2}$$

where H is the Lagrange invariant, see Eq. (3.10). In the above formulas, the refractive index of the medium after surface i is denoted by n_i and the curvature of surface i by c_i. When imposing the Petzval condition, it is convenient to use the other paraxial constraints to eliminate distances as independent variables instead of curvatures.

4.1.2 Imposing the constraints

Some of the constraints discussed in Section 2.2.4 that a reflective projection system has to fulfill can be imposed on a paraxial matrix model, see Section 3.1. A comparable approach is employed by Howard et al[4-7] for non-rotational symmetric systems, which require a more complex Hamiltonian approach.

We only consider stop surfaces coinciding with one of the reflecting surfaces (excluding the first and last surfaces). Separate stop surfaces only add one distance to the number of variables. The chances for obscuration however increase drastically. We now consider a system with an even number N mirrors and the aperture stop positioned on surface s.
Each transfer and reflection matrix depends on a single variable, either a curvature or a distance. Before we impose any constraints, the model consists of $2N$+1 variables. The order of solving variables (i.e. making them dependent on the other variables) by using the constraint equations is important. In our approach, we start by imposing the Petzval sum condition, see Eq. (4.2). The curvature of the stop surface c_s becomes a function of the other

curvatures. As a result of the telecentricity and the *quasi-telecentricity* constraints, the principal ray propagate as

$$\mathbf{M}\cdot\begin{pmatrix} h \\ \pm NO \end{pmatrix} = \mathbf{T}_N\cdot\mathbf{R}_N\cdot\mathbf{T}_{N-1}\cdot\mathbf{R}_{N-1}\cdot\ldots\cdot\mathbf{T}_s\cdot\mathbf{R}_s\cdot\begin{pmatrix} 0 \\ \bar{u}_s \end{pmatrix} = \begin{pmatrix} M_t\, h \\ 0 \end{pmatrix}, \tag{4.3}$$

where M_t is the transverse paraxial magnification, NO the numerical aperture on the object side, and $\bar{u}_s$ the angle of the principal ray with the optical axis at the aperture stop. The variable of either R_1 or T_1 can be solved to impose the angle of the principal ray in the object space. The other matrices cannot be used for this purpose because they sometimes describe the ray behavior after the stop surface. Note that we locate the aperture stop on a surface with number s, with s equal or in between 2 and N-1. The matrix T_0, whose variable is the object distance, will be used to fix the magnification. Note that our choice of making the angle of the principal ray equal to the numerical aperture on the object side simplifies the formulas. In the image space, the variable belonging either to $T_{N\text{-}1}$ or to R_N can be solved to impose an exit pupil at infinity. Here again the matrices belonging to earlier surfaces are less appropriate, while the last matrix fixes the image plane. Next, we consider one of the extreme meridional rays, which propagates as

$$\mathbf{T}_N\cdot\mathbf{R}_N\cdot\mathbf{T}_{N-1}\cdot\mathbf{R}_{N-1}\cdot\ldots\cdot\mathbf{T}_2\cdot\mathbf{R}_2\cdot\mathbf{T}_1\cdot\mathbf{R}_1\cdot\mathbf{T}_0\cdot\begin{pmatrix} h \\ 0 \end{pmatrix} = \begin{pmatrix} M_t h \\ M_a NO \end{pmatrix} = \begin{pmatrix} M_t h \\ NA \end{pmatrix}, \tag{4.4}$$

where NA is the numerical aperture and M_a the angular magnification, related to the transverse magnification by Eq. (3.37). With the help of this formula the distance in the last transfer matrix T_N can be solved to make the last plane coincide with the image plane. Concluding, by imposing the constraints the number of independent variables in our model now decreases from 2N+1 to 2N-4 variables

4.2 Implementations

The implementation Section demonstrates our use of the paraxial model. With the aid of software developed by us, we look at the consequences of changes in intermediate distances, curvatures, intersection heights of the principal ray etc. in the (constrained) paraxial model. We show results of two-dimensional analyses of the solution space. The two-dimensional analyses using paraxial and finite rays give interesting information on how properties as the wavefront error change in the solution space. With the systematic exhaustive search, we claim to find all useful classes of systems.

4.2.1 Two-dimensional cuts of the solution space

To investigate the multidimensional unobstructed solution space we make two-dimensional cuts through it. Two variables vary between realistic ranges, whereas the other independent variables remain constant. The results are two-dimensional plots in which each pixel represents a given combination of the two changed variables. All these different configurations are evaluated, with paraxial and / or finite ray tracing. Several properties were examined with these two-dimensional cuts:

- the number of obstructions of a beam, with paraxial or finite ray tracing,
- the obstruction of the beam, to compare paraxial and finite ray tracing,
- the wavefront error of systems, with finite ray tracing,

- the difference between the finite and paraxial beam propagation,
- the third-order aberrations,
- the magnitude of the incidence angles, with paraxial or finite ray tracing,
- the class of a system.

In the next subsections, we present some of these two-dimensional analyses of the solution space. With these analyses, we also seek for paraxial characteristics that can be used to locate the best starting point within an unobstructed domain. The original system under evaluation in the examples is shown in Fig. (4.2).

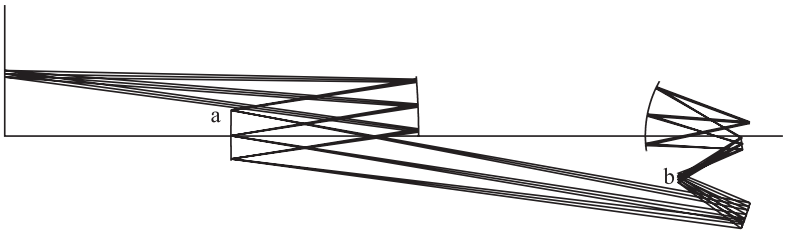

Fig. (4.2) Positive six-mirror system in class 37+. The object heights are between 108 and 120 mm. The ray trajectories correspond to a numerical aperture of 0.3.

Size of solution spaces

To improve the resolution of the projection system, the numerical aperture is as large as the obstruction allows it to be. A typical situation of very limited possibilities to move a mirror is near the image. Here the difference in ray angles with the optical axis relates directly to the numerical aperture and is large, often even the maximum of the system. The last mirror must be close to the image to limit its diameter. The penultimate mirror frequently tends to cause obstruction by cutting the beam between the last mirror and the image, a limiting situation is shown in Fig. (4.3).

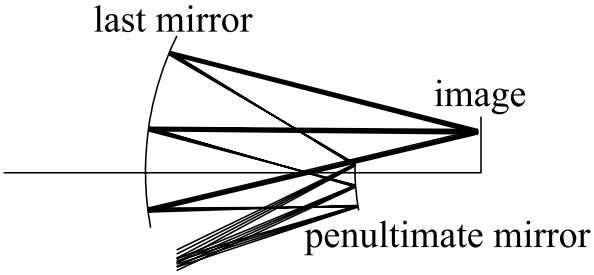

Fig. (4.3) The demand of a large numerical aperture limits the possible mirror positions near the image.

The two-dimensional cross-sections of the solution space show that the consequence of the prohibition of obstructions is a large reduction of the allowed space. The imposed constraints (see Section 4.1.1) drastically further decrease the volume of the unobstructed domains. The size of the unobstructed domain while the number of constraints is increased is shown in Fig. (4.4).

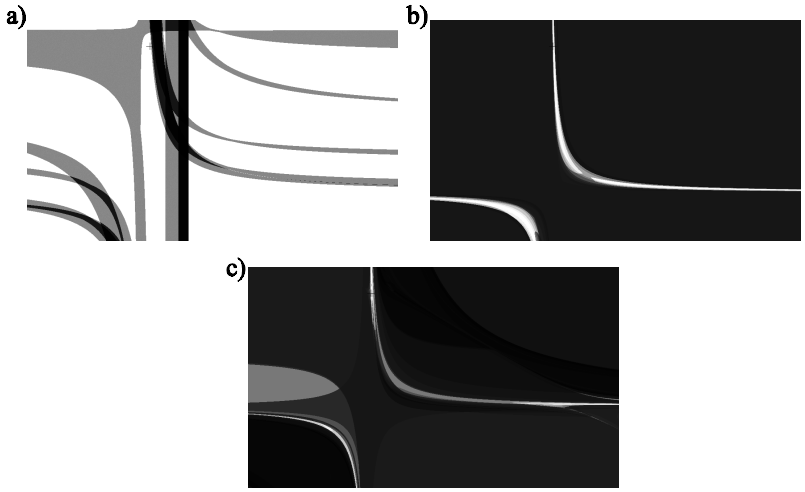

Fig. (4.4) The graylevel indicates the number of obstructions present, starting with unobstructed white regions towards regions with many obstructions in black. In a) no constraints, in b) four constraints, and in c) five constraints are imposed. Increasing the number of constraints imposed on the system reduces the size of the unobstructed domains. The six-mirror system evaluated in this example belongs to the class 37+, see Fig. (4.2). On the horizontal axis the curvature of the third surface and on the vertical axis the curvature of the fifth surface is varied. These two variables are varied within the same domains in the three plots.

Comparison paraxial and finite unobstructed domains

To investigate the differences between paraxial and finite ray tracing, we examine the contours of the obstruction-free domains. Fig. (4.5) shows an example of a two-dimensional analysis of a system. We observe that in general the paraxial unobstructed domains tend to include the smaller finite unobstructed domains. This feature has been observed in many other cases and seems to be typical.

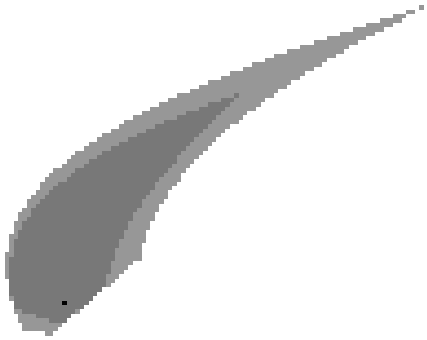

Fig. (4.5) A two-dimensional analysis of the solution space for six-mirror systems in class 37+. The distance between the first and second reflective surfaces varies horizontally. Along the vertical axis, the curvature of the fourth reflective surface varies. The other variables remain constant or are solved by constraints. The lightgray points are paraxially unobstructed, the dark gray points appear unobstructed with both paraxial and finite ray tracing. The black dot marks the original system, see Fig. (4.2).

The obstruction borders are a result of the equality between two (paraxially) computed ray heights

$$y_{j,k} = y'_{j,k'}, \tag{4.5}$$

where j is a surface number and k and k' are 0 or 1 and distinguish between the upper and lower meridional ray, see the subsection on **obstruction** (page 25). An example of these equalities defining the borderlines of an unobstructed domain is shown in Fig. (4.6).

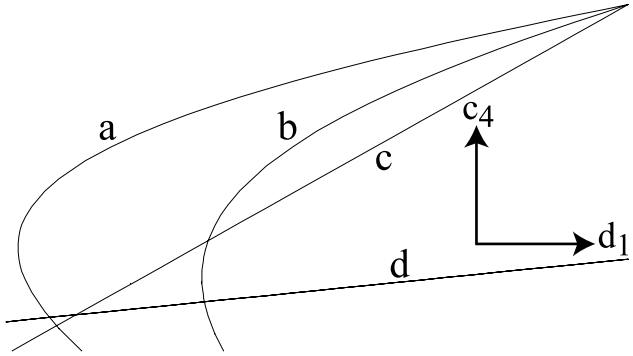

Fig. (4.6) An analytical exploration of the paraxial obstruction borders for six-mirror systems in class 37+. The lines in this Figure are the obstruction boundaries, found with Eq. (4.5). Mirror number 4 obstructs the beam between mirrors number 2 and 3 on lines **a**, **b**, and **d**, see point b in Fig. (4.2). The difference is in the location of the intermediate image: in **a** the intermediate image is after mirror number 4, in **b** the intermediate image is between mirror number 2 and 3, and in **d** the intermediate image is between mirror number 3 and 4. On line **c,** mirror number 2 borders the beam between the object and the first mirror, see point **a** in Fig. (4.2).

Wavefront error

An example of the two-dimensional analysis of the wavefront within the same domain as in Fig. (4.5) is shown in Fig. (4.7). The difference between the evaluation of obstruction with paraxial and finite ray tracing seems to be correlated with the size of the wavefront error. Typically, when the wavefront error is small, the obstruction border found with paraxial and finite ray tracing agree well. The wavefront aberrations vary smoothly and slowly within the domains.

The wavefront aberrations in the finite unobstructed domain show another unexpected feature. In many cases, the equimagnitude lines of the wavefront follow one or more of the obstructed borders. We have also observed that the systems on the line of the minima differ only slightly between them and predominantly rescale.

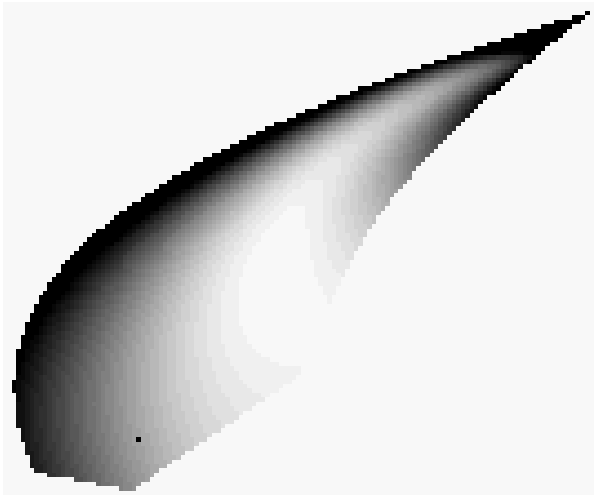

Fig. (4.7) A two-dimensional analysis of the solution space for six-mirror systems in class 37+. The distance between the first and second reflective surfaces varies horizontally. The curvature of the fourth reflective surface changes along the vertical axis. The other variables remain constant or are solved by constraints. The grayscale indicates the wavefront error of unobstructed systems, which varies from small wavefront errors in lightgray to high wavefront errors in black. The difference between the evaluation of the obstruction with paraxial and finite ray tracing is small in the regions with low wavefront errors, compare this Figure with Fig. (4.5).

Maximum numerical aperture

Many optical characteristics are highly nonlinear in the solution space. However, we have observed that these characteristics often vary slowly along lines parallel to the obstruction borders. On these lines, the system itself, as seen on the system drawings, also changes relative slowly compared to variations of the parameters in other directions. One of these characteristics is the maximum numerical aperture at which the system is unobstructed in the paraxial approximation.

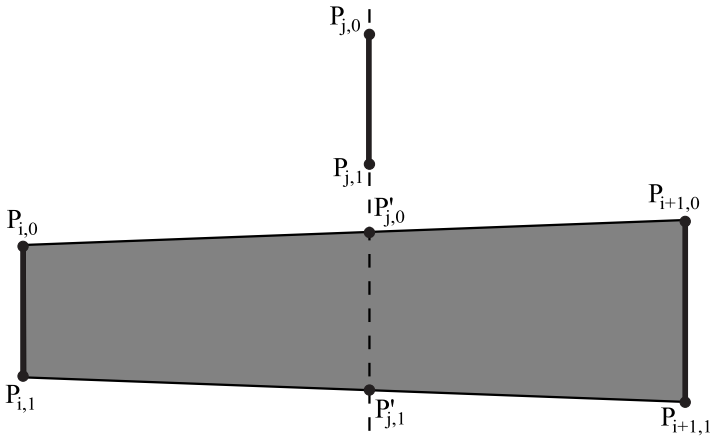

Fig. (4.8) In the Figure the dots are the intersections of an extreme meridional ray with a surface. The surfaces are in the paraxial approximation replaced by their tangent planes. The coordinates of the two points $P'_{j,0}$ and $P'_{j,1}$ are found by solving Eq. **(2.5)**.

The maximum numerical aperture NA_{max} is found by evaluating all combinations of a beam segment between mirror i and i+1 and a mirror j

$$NA_{max} = \mathrm{Min}_{\substack{i=0..N \\ j=1..N \\ j\neq i \\ j\neq i+1}} \left(\frac{\left| y'_{j,0} + y'_{j,1} - y_{i,0} - y_{i,1} \right|}{\left| y'_{j,0} - y'_{j,1} \right| + \left| y_{i,0} - y_{i,1} \right|} \right). \tag{4.6}$$

All systems are scaled along the optical axis to a fixed overall length before calculating the maximum numerical aperture. In Fig. (4.9), the two-dimensional behavior of the maximum numerical aperture is shown. One of the contours with an equal maximum numerical aperture is identical with the paraxial obstruction border. The other equal magnitude lines have resembling but scaled shapes.

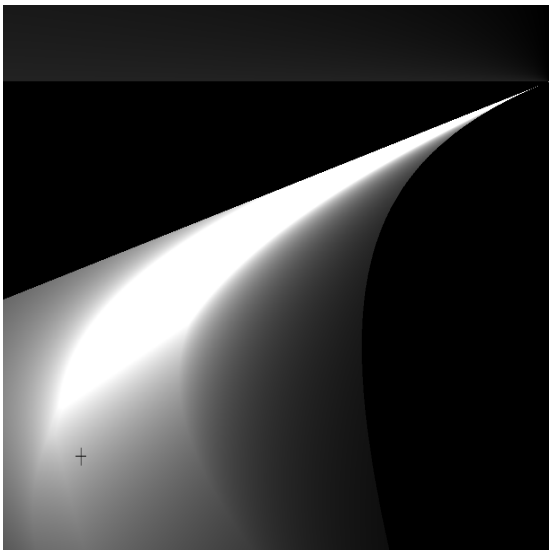

Fig. (4.9) The systems in this two-dimensional analysis are the same as those in Fig. (4.5) and Fig. (4.7). The grayscale indicates the maximum numerical aperture of the system without obstruction. The scale goes from always obstructed in black (zero values of maximum numerical aperture) to high values of the maximum achievable numerical aperture in white.

Scale

The difference between the shapes of the unobstructed domains with finite and paraxial ray tracing varies when the system is scaled perpendicular to the optical axis, see Section 3.6. The paraxial characteristics of the system remain unchanged, as for instance the occurrence of obstruction. A smaller numerical aperture and object height decrease all incidence angles and therefore the validity of the paraxial approximation improves. The shapes of the finite obstruction borders with different object heights and numerical heights remain similar, see Fig. (4.10).

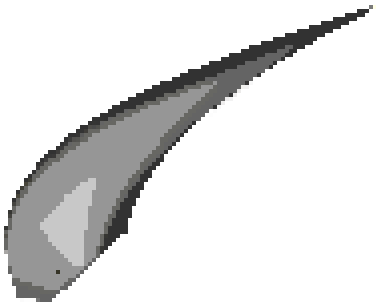

Fig. (4.10) The systems in this two-dimensional analysis are the same as those in Fig. (4.5), Fig. (4.7), and Fig. (4.9). The darkest region is the paraxial unobstructed region. The other domains are unobstructed domains with finite ray tracing. The object height and the numerical aperture have increasing values (dark gray, middle gray, and light gray), but their ratio and therefore the paraxial obstruction domain remain unchanged. The area of the finite unobstructed domain decreases when the object height and numerical aperture increase.

Difference between finite and paraxial ray tracing

The two-dimensional analysis of the wavefront error shows that we have a relation between the validity of the paraxial approximation and the wavefront error. The obstruction borders evaluated with paraxial and finite ray tracing match better for systems with lower wavefront errors. To examine the relation between the wavefront error and the correctness of the paraxial approximation the agreement of the paraxial and finite beam propagation is quantified. All intersections of the two extreme meridional rays and the principal ray with the surfaces in the system are evaluated. The paraxial and finite rays have different intersections with a surface. The total difference B is defined as the sum of the square of all distances between the intersection of a paraxial ray P_p and the intersection of a finite ray P_f with all surfaces in the system

$$B = \sum_{i=1}^{3} \sum_{s=1}^{N} \left(\mathbf{P}_{f,i,s} - \mathbf{P}_{p,i,s} \right)^2 , \tag{4.7}$$

where i indicates a ray and s a surface number, see Fig. (4.11).

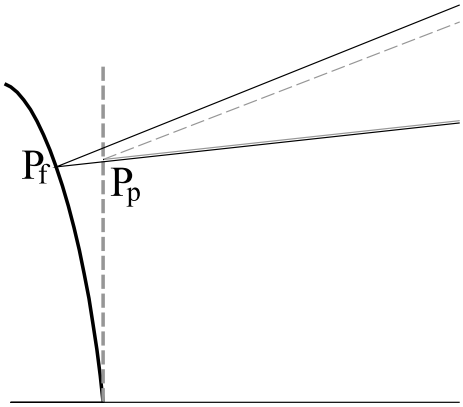

Fig. (4.11) The segments $P_f P_p$ between the intersection point of the paraxial ray (P_p) and the intersection of a finite ray (P_f) with a surface quantifies the difference between the paraxial and the finite trajectory of the beam through the system.

The behavior of the difference *B* between the paraxial and finite beam propagation shows similarities with the wavefront error, compare Fig. (4.7) with Fig. (4.12).

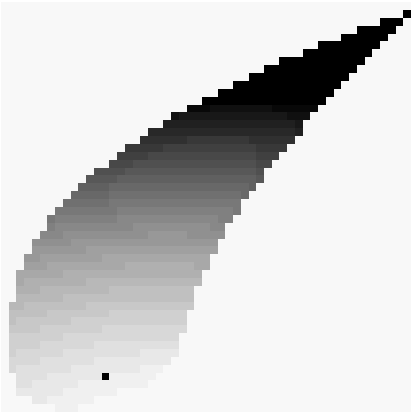

Fig. (4.12) The difference between the paraxial and finite beam propagation through the system is small in the light gray region and increases within the darker regions. The systems correspond to the systems in Fig. (4.5), Fig. (4.7), Fig. (4.9), and Fig. (4.10).

Angles of incidence

The reflectors in the projection systems consist of multilayers. One of the demands of multilayers is a low incidence angle, as is discussed in Section 7.2. The low-incidence-angle condition can be evaluated in the paraxial approximation as well. An example of the two-dimensional analysis of the angles of incidence in a system is shown in Fig. (4.13). In these analyses, the characteristic evaluated of a system is the sum of all incidence angles of the principal ray (i=1) and the two extreme meridional rays (i=2,3)

$$\sum_{i=1}^{3}\sum_{s=1}^{N}\theta_{i,s}^{2} \ . \tag{4.8}$$

In this example and also in other two-dimensional analyses the systems with the lowest incidence angles also have the lowest wavefront error, see Fig. (4.7), and the smallest difference between the paraxial and finite ray tracing, see Fig. (4.12).

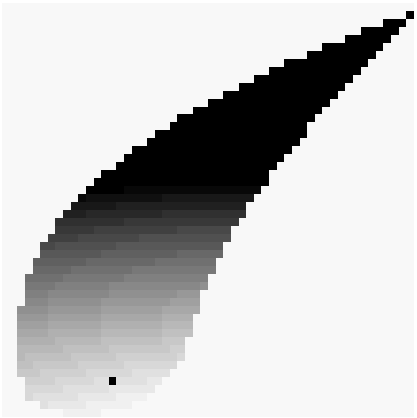

Fig. (4.13) The systems correspond to the systems in Fig. (4.5), Fig. (4.7), Fig. (4.9), Fig. (4.10), and Fig. (4.12). The light gray region has the lowest, the dark gray region the highest incidence angles.

Third-order aberrations

Other characteristics of the systems that can be analysed with the two dimensional analyses are the third order Seidel aberrations. In Fig. (4.14) and Fig. (4.15) the spherical aberration, tangential coma, tangential astigmatism, sagittal astigmatism, Petzval sum, and distortion are evaluated. In Fig. (4.14) all systems have spherical surfaces, whereas the systems in Fig. (4.15) have aspherical surfaces and better imaging properties. By comparing Fig. (4.14 a) and Fig. (4.15 a) we note that the spherical aberration changes sign, but the line of systems with no spherical aberrations remains more or less in the same place. The Petzval curvature is independent of aspherical coefficients and therefore Fig. (4.14 e) and Fig. (4.15 e) are identical.

The black dot marks the location of the optimised system in the Figures. This system corresponds in Fig. (4.15) to the point where all Seidel aberrations are zero.

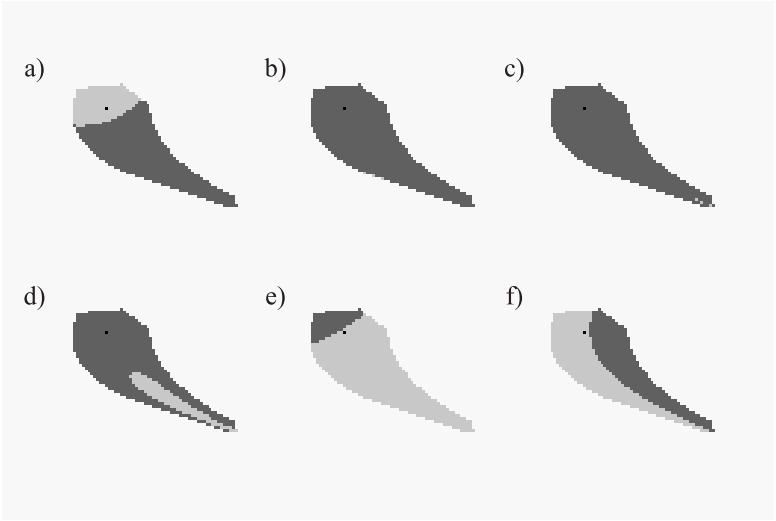

Fig. (4.14) Two-dimensional analysis of the solution space for six-mirror systems in class 45+ with spherical surfaces. The distance between the first and second reflective surfaces varies horizontally. On the vertical axis, the distance between the second and third reflective surfaces changes. The light gray regions have negative and the darker regions have positive aberrations. The separation lines between the light and dark gray are the lines with zero aberrations. Six plots of third-order aberrations are shown here: a) spherical aberration, b) tangential coma, c) tangential astigmatism, d) sagittal astigmatism, e) Petzval sum, and f) distortion. The black dot marks the original system. This point turns out to be close to the intersection of the zero lines for spherical aberration, Petzval sum, and distortion.

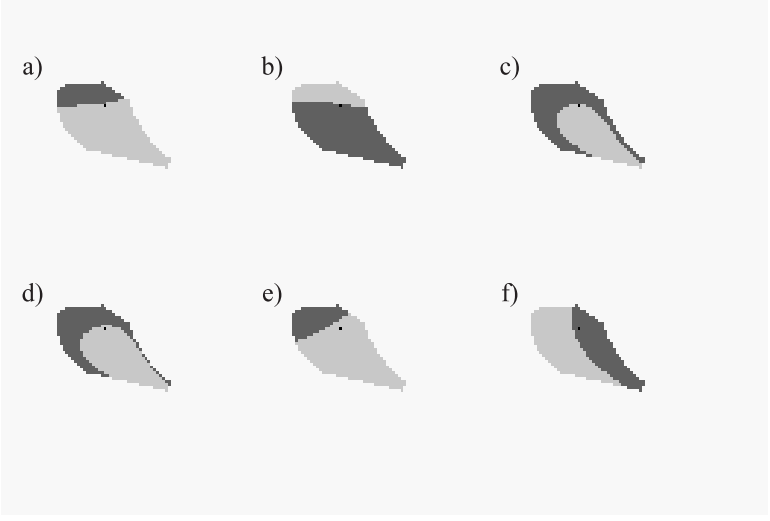

Fig. (4.15) Two-dimensional analysis of the solution space starting with the same system as in Fig. (4.14). The surfaces are in this case aspherical. Six different types of third-order aberrations are represented: a) spherical, b) tangential coma, c) tangential astigmatism, d) sagittal astigmatism, e) Petzval sum, and f) distortion. The black dot marks the original system. This is the only point where all third-order aberrations are close to zero.

4.2.2 Systematic paraxial search

The two-dimensional analysis of the solution spaces shows us the small sizes of the unobstructed domains. These domains are bounded by obscurations and the borders can be found analytically by solving systems of equations that take into account all obscuration possibilities. Here we opt for a more straightforward systematic paraxial search in which all variables vary within realistic ranges. All these configurations are tested for obscuration, workspace, and the feasibility of the system (we eliminate systems for which the solved variables have unrealistic values).

Since the paraxial model with the reduced number of variables is used for these calculations instead of the finite one, we can quickly evaluate many systems. A personal computer is able to check several million paraxial systems per minute. Several computers have been used simultaneously to perform the searches shown in Table (4.1). The evaluation of these twenty billion systems demands approximately a week of computation for each exhaustive search. Obviously, the evaluation of a system with more variables takes even more time. The number υ of configurations to be evaluated increases drastically when the number of surfaces N increases

$$\nu = (N-2)b^{2N+1-q}, \tag{4.9}$$

where b is the number of values tried for each parameter (the sample rate) and q is the number of constraints imposed on the system.

In these systematic searches, only a small fraction of the evaluated systems proves useful (i.e. unobstructed). As an illustration consider a six-mirror system with seven distances and six curvatures, see Table (4.2); eight free variables remain when the constraints are introduced. One of the searches for systems with the stop on the second, third, fourth, and fifth surface and thirty steps for each of the eight independent variables leads to the evaluation of $4 \cdot 30^8 \approx 2.6 \cdot 10^{12}$ configurations. For a numerical aperture to ring-radius ratio at the object side of $^{0.06}/_{116}$ and a transverse magnification of -0.2, only $1.6 \cdot 10^6$ systems fulfil all conditions mentioned. This means that only one system out of $1.6 \cdot 10^6$ evaluated systems is unobstructed in this example of an exhaustive search.

Four-mirror systems

Two searches, one with a negative and one with a positive magnification were done for four-mirror systems. With four constraints, the paraxial model becomes a function of five parameters. Out of the twenty billion systems evaluated only a small part fulfilled the conditions, see Table (4.1).

Table (4.1) The exhaustive searches performed for four-mirror systems. The transverse magnification is M_t, NA is the numerical aperture, and $\bar{u}_0$ is the angle of the principal ray at the object. The column named constraints gives the number of constraints imposed on the system. The number and percentage of paraxial systems fulfilling all requirements are given in the last two columns.

sample rate	M_t	NA	$\bar{u}_0$	constraints	result	%
100	-0.20	0.20	-0.07	4	116888	5.8e-4
100	0.20	0.20	-0.07	4	366423	1.8e-3

The classes of the systems found are shown in Fig. (4.16). Systems belonging to the classes 2-, 6-, 10-, and 9+ are published in the patent literature, see the subsection on **existing designs** (page 28).

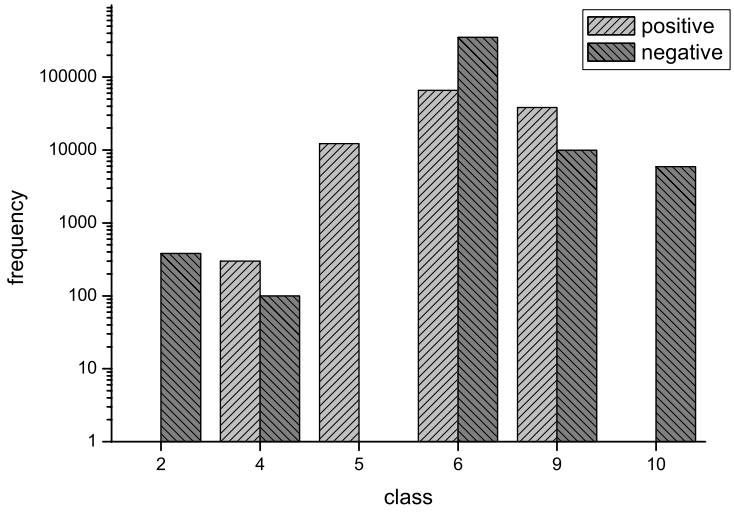

Fig. (4.16) The classes found in paraxial searches for four-mirror systems with positive and negative magnification. The height of the bars corresponds to the paraxial volume of an unobstructed domain.

Six-mirror systems

The first commercial extreme ultraviolet lithographic systems will probably have six-mirror systems. The numerical aperture of these systems will be between 0.24 and 0.35, depending on the quality of the optical design and the compromise made between the slit width and the numerical aperture. Several exhaustive searches for six-mirror systems were done.

An increase of the number of constraints imposed on the system seems to increase the success rate, see Table (4.2). However, this is largely dependent on how the different values of the parameters are chosen and the method to solve the constraints.

Table (4.2) The exhaustive searches performed for six-mirror systems.

sample rate	M_t	NA	$\bar{u}_0$	constraints	Result	%
13	-0.25	0.24	-0.07	4	224	5.3e-7
13	0.25	0.24	-0.07	4	235	5.5e-7
17	0.25	0.24	-0.07	4	970908	2.1e-4
20	0.20	0.30	0.07	5	744649	7.3e-4
20	0.20	0.30	-0.07	5	184995	1.8e-4
20	-0.20	0.30	0.07	5	401220	3.9e-4
30	-0.20	0.30	-0.07	5	1600862	6.1e-5

The *quasi-telecentricity* constraint at the wafer demands an absolute value of the incidence angle $\bar{u}_0$ of the chief ray on the mask that is slightly larger than the numerical aperture on the object side. Most of the times, $\bar{u}_0$ at the mask is negative. This simplifies the combination of the condensor optics with the projection optics. Nevertheless, for some designs a positive sign of $\bar{u}_0$ at the wafer could be preferable. Exhaustive searches with both configurations were done for six-mirror systems with positive (see Fig. (4.17)) and negative

magnification (see Fig. (4.18)). These searches produced the same classes. The configurations with positive and negative incidence angles exist in the same unobstructed solution space, since the condensor optics is not considered when evaluating the presence of obstruction. In the first stage of an optimization, a system can switch the sign of the incidence angle of the principal ray at the mask, $\bar{u}_0$.

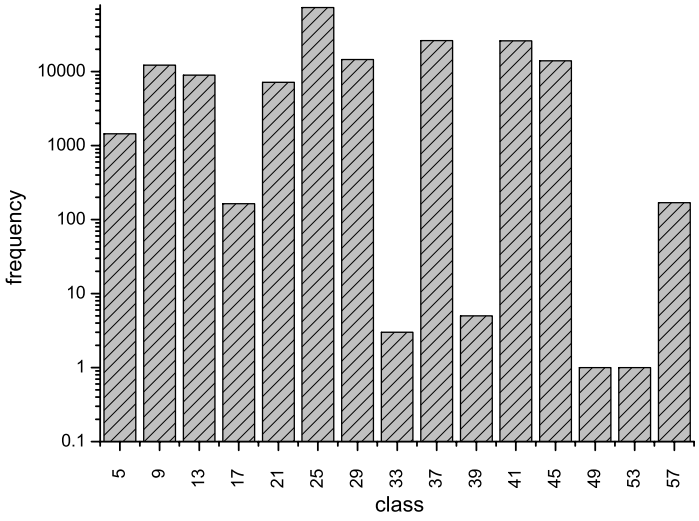

Fig. (4.17) The classes found in a paraxial search for six-mirror systems with a positive magnification. Systems belonging to the classes 41+ and 45+ are published in the patent literature.

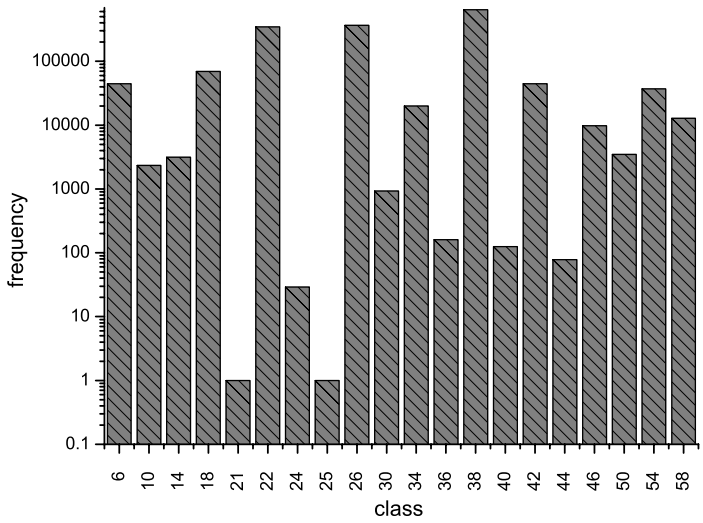

Fig. (4.18) The paraxial results of a search for six-mirror systems with a negative magnification. The patent literature contains systems in the classes 26-

Eight-mirror systems

An eight-mirror system with five constraints and eight different values per parameter demands the evaluation of more than four hundred billion configurations, see Eq. (4.9). Although this sample rate of eight seems small, approximately twenty times more systems have to be evaluated than in the case of a sample rate of a hundred in a four-mirror system with four constraints. With the numerical aperture to object height ratio used in the exhaustive searches, we can control the relative number of systems found without obstruction. Obtaining too many systems is cumbersome, obtaining too few increases the risk of overlooking classes. The searches for four- and six-mirror systems were done with the intended numerical aperture. However, the searches for eight-mirror systems were done with a slightly decreased numerical aperture. Then, due to the lower numerical aperture, the ratio between unobstructed and obstructed systems increases. The results of the different exhaustive searches done for eight-mirror systems with positive and negative magnification are presented in Table (4.3), Fig. (4.19), and Fig. (4.20).

Table (4.3) The exhaustive searches performed for eight-mirror systems

sample rate	M_t	NA	$\bar{u}_0$	constraints	Result	%
6	0.24	0.24	-0.07	4	54	6.7e-8
6	-0.20	0.30	-0.07	5	189	1.5e-6
8	0.20	0.30	-0.07	5	11752	2.9e-6

The patent literature mentions only eight-mirror systems in the 165+ and 150- classes. The possibility to use eight-mirror extreme ultraviolet projection systems is probably only recently considered. Therefore, it is reasonable to believe that more eight-mirror systems will be published in the patent literature in the near future.

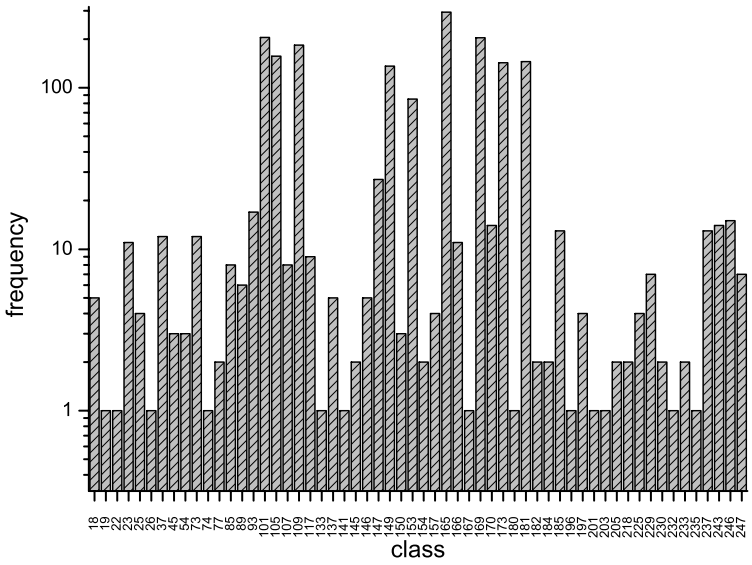

Fig. (4.19) The paraxial results of a search for eight-mirror systems with a positive magnification.

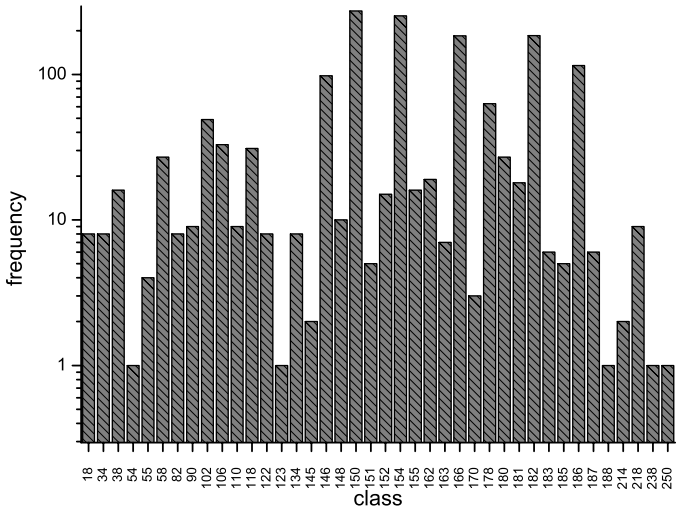

Fig. (4.20) The paraxial results of a search for eight-mirror systems with a negative magnification.

Maskless six-mirror systems

Instead of a reflective mask, an array of very small mirrors with surfaces of 1 μm^2 to 256 μm^2 can be used, see subsection 2.2.3. These arrays of micromirrors are named Spatial Light Modulators (SLM). The projection system of a maskless system images one or more mirrors on a small area with the dimensions of the feature size. The required absolute demagnification is therefore of the order of $^1/_{160}$. This is the demagnification used in the exhaustive searches for maskless projection systems, see Table (4.4). The different classes and the volume of these classes are shown in Fig. (4.21) and Fig. (4.22).

To limit the width of the projection system, the image height is reduced to 3 mm with a corresponding object height of 480 mm. The numerical aperture is nevertheless 0.30, necessitating a strongly curved last mirror and restricting the workspace near the wafer. The image distance and the distance after the penultimate surface were fixed in the searches. The paraxial workspace at the wafer side was chosen to be 4 mm. The stop surface was the penultimate surface.

The large demagnification decreases the numerical aperture on the object side and the bundle diameter drastically. The small bundle diameter combined with the large optical system (with dimensions larger than 480 mm by 240 mm) causes a large ratio of systems to be paraxially unobstructed in the exhaustive searches.

Table (4.4) The exhaustive searches performed for maskless six-mirror systems.

sample rate	M_t	NA	$\bar{u}_0$	constraints	Result	%
$20^5 \cdot 21^2$	$^1/_{160}$	0.30	-0.06	4+2	3.1e8	21.9
$20^5 \cdot 21^2$	$^{-1}/_{160}$	0.30	-0.06	4+2	3.1e8	21.9

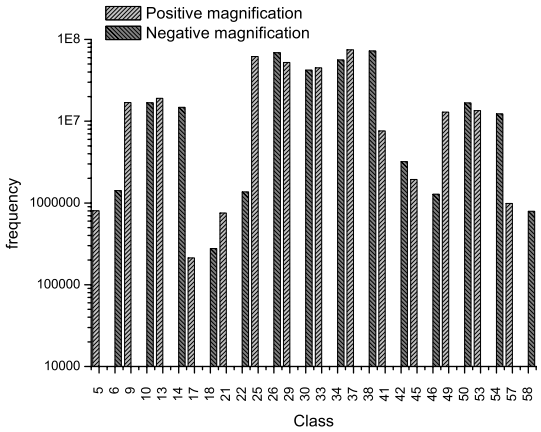

Fig. (4.21) The paraxial results of a search for six-mirror systems with an absolute magnification of 160.

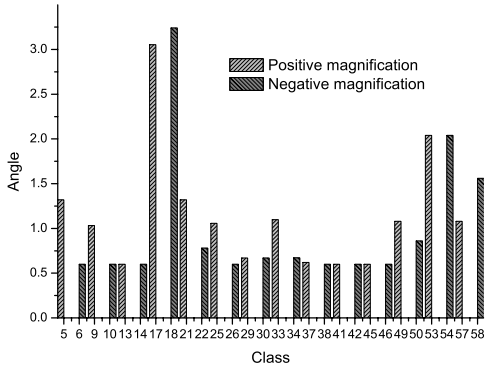

Fig. (4.22) The paraxial evaluation of the minimal value in a class of the maximum angle of incidence in a system. The systems are the same as in Fig. (4.21).

4.3 Conclusion

The chapter started with a discussion of a paraxial model of extreme ultraviolet projection systems where the number of parameters is reduced as compared to real system. The validity of this model in the first design stage was evaluated in two-dimensional cuts through the solution space. For all useful system characteristics, these analyses show a good agreement between paraxial and finite ray tracing. One of these characteristics, which severely limit the usable solution space, is the absence of obstruction. The unobstructed domains found with paraxial ray tracing tend to include the unobstructed domains found with finite ray tracing. Therefore, we can be reasonably sure that useful domains of starting configurations are not lost by choosing paraxial starting configurations

We use this convenient property in exhaustive searches. By evaluating all combinations of different values of the parameters, we claim to find all unobstructed domains. The histograms in Fig. (4.16), Fig. (4.17), Fig. (4.18), Fig. (4.19), and Fig. (4.20) present the different classes found in our exhaustive paraxial searches for systems with four, six or eight mirrors with either positive or negative magnification. The height of the bars in the histograms indicates the volume of a class. This volume is important because, obstruction and constraints are less likely to interfere with optimization for systems in large classes than with those in small classes. Properties as the average angle of incidence, the total track length, a large unobstructed volume for a broad range of numerical aperture to object height ratios, symmetry, the wavefront error of non-optimized systems indicate the potential of classes. All systems published in the patent literature belong to one of the larger classes of projection systems.

4.4 References

1. S. A. Lerner, J. M. Sasian, and M. R. Descour, "Design approach and comparison of projection cameras for EUV lithography," Opt. Eng. **39**(3), 792-802 (2000).
2. OSLO, Lambda Research Corporation, Littleton, 2001.

3. Code V, Optical Research Associates, Pasadena, 2001.
4. J. M. Howard and B. D. Stone, "Imaging a point with two spherical mirrors," J. Opt. Soc. Am. A **15**(12), 3045-3056 (1998).
5. J. M. Howard and B. D. Stone, "Imaging a point to a line with a single spherical mirror," Appl. Opt. **37**(10), 1826-1834 (1998).
6. J. M. Howard and B. D. Stone, "Imaging with three spherical mirrors," Appl. Opt. **39**(19), 3216-3231 (2000).
7. J. M. Howard and B. D. Stone, "Imaging with four spherical mirrors," Appl. Opt. **39**(19), 3232-3242 (2000).

5 Optimization

The performance of an optical system is a function of the parameters of the system. The target of optimization is to find the values of the parameters that result in the best performance of the system. The optical systems are too complex to solve the parameters corresponding to the optimal performance of the system analytically. An error function represents numerically the difference between the actual and desired characteristics of the system. A repeated adjustment of the values of the parameters yields ever lower error function values. Because of the large amount of parameters and the extensive calculations needed, optical design computer programs nowadays perform such an iterative approach.

A typical local optimization algorithm starts to calculate a direction of movement, prescribed according to some algorithm. Algorithms to compute the direction are e.g. Powell's method, the simplex method, the gradient method, the damped-least-squares method, the Newton's method, see e.g. *Handbook of Optics*[1]. The next step is to determine the appropriate step size, either mathematically or by a minimum search along the direction vector.

The majority of the optical design programs use the damped-least-squares optimization algorithm. Other methods used are the orthonormalization method, to overcome stagnation in some damped-least-squares problems. In the next sections, we give a short description of these different optimization algorithms.

5.1 Variables

The variables are those constructional parameters of a system of which the best value is unknown. In the next sections, the vector **x** represents the set of optimization variables. Variables include surface curvatures, thicknesses, refractive indices, aspherical coefficients, etc. Often the values of the variables are restricted, either by physical requirements (e.g. positive thickness) or by the given specifications (e.g. maximum total length or width of the system). These restrictions are called boundary conditions or constraints.

5.2 Error function

In optical design programs, the performance of a system is measured by a single number, known as the merit or error function. The error function is the sum of squares of quantities called operands that characterize different properties of the system. The error function sums the weighted squares of the operands

$$\Phi(\mathbf{x}) = \sum_{i=1}^{m} w_i f_i^2(\mathbf{x}) \,, \tag{5.1}$$

where f_i is the operand number i weighted with a factor w_i in a system. In vector notation, the error function is

$$\Phi(\mathbf{x}) = (\mathbf{w} \cdot \mathbf{f})^T \cdot \mathbf{f} \,, \tag{5.2}$$

where **f** is the set of operands and **w** the set of weighting factors. Examples of operands are aberration coefficients, ray displacements, ray angles, physical realizability, and paraxial properties. These operands are nonlinear functions of the design parameters. The minimum

value of the error function is achieved when all operands are close to zero, although all operands can not typically reach their minimum simultaneously.

To calculate the performance of an optical system a suitable pattern of rays in the entrance pupil is used for each field point. All these rays are then traced to the exit pupil or the image plane. The ray displacements in the image plane give the root-mean-squared transverse error. The set of transverse errors for a field point leads to the geometrical spot size. The optical path length differences with respect to the reference sphere in the exit pupil leads to the root-mean-squared wavefront error, that is the wavefront variance. The optical path differences can also be fitted to a set of polynomials, for instance the Zernike polynomials, in order to include specific aberrations in the error function, see Section 3.3. Another method to express the image quality of an optical system is by using the Strehl ratio. The Strehl ratio is defined as the ratio between the maximum intensity of the actual point spread function and the maximum intensity of the aberration-free point spread function.

A trade off between better accuracy and quicker optimization cycles must be made. The accuracy of the error function specification should depend on the expected system performances.

5.3 Constraints

A system with N variables can be seen as a vector in a N-dimensional space. Each active constraint limits the system to a hypersurface and reduces the number of variables by one. Inequalities simply divide the space of variables into an allowed and a forbidden region. Usually, both the operands and the constraints are non-linear functions of the variables.

Two different methods exist to handle constraints in optical design programs. The first method adds one or more operands to the error function when constraints are violated. These extra constraint operands then express the difference between actual and desired values. In the case of boundary violation, a fixed value or a special function with a magnitude that depends on how far the system is in the forbidden region is used for the operand; otherwise the constraint operand has zero value. The second method enforces the conditions exactly, using the Lagrange multiplier method[2].

The first method is more flexible and faster than the latter method. On the other hand, the latter method gives more precise control over the constraints, which are then exactly satisfied. Both methods are commonly used in optical design software.

5.4 Local optimization

Modern optical design software automatically changes the variables to improve the optical performance of a system represented by the error function. From a mathematical point of view, a system with N-variables is a point in a N-dimensional space. Including the error function adds another dimensionality; all systems are on a surface with a height given by the error function. In a local optimization, the error function is repeatedly decreased in small steps along a declining path, until the lowest point is found. The location of the initial system determines the minimum found in an optimization and the set-up of the resulting system looks typically roughly similar to the set-up of the starting point. The variable

vector $\mathbf{x}'$ represents a system in a minimum if for all points $\mathbf{x}$ in the neighborhood of $\mathbf{x}'$, the error function fulfills

$$f(\mathbf{x}') < f(\mathbf{x})\,. \tag{5.3}$$

The minimum is a stationary point and not an extreme created by constraints when

$$\nabla f(\mathbf{x}) = 0\,. \tag{5.4}$$

An optimization routine can also end up on a plateau, where the error function remains almost constant. Truncation or roundoff errors in the mathematical routines of the optimization can produce the same effect. A minimum differs from a plateau by the property

$$\nabla^2 f(\mathbf{x}) > 0\,. \tag{5.5}$$

The global minimum is the local minimum with the lowest value of the error function. The different types of minima are shown in Fig. (5.1).

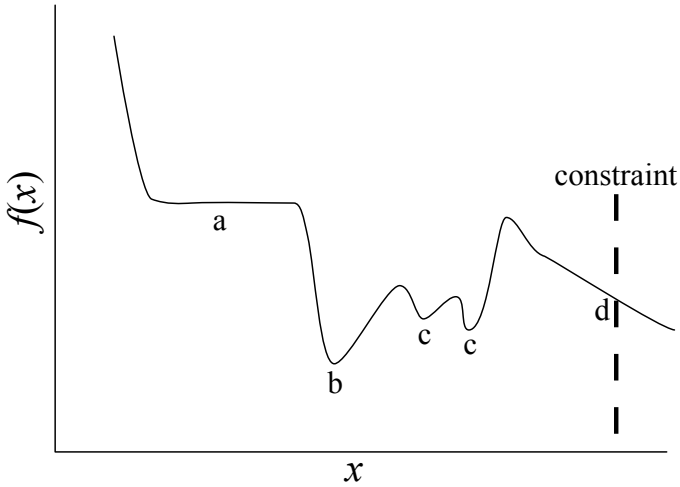

Fig. (5.1) Local optimization can end up in different types of minima. In the Figure *a* is a plateau, *b* is a global minimum, the points *c* are local minima, and *d* is an extreme minimum. The global minimum is the system with the best performance according to the error function.

5.4.1 Damped-least-squares method

The damped-least-squares or Levenberg-Marquardt method is a local optimization method[3]. A typical optimization has more operands than variables, therefore in general no solution exists for which all the operands are simultaneously zero. The state of the system for which the operands are collectively as close to their targets as possible is a least squares solution based on Eq. (5.1). The first order expansion of the operands are

$$f_i(x_j + \Delta x_j) = f_i(x_j) + \Delta x_j A_{i,j} + O(2)\,, \tag{5.6}$$

$$\mathbf{f}' = \mathbf{f} + \mathbf{\Delta x} \cdot \mathbf{A} + O(2)\,, \tag{5.7}$$

where $\mathbf{\Delta x}$ is the change vector of the variables. The matrix $\mathbf{A}$ consists of the derivatives of all the operands with respect to each of the variables

$$A_{i,j} = \frac{\partial f_i}{\partial x_j}, \tag{5.8}$$

$$\mathbf{A} = (\nabla \cdot \mathbf{f}^T)^T . \tag{5.9}$$

When the derivatives are zero (Eq. (5.4)), the error function has a minimum. The non-linearity of the operands limits the range of validity of Eq. (5.6) and Eq. (5.7). A damping factor μ or matrix $\boldsymbol{\mu}$ is added to decrease the length of the variable changes $\boldsymbol{\Delta x}$ and to accelerate the convergence of the optimization[4]. The change of the variables $\boldsymbol{\Delta x}$ is, as a function of the derivative matrix $\mathbf{A}$ and the operands $\mathbf{f}$, a solution of

$$\mathbf{A}^T \mathbf{A}\, \boldsymbol{\Delta x} = -\mathbf{A}^T \mathbf{f} . \tag{5.10}$$

The penalty for larger step sizes that is the damping factor or matrix can now be included in Eq. (5.10)

$$(\mathbf{A}^T \mathbf{A} + \mu \mathbf{I})\boldsymbol{\Delta x} = (\mathbf{A}^T \mathbf{A} + \boldsymbol{\mu})\boldsymbol{\Delta x} = -\mathbf{A}^T \mathbf{f} . \tag{5.11}$$

The terms along the diagonal of the damping matrix penalize the step size of the concerning operand and influence the convergence and direction of the optimization algorithm[5]. Sometimes these damping terms are made proportional to the second derivative terms. The damping has also been used to increase the magnitude of the change vector to escape from a local minimum. The location of the minimum remains nevertheless unchanged by the choice of the damping matrix.

5.5 Global optimization

Most local optical optimization routines like damped least squares and simplex methods are deterministic algorithms. Many of the global optimization techniques contain stochastic elements, e.g. simulated annealing, while some are deterministic, for instance the escape method, see Fig. (5.2). Stochastic methods typically need more evaluations of the system, whereas the deterministic methods have more problems to locate a global minimum.

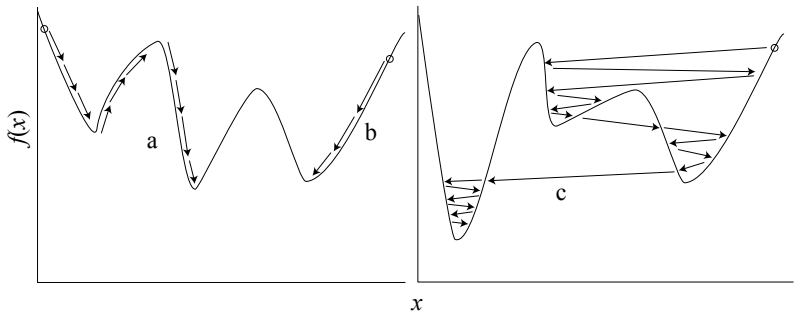

Fig. (5.2) The Figure shows a difference between deterministic and stochastic algorithms: a deterministic method mathematically determines and follows the direction of evolution (e.g. of steepest descent) of the error function. Stochastic methods arbitrarily choose a direction. In the Figure *a* shows a global deterministic, *b* a local deterministic and *c* a global stochastic optimization.

The minimum found with local optimizing methods depends heavily on the configuration of the initial design. In the absence of an analytic solution, the question always remains if the minimum found is the global optimum.

Several global optimization methods are actually extended local deterministic optimization routines. The downhill optimization methods are combined with algorithms that allow the solution to move away from the current local minimum. Examples include the use of the damping factor of the damped-least-squares method and the global explorer algorithm.

The grid search belongs to another group of global optimization methods. All the points on a regularly spaced grid in the specified domains of the variables are evaluated by the error function. The point with the lowest error function serves as the starting point of a local optimization. This kind of approaches becomes however rapidly unworkable with an increasing number of variables.

Stochastic global optimization methods use a controlled random search, called Monte Carlo method. The steps of the variables are randomly, as a function of several parameters, chosen by the algorithm. An example is the simulated annealing routine.

5.5.1 Global explorer

The global explorer algorithm is based on a local optimization routine[6, 7]. Each time the local optimization algorithm finds a minimum, an *escape function* is added to the error function, see Fig. (5.3), to remove a local minimum from its basin of attraction. Reiteration should ultimately lead to the complete set of local minima.

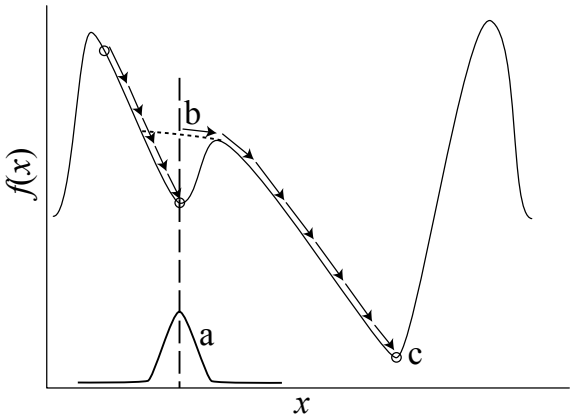

Fig. (5.3) The global explorer uses a downhill optimization. After a local minimum is found, an escape function (*a* in the Figure) is added to the error function. The local minimum disappears (*b* in the Figure) and the downhill optimization method searches for the next local minima. In the best case, all local minima and therefore the global minimum are found successively with the global explorer.

The error function with an escape function is defined by

$$\Phi_E = \Phi + \sum_i \sqrt{\beta_{i,1}} e^{\frac{-\sum_{j=1}^{m} w_j \left(x_j - x'_{i,j}\right)^2}{2\beta_{i,2}^2}}, \tag{5.12}$$

where $\beta_{i,1}$ determines the height and $\beta_{i,2}$ the width of the escape function of the i^{th} local minimum. The choice of the values of these two parameters is difficult and is found by trial and error. If the local optimization routine comes up with a local minimum similar to a previous minimum, two scenarios are possible: all local minima of the error function were

found or the values of the parameters were inadequately chosen. In the latter situation the two parameters $\beta_{i,1}$ and $\beta_{i,2}$ are adapted.

5.5.2 Simulated annealing

The concept of the simulated annealing optimization algorithms is based on the manner in which solids recrystalize into a minimum energy crystalline structure. A melt is in the beginning at high temperature and disordered. If the temperature decreases slowly enough, the system remains at any time in the thermodynamic equilibrium. The system improves its order and reaches a ground state below some critical temperature value. A too quick cooling or a too low initial temperature of the system leads to a crystal structure in a metastable state or with defects.

In the simulated annealing optimization algorithm, the energy corresponds to the error function of the optical system. The magnitude of the stochastic variations of the variables decreases gradually during the optimization as a function of the temperature. The error function is recomputed after each change of the variables. A decrease of the error function is always accepted. The temperature regulates the maximum allowed relative increase of the error function. If the *temperature* decreases sufficiently slowly, this stochastic optimization method is expected to escape from the higher valleys and to terminate at the global minimum. However if the temperature of the optimization process decreases too quickly the system risks to be trapped in a local minimum.

In simulated annealing the temperature is in general given by the user, in contrast to adaptive simulated annealing algorithms where an algorithm reduces the temperature. Other difficulties include the proper sampling of all variables, since the variables have different units and affect the error function in different ways and non-linearly.

5.6 Optimization of EUV systems

This Section describes how the paraxial starting points are optimized to, ultimately, diffraction-limited projection systems. In the next chapter we illustrate the possibilities of our search method by presenting search results for four-, six-, and eight-mirror systems.

5.6.1 Transition from paraxial to finite ray tracing and optimization

If in the paraxial calculations all distances perpendicular to the optical axis and all angles of the rays with the optical axis are simultaneously scaled with the same factor, the obstruction borders are not affected (see the first scale method in Section 3.6). Consequently only the numerical aperture to field ratio matters, when we calculate properties as for instance obstruction borders and workspace. However, a proportional decrease of the numerical aperture and object heights reduces the difference between the paraxial and the finite rays and also reduces the wavefront error. The finite system performs better, more similarly to the paraxial approximation, while the paraxial system only scales along the vertical y-axis.

Many of the systems resulting from a paraxial exhaustive search are obstructed or even impossible, when finite ray tracing is attempted directly. This happens for instance when the ray misses a curved surface, because the ray height at the surface is too large, or when mirror surfaces intersect each other, because the distance between them is too small. As

mentioned before, the difference between the finite and the paraxial rays in a well-corrected system is small. Those systems that even without adjustment are also unobstructed with finite ray tracing belong mainly to the classes of the known systems from literature. Some unknown but large classes contain a convex first mirror (counted from the mask).

In order to adjust those systems where the paraxial ray tracing diverges seriously from finite ray tracing, we tried several conversion methods:

- the optimization is used with an error function in which the operands are all distances between the finite and paraxial intersection points of rays with the surfaces. Most of the times, after a few optimizations, the finite ray tracing indeed converges towards the paraxial ray tracing. With those systems, we continued optimization with a traditional error function.
- the reduction of the numerical aperture and object heights while keeping their ratio fixed is used to decrease the difference between the paraxial and the finite rays. The intermediate systems of this technique are first optimized. The numerical aperture and object heights are later scaled to larger or the original values. This method avoids obstructed or even impossible systems and ray-failure when rays miss a curved surface.
- the numerical aperture is decreased to emphasize the lower-order aberrations in the optimization. The numerical aperture is later increased. A disadvantage is the risk to create obstructions while increasing the numerical aperture back again.
- the error function contains only third order Seidel aberrations, as these are computed with paraxial ray data and do not suffer from ray failures.

5.6.2 Standard optimisation

For a standard optimisation we initially use the transverse aberrations. In the final stage of an optimisation the transverse aberrations are replaced by the wavefront aberrations corresponding to several relevant field-points. Other constraints included comprise the quasi telecentricity near the mask, the telecentricity near the wafer, the free working space, and the absence of obstruction. The evolution of the error function during several stages of optimization is shown in Fig. (5.4) and Fig. (5.5). In point B of Fig. (5.4) and Fig. (5.5) the optimization converges to a system where the gradient of the error function is almost zero. Probably, this is an example of the presence of saddlepoints in the solution domain of these optical systems[8].

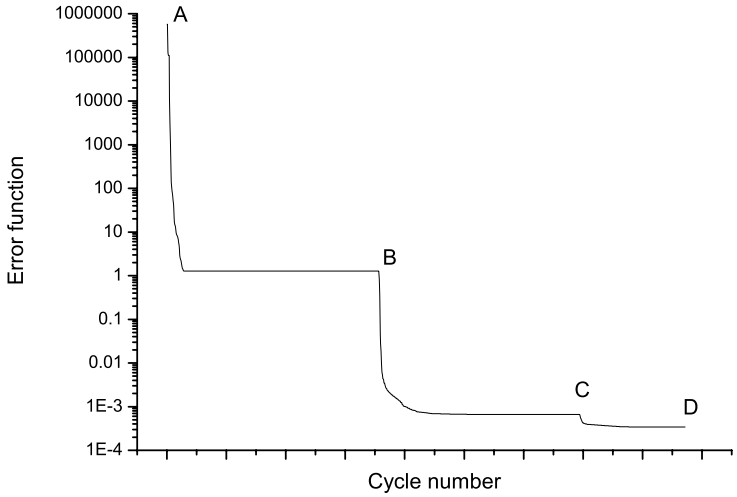

Fig. (5.4) The evolution of the error function during a typical optimization of an optical system. On the vertical axis, the value of the error function is plotted as a function of the cycle number on the horizontal axis. The system under optimization is an eight-mirror system in the class 169+. In point A, the system has spherical mirrors and a root-mean-squared wavefront error of 3860. The optimization reduces the error function quickly to the level of point B and remains constant. In point B, exactly the same optimization is started as in point A. In point C, the set of variables is extended with the 10^{th} order aspherical coefficients. Point D is the final configuration.

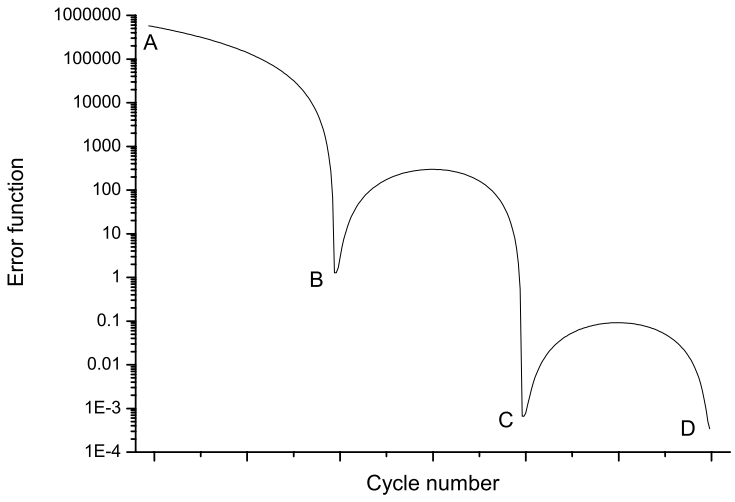

Fig. (5.5) The development of the error function along a straight line in the parameter space between the systems in points A, B, C, D of Fig. (5.4). The non-linearity of the solution space increases near the minimum of the error function.

5.6.3 Problems encountered during optimization

One of the problems that occurs during the optimization of a reflective projection system is the absence of an intersection of a ray with a surface. Also, sometimes the optical design program uses the wrong intersection point of the ray with the surface. Each sphere has at most two intersection points with a ray. Normally the intersection point needed is the one closest to the previous surface.

A method is to take the intersection with the shortest distance to the previous intersection of the ray with a surface. Another method, is to use the distance to the surface intersection of the previous surface with the optical axis, the so-called pole of the surface, which leads in extreme situations with strongly curved surfaces to the wrong selection of the intersection, see Fig. (5.6).

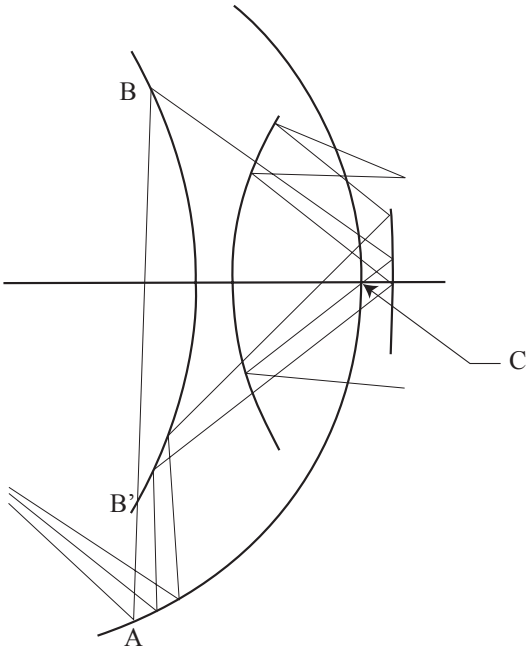

Fig. (5.6) One of the causes of ray failure in optical design programs is the erroneous choice of the intersection point of a ray with a surface. The choice for intersection B or B' depends on whether the distance is calculated from the precious intersection of the ray with a surface (point A) or from the pole of the previous surface (point C). The distance AB' is shorter than the distance AB, while the distance CB is shorter than the distance CB'.

Although some starting points were reasonably well behaved, the optimization of extreme ultraviolet systems remains laborious. In general, the performance of a system obtained with general optimization cycles remains far from diffraction limited, even if resembling systems exist in literature. This is especially the case for six-mirror systems. In Fig. (5.7), the evolution of the root means square wavefront error during optimization of some six-mirror systems is shown. The graphs suggest the presence of local minima, flat plateaus or saddlepoints that hinders the optimization.

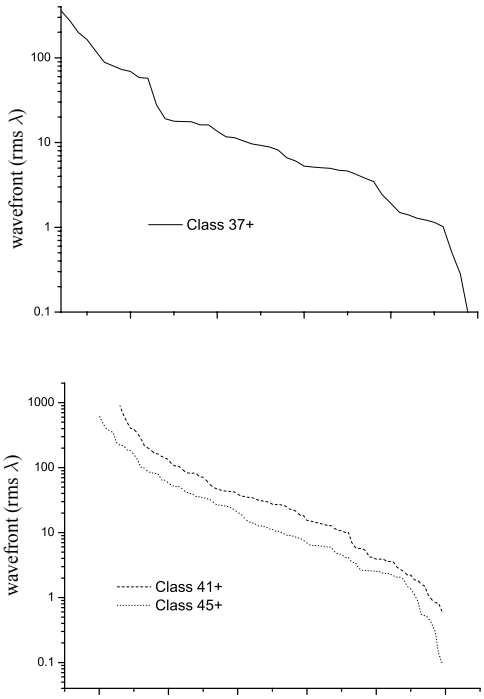

Fig. (5.7) The course of the wavefront error of six-mirror systems during optimization. The horizontal parts in the graphs suggest the presence of local minimum and / or stagnation of the optimization routine.

The convergence of the optimization can be disturbed by truncation errors. The different operands can have opposite dependence on variables or sets of variables. Changing or even reducing the set of variables often proves to be a good remedy for a stagnated optimization process.

The implementation of the optimization procedures in the various optical design programs differs: some programs generate a list of operands the user can see and adjust individually[9] while other programs work with predefined overall error functions[10]. The first type of programs demands more input and adjustments from the user, while the second type of programs works more independently. To deal with the convergence problems mentioned above, we prefer the more open algorithms and programs.

A plausible cause of the convergence problems is the optimization of the aspherical coefficients in a system. The aspherical coefficients are not dimensionless. Also, the higher order coefficients can become very small although their influence on the optical performance is still important. The small values of the aspherical coefficients could lead to mathematical inaccuracy and cause truncation errors.

A final system needs many variables to achieve the demanded optical performance. These variables can be aspherical coefficients. Including orders of aspherical coefficients which do not improve the system performance, can also create convergence problems of the optimization.

Several global optimization algorithms were tried, without much success. Probably the unobstructed solution domains are too small in the solution space defined by a large number of variables.

5.7 Conclusion

In this chapter, we discussed several optimization techniques utilized in optical design programs to improve the optical performance of systems. Optimization algorithms repeatedly make fine adjustments to the parameters in order to find a minimum. In local optimization algorithms, the direction of each change is done deterministically. Global optimization algorithms are either based on deterministic or stochastic methods; we discussed examples of both types.

By using the methods described in chapter 4, we obtained a collection of systems in different classes for each type of configuration. The characteristics of these systems were evaluated with the paraxial approximation. The difference between the finite and paraxial ray tracing is not always small enough to make direct optimization of these systems possible. Different methods were discussed that enable the transition to finite ray tracing and optimization. One of these methods is for instance to scale the numerical aperture and object heights to lower values, do a few optimization cycles, and subsequently to restore the original values of the numerical aperture and object heights. The optimization of systems towards diffraction limited systems with minimal distortion is done with local optimization algorithms.

5.8 References

1. *Handbook of optics,* 2 ed. (Mc Graw Hill, Belfast, 1995), Vol. 1.
2. M. Born and E. Wolf, "Principles of Optics," 7 ed., Cambridge University Press 64-74 (1999).
3. M. J. Kidger, "Use of the Levenberg-Marquardt (damped least-squares) optimization method in lens design," Optical Engineering **32**(8), 1731-1739 (1993).
4. D. Sinclair, *Optics Reference,* 6.1 ed., OSLO, Optics Software for Layout and Optimization (Lambda Research Corporation, Littleton, MA, 2001).
5. H. Matsui and K. Tanaka, "Determination method of an intial damping factor in the damped least squares problem," Applied Optics **33**(13), 2411-2418 (1994).
6. M. Isshiki, H. Ono, K. Hiraga, J. Ishikawa, and S. Nakadate, "Lens design: Global optimization with escape function," Optical Review **2**(6), 463-470 (1995).
7. M. Isshiki, H. Ono, and S. Nakadate, "Lens Design - an Attempt to Use Escape Function as a Tool in Global Optimization," Optical Review **2**(1), 47-51 (1995).
8. A. Serebriakov, F. Bociort, and J. Braat, Delft (personal communication, 2002).
9. Code V, Optical Research Associates, Pasadena, 2001.
10. OSLO, Lambda Research Corporation, Littleton, 2001.

6 New EUV projection system designs

The exhaustive searches presented in chapter 4 resulted in many new possible types of four-, six-, and eight-mirror systems. Some of these systems were optimized with methods discussed in the previous chapter. This chapter presents examples of systems from different classes and discusses their potential. More systems can be found in the appendix in chapter 1.

Some systems are composed of imaging subsystems, which are separated by an intermediate image. One of these subsystems is a three-mirror system with unit magnification, designed by Offner[1]. The three mirrors in the Offner relay have their centers of curvature coinciding at a single point, see Fig. (6.1). Consequently, this combination is free from spherical aberration, coma, and distortion. When the first and third mirrors have twice the radius of curvature of the convex second mirror, the system is also free from third order astigmatism and field curvature. The first and last concave mirrors can be realized by a single mirror surface.

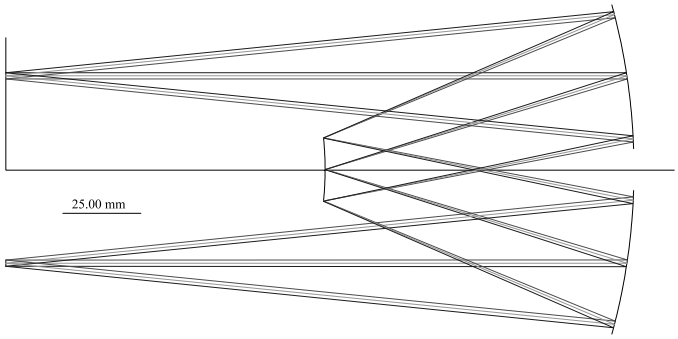

Fig. (6.1) A three-mirror system designed by Offner with unit magnification.

6.1 Four-mirror designs

Many patent publications discuss four-mirror systems. Most of these four-mirror systems belong to the 10- class[2-11] . Some four-mirror systems presented by Dinger[12] fall in the classes 2-, 6-, and 10-.

The four-mirror systems with a class number lower than eight, have a convex first mirror (that is the first mirror after the mask), see Fig. (6.2). A first convex mirror results in a large second mirror located further away from the optical axis than the object. A disadvantage of this kind of arrangement of the first two mirrors is that the illumination coming from the condensor system passes through the projection system. Combining the condensor system (especially the last mirror of the condensor system) with such an projection system becomes more difficult. The main advantage is however that these systems with a first convex mirror have in general large unobstructed solution spaces.

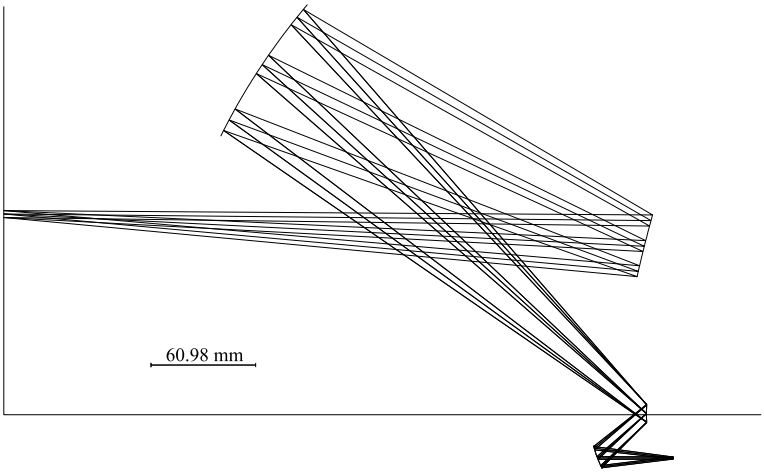

Fig. (6.2) This four-mirror system serves as an example for the class 2-.

Although most of the systems we present in this chapter need more development before they can be implemented in an extreme ultraviolet lithographic system, some of them fulfil the specifications. The system shown in Fig. (6.3) for instance has a Strehl ratio larger than 0.972, and a maximum distortion of 12 nm in a slit width of 1 mm on the wafer side. The optical performances of the system shown in Fig. (6.3) are given in Table (6.1).

Table (6.1) The optical performance of our four-mirror system belonging to class 6-.

object height	*Strehl ratio*	rms wavefront
-22.8	0.972	0.027
-23.0	0.982	0.021
-23.2	0.983	0.021
-23.4	0.984	0.020
-23.6	0.984	0.020
-23.8	0.975	0.025

Other characteristics of the system are the numerical aperture of 0.15 and a magnification of -0.2. The numerical aperture and the slit are both relatively large; the numerical aperture is in fact the largest of all systems found in the patent literature, see the subsection on **existing designs** (page 28).

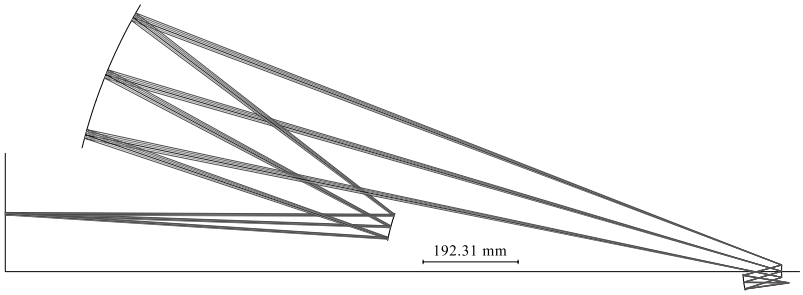

Fig. (6.3) Four-mirror system in class 6- with root-mean-squared wavefront errors below 0.0266 λ and distortion below 12 nm. Designed at a wavelength of 13 nm, this system has a numerical aperture of 0.15, and object heights between 114 and 119 mm.

For comparison, the projection system in the engineering test stand (ETS)[6], see the subsection on **Extreme UltraViolet (EUV)** (page 20), is a four-mirror system in class 10-. The system consists of three aspherical surfaces and a spherical third mirror combined with the aperture stop, see Fig. (6.4). The characteristics of this system are, at a wavelength of 13.4 nm, a numerical aperture of 0.10, a ring width of 1.5 mm on the wafer side and a -0.25 magnification. The distortion of the design is approximately 15 nm and the composite root-mean-squared wavefront error is 0.014 λ.

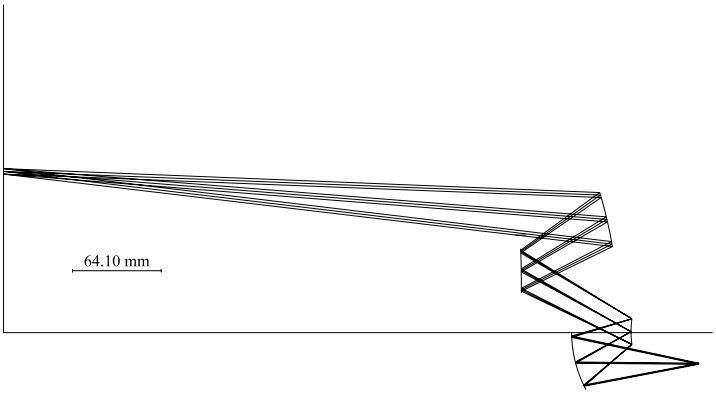

Fig. (6.4) A four-mirror system in class 10-, a system in the same class is used in the Engineering Test Stand.

Of the positive classes, only class 9+ (see Fig. (6.5)) leads to usable finite systems. Lerner et al[13] presented together with other systems a four-mirror system in the 9+ class, called the *symmetric design*. However, the illumination of the mask is in their system far from telecentric.

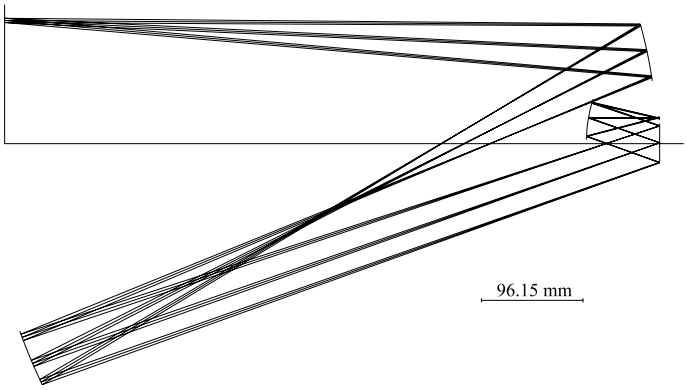

Fig. (6.5) A four-mirror system in class 9+, where in this system the stop is located on the third surface.

6.2 Six-mirror designs

This Section presents six-mirror systems found with the exhaustive searches with positive and negative magnification. The different classes of systems were presented in the

subsection on **six-mirror systems** (page 65). The patent literature mentions systems in the six-mirror classes 26-[14], 41+, 45+[15].

The six-mirror systems with a class number smaller than thirty-two have a convex instead of a concave first mirror, just as the four-mirror systems with a class number smaller than eight. Examples include the four-mirror system in class 6-, the six-mirror system in class 26- and the six-mirror system in class 9+. Relatively large unobstructed domains but small free working spaces are typical for these classes, see Fig. (6.6) and Fig. (6.8). Nevertheless, the class of negative six-mirror systems with the largest paraxial unobstructed domain has a concave first mirror, see Fig. (6.7).

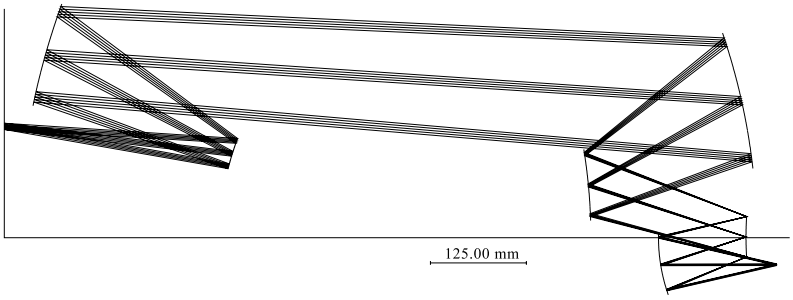

Fig. (6.6) This six-mirror system in the class 26- resembles the four-mirror system in class 6- shown in Fig. (6.3), but has an additional pair of mirrors in the group on the image side. As in the four-mirror design, the two mirrors nearest to the wafer almost cause obstruction. However, the additional pair of mirrors enables a larger numerical aperture.

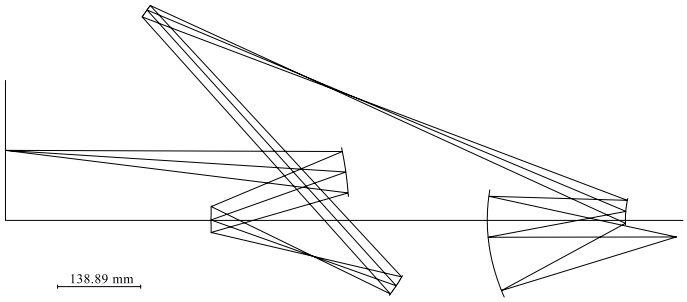

Fig. (6.7) The class 38- has the largest unobstructed domain of all six-mirror classes with a negative magnification. This example also has two intermediate images and tree intersections of the principal ray with the optical axis.

Sometimes unexpected design possibilities emerge from the paraxial searches. An example of an exotic design is shown in Fig. (6.8). These designs show the remarkable flexibility of the obstruction borders, despite the severely limited design space. Large incidence angles make some new unobstructed classes less attractive, since large angles make aberration correction more difficult and the reflectivity of the multilayers smaller.

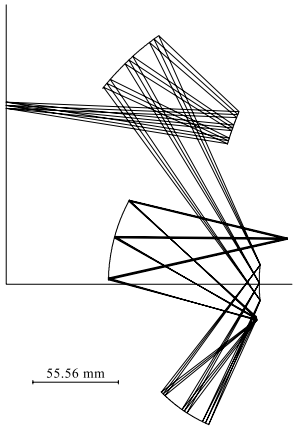

Fig. (6.8) A positive six-mirror system in class 9+. Although the angles of incidence are large, the root-mean-squared wavefront error can decrease down to $\lambda/2$. The object heights are between 114 and 118 mm, the numerical aperture is 0.3 but the distortion is large.

Two well-known six-mirror systems belong to the 41+ and 45+ classes. In class 37+, two mirrors form a group near the object and four mirrors cluster near the image in the systems see Fig. (6.9). Conversely, in class 41+, mirrors cluster near the object and two mirrors near the image, see Fig. (6.10). The intermediate image is in both classes most of the times approximately in between the two groups of mirrors.

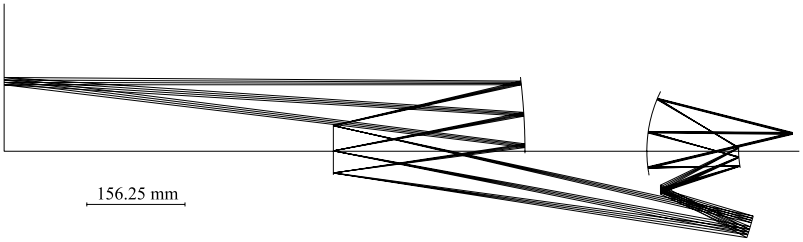

Fig. (6.9) A six-mirror system in class 37+, with the stop on the second surface. The object heights are between 108 and 120 mm, the numerical aperture is 0.3, the root-mean-squared wavefront is below 0.15 λ, but the distortion is large.

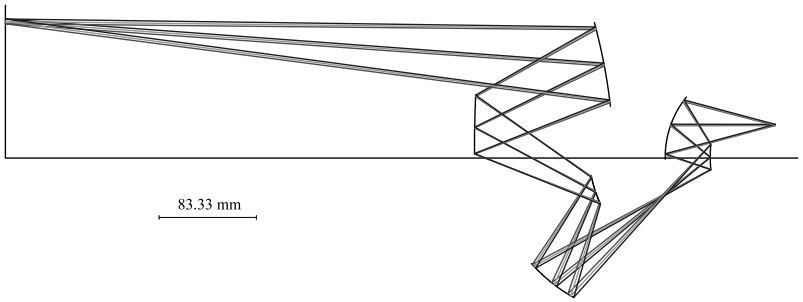

Fig. (6.10) This six-mirror system belongs to the class 41+, just as the systems designed by Braat[16], Hudyma[17], Mann[18], and Shafer[19].

Williamson designed a well-known six-mirror system in class 45+. The system started as a combination of two three-mirror systems designed by Offner[1]. Optimization at the required

magnification of 0.25 resulted in the system shown in Fig. (6.11). The intermediate image is still between the third and fourth mirror.

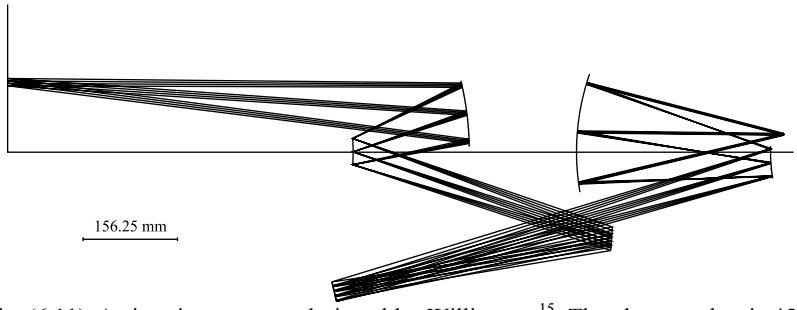

Fig. (6.11) A six-mirror system designed by Williamson[15]. The class number is 45+, the numerical aperture 0.25, with a magnification of 0.25 and a wavelength of 13 nm.

6.3 Eight-mirror designs

This Section shows eight-mirror systems in different classes with negative and positive magnification. Two eight-mirror systems in class 165+ designed by Mann et al[20] and Shafer[19] were found in the patent publications. Another class of eight-mirror systems was represented by a system also designed by Mann et al[20].

In most eight-mirror systems, one or more intermediate images divide subsystems. Some of these subsystems resemble known four- and six-mirror systems. An example is shown in Fig. (6.12), where the first subsystem is an inverted negative six-mirror system in the class 26-. In Fig. (6.13), two four-mirror systems are separated by an intermediate image located between the fourth and fifth mirror. The subsystem near the object is in class 9+, while the subsystem near the image is in class 6-.

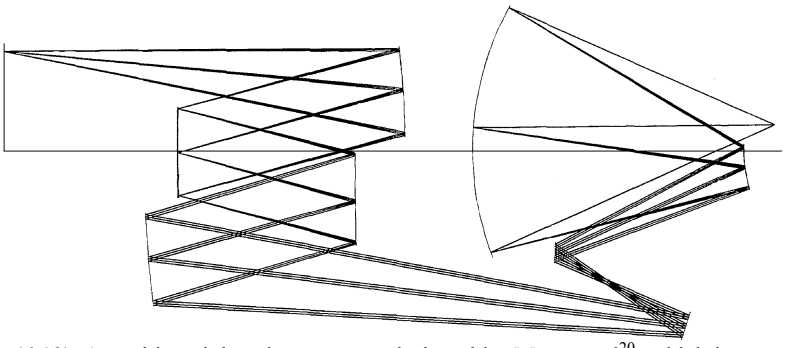

Fig. (6.12) A positive eight-mirror system designed by Mann et al[20], which is part of the eight-mirror class 165+.

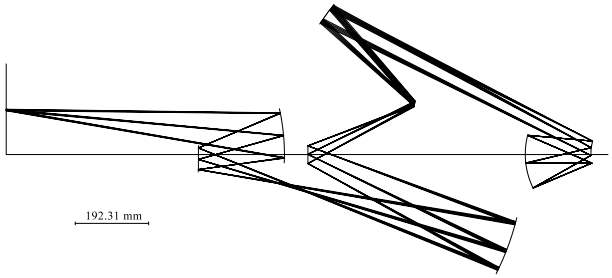

Fig. (6.13) The class 150- includes this negative eight-mirror system, with a distortion smaller than 2.6 nm, object heights between 114 and 118 mm, a numerical aperture of 0.4, and a root-mean-squared wavefront error smaller than 0.19 λ.

Normally, the rays in an extreme ultraviolet projection system have small angles with the optical axis. A curious example of a system with large angles is shown in Fig. (6.14), where some of the rays are almost perpendicular to the optical axis.

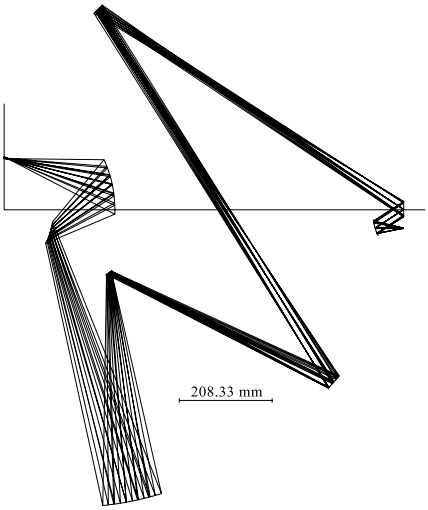

Fig. (6.14) Negative eight-mirror system in class 150-, with a large distortion, object heights between 114 and 118 mm, a numerical aperture of 0.4, and a root-mean-squared wavefront error varying between 1.13 λ and 5 λ within the slit.

The first subsystem in the negative eight-mirror system in Fig. (6.15) is a six-mirror arrangement that has some resemblance with the system designed by Williamson, see Fig. (6.11).

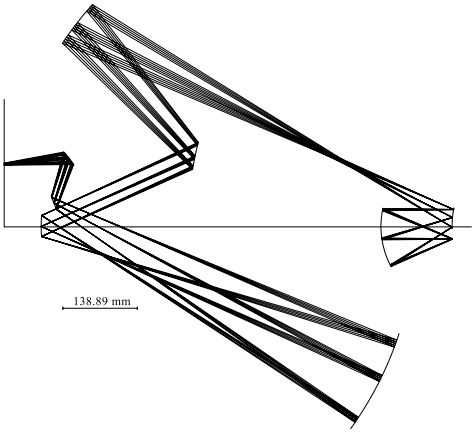

Fig. (6.15) Negative eight-mirror system in the class 182-, object heights between 114 and 118 mm, and a numerical aperture of 0.4. The root-mean-squared wavefront error is smaller than 1 λ, the distortion is smaller than 2.18 nm.

Between each intersection of the principal ray with the optical axis the system forms an intermediate image, see the subsection on **intermediate images** (page 47). The negative eight-mirror systems typically have two intermediate images and cross the principal ray crosses the optical axis three times. The positive eight-mirror systems have in general a single intermediate image and two intersection of the principal ray with the optical axis. An exception of this observation is the positive eight-mirror system shown in Fig. (6.16).

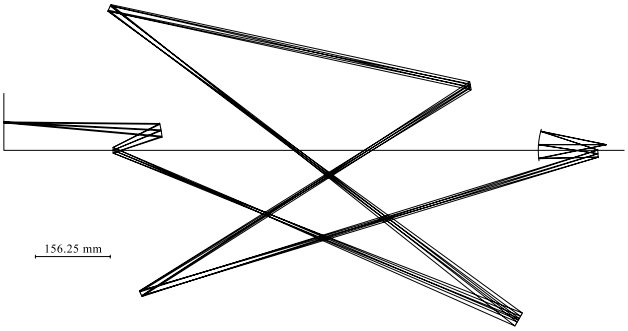

Fig. (6.16) Positive eight-mirror system in the class 153+ with four intersection of the principal ray with the optical axis and three intermediate images.

In previous eight-mirror examples, subsystems could be distinguished, that resemble other four- and six-mirror systems. Instead of a combination of two systems, sometimes a pair of mirrors seems to be added to a design. The positive six-mirror systems in class 41+, see Fig. (6.10), consists of a second group of two mirrors and a first group of four mirrors equivalent to the four-mirror system in class 10-, see Fig. (6.4). In the system in Fig. (6.17) the first group of four mirrors is extended with an additional pair of mirrors.

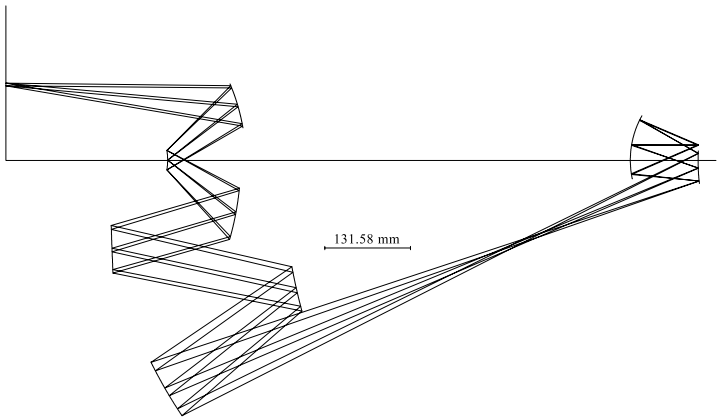

Fig. (6.17) Positive eight-mirror system in the class 169+, with a distortion smaller than 2.9 nm, object heights between 114 and 118 mm, numerical aperture of 0.4, and a root-mean-squared wavefront error smaller than 0.3 λ.

A similar situation is shown in Fig. (6.18), which looks like the Williamson system in Fig. (6.11) with an additional pair of mirrors near the object.

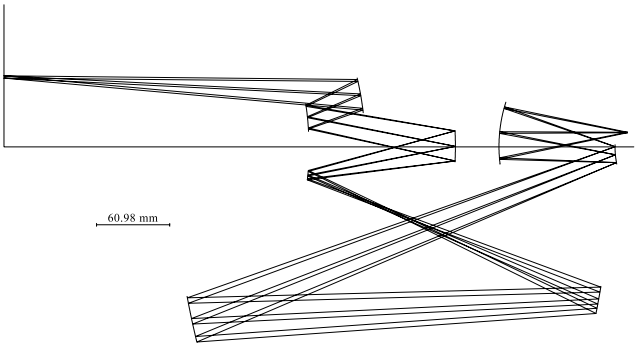

Fig. (6.18) Positive eight-mirror system in the class 173+, with a distortion smaller than 15 nm, object heights between 114 and 118 mm, numerical aperture of 0.4, and a root-mean-squared wavefront error smaller than 0.4 λ.

6.4 Maskless extreme ultraviolet systems

The projection system images a structure of features on the wafer. Both for extreme ultraviolet and for longer wavelengths, a mask typically contains the information of the desired patterns. An alternative method is to use a large array of very small mirrors (10^5-10^8 mirrors), see Section 2.2.3. One of the more promising projection systems we found which could be used in a maskless extreme ultraviolet system is shown in Fig. (6.19).

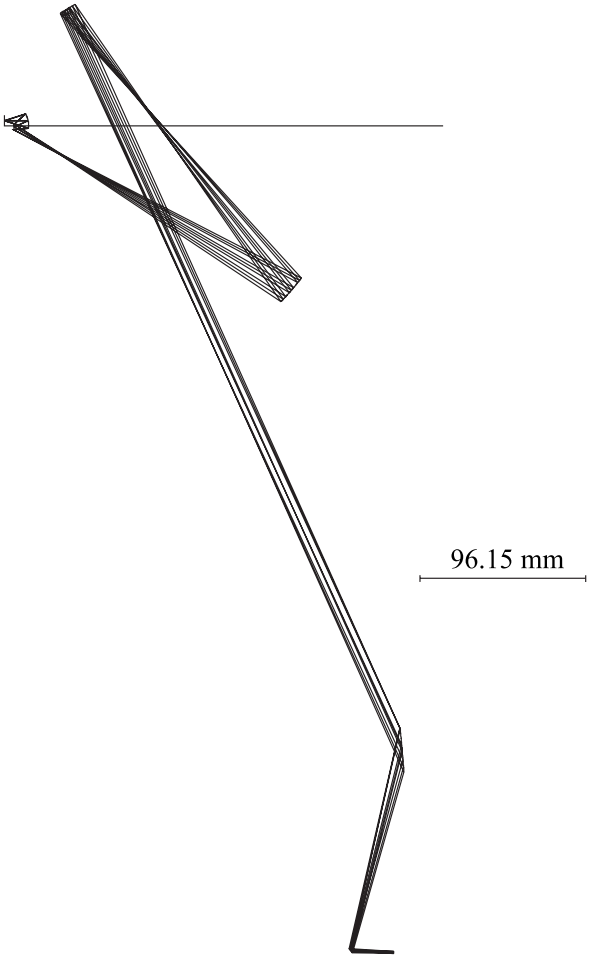

Fig. (6.19) A negative six-mirror system in the class 46-, which could be used in a maskless extreme ultraviolet lithographic system. The root-mean-squared wavefront error is smaller than 0.6 λ, when the numerical aperture is equal or smaller than 0.3. The restriction to the distortion limits the maximum slit width; the image heights are between 2.95 and 3.05 mm. The beam crosses the optical axis three times and has two intermediate images.

6.5 Conclusion

Our four- and six-mirror system searches have found many new and surprising mirror arrangements, but show that the most promising classes are already known. An exception is a four-mirror system with good optical performance and a relatively large numerical aperture we found in class 6-. For eight-mirror systems, some new classes, not found in patent publications, seem very promising. For most of the examples presented, the given values of the root-mean-squared wavefront error and distortion are only the results of a very crude first optimisation attempt. With further design effort, these performances could probably substantially be improved for many of these systems.

6.6 References

1. A. Offner, "Unit power imaging cataptric anastigmat," US patent 3,748,015 (Jul. 24, 1973).
2. D. Williamson, "Four mirror EUV projection optics," US 5,956,192 1997).
3. W. Sweat, "Ringfield lithographic camera," US patent 5,805,365 (Sept. 8, 1998).
4. D. Shafer, "Reflective optical imaging system," US patent 6,014,252 (Jan. 11, 2000).
5. D. Shafer, "Reflective projection system comprising four spherical mirrors," US 5,410,434 1993).
6. T. Jewell and K. Tompson, "X-ray ringfield lithography," US 5,315,629 1991).
7. T. Jewell and J. Rodgers, "Apparatus for semiconductor lithography," US patent 5,063,586 (Nov. 5, 1991).
8. R. Hudyma, "Reflective optical imaging systems with balanced distortion," US patent 6,226,346 (May 1, 2001).
9. S. Cohen, H. Jeong, and D. Shafer, "Four-mirror extreme ultraviolet (EUV) lithography projection system," US patent 6,142,641 (Nov. 7, 2000).
10. H. Chapman, R. Hudyma, D. Shafer, and D. Sweeney, "Reflective optical imaging system with balanced distortion," US patent 5,973,826 (Oct. 26, 1999).
11. J. Bruning, A. Phillips, D. Shafer, and A. White, "Lens system for X-ray projection lithography camera," US patent 5,353,322 (Oct. 4, 1994).
12. U. Dinger, "Ringfeld-4-Spiegelsysteme mit konvexem Primarspiegel fur die EUV-Lithography," European patent EP 0 962 830 A1 (Dec. 8, 1999).
13. S. A. Lerner, J. M. Sasian, and M. R. Descour, "Design approach and comparison of projection cameras for EUV lithography," Opt. Eng. **39**(3), 792-802 (2000).
14. J. Braat, "Lithographic apparatus comprising a dedicated mirror projection system," US 6,396,067 1999).
15. D. Williamson, "High numerical aperture ring field optical reduction system," US patent 5,815,310 (Sep. 29, 1998).
16. J. Braat, "Mirror projection system for a scanning lithographic projection apparatus, and lithographic apparatus comprising such a system," US 6,255,661 1999).
17. R. Hudyma, "High numerical aperture ring field projection system for extreme ultraviolet lithography," US patent 6,318,869 (Nov. 20, 2001).
18. H.-J. Mann and R. Hudyma, "Projection system for EUV lithography," WO patent 02/48796A2 (Jun. 20, 2002).
19. D. Shafer, "Projection lithography system and method using all-reflective optical elements," US 5,686,728 1996).
20. H.-J. Mann, W. Ulrich, and G. Seitz, "8-mirrored microlithographic projector lens," World Intellectual Property Organisation WO patent 02/33467A1 (Apr. 25, 2002).

7 Multilayers in EUV systems

Extreme ultraviolet projection systems consist of reflective optics, as the 13.4 nm illuminating radiation is highly absorbed in all materials. The high reflectivity is obtained with the aid of multilayers, which typically consist of alternating layers of molybdenum and silicon or molybdenum and beryllium. The phase and amplitude of the optical field reflected by a multilayer depend on the angle of incidence and the multilayer characteristics. The phase and amplitude variations produced by multilayers entail important consequences for the imaging properties of the projection system, e.g. the resolution, the depth of focus, and the tolerances.

The optical performance of extreme ultraviolet projection systems is modeled, evaluated, and optimized in optical design programs. The influence of the multilayers on the reflecting properties can be neglected in the early design stage. Once the projection system reaches the diffraction limit, the effect of the multilayers must be included. We propose two different methods to incorporate multilayers in an extreme ultraviolet projection system. The first method is to make the top of the multilayer stack coincide with the single transition surface. In the second method, the so-called *effective reflection depth* coincides with the single transition surface. The effective reflection depth is a fictitious surface, where the multilayer reflection is thought to be concentrated. At the end of this chapter, we present a new method to calculate the spatially varying optimum thickness of multilayers (*grading*). This method makes the numerical optimization of the grading of the multilayers superfluous. Using the new method, we are able to fully optimize the wavefronts exiting from the projection system towards the image plane. The only residual variation present is found in the intensity distribution over the wavefront.

7.1 Thin-film approach

The thin film calculations are implemented in the symbolic mathematical program of Mathematica[1]. The results are systematically checked with the thin-film calculation program TFCalc[2]. This Section starts with a short review of the standard matrix method for thin-film calculations. The matrix method for thin-film calculations assumes an infinitely extending plane wave incident on a multilayer consisting of plane parallel surfaces. The tangential components of both the electric field and the magnetic field must be continuous across all boundaries between the different media

$$E_j^t = E_{i,j}^t + E_{r,j}^t = E_{t,j}^t + E_{r,j+1}'^t \,. \tag{7.1}$$

In this equation and in Fig. (7.1), $E_{i,j}$ marks the incident electric field at transition j, $E_{t,j}$ marks the transmitted electric field, and $E_{r,j}$ marks the reflected electric field at transition j. A comparable relation holds for the magnetic field components. Note that in subsequent equations, E_j marks the tangential field component at the transition with number j.

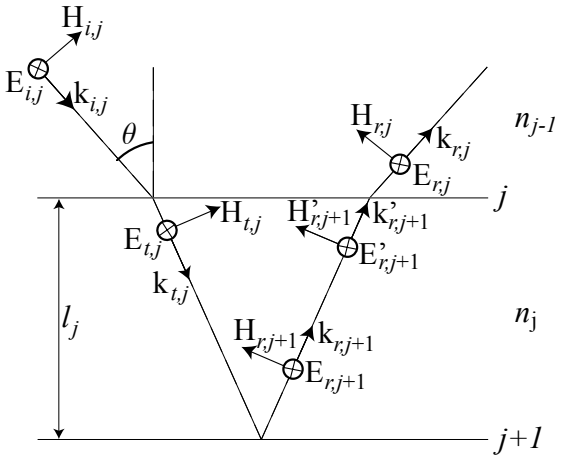

Fig. (7.1) The tangential field components at both sides of a transition with number j in a multilayer are equal. The layer with number j has a thickness l_j, and a refractive index n_j.

The derivation of the following formulas is based on these boundary conditions[3]. The complex refractive indices n of the various materials used are taken from the website presenting the work of Henke et al[4] and shown in Table (7.1). Two orthogonal polarization states are distinguished. In the case of s-polarization, the electric field is perpendicular to the plane of incidence defined by the wavevector $k_{j,1}$ and the normal to each interface. The p-polarization implies that the electric vector is located in the plane of incidence. The two polarization states have to be analyzed separately, by choosing the appropriate expression for γ from Eq. (7.4).

Table (7.1) The complex refractive indices n of the materials used in extreme ultraviolet multilayer stacks.

Material		Optical constants	
Symbol	Name	Re(n)	Im(n)
B	boron	0.9689	-0.0040
Mo	molybdenum	0.9227	-0.0062
Pd	palladium	0.8780	-0.0443
Pt	platinum	0.8928	-0.0577
Rb	rubidium	0.9941	-0.0007
RbCl	rubidium-chloride	0.9895	-0.0019
Rh	rhodium	0.8775	-0.0296
Ru	ruthenium	0.8898	-0.0165
Si	silicon	0.9999	-0.0018
SiO_2	silicon-dioxide	0.9787	-0.0106
Sr	strontium	0.9880	-0.0013
Y	yttrium	0.9742	-0.0023

The characteristic matrix of a layer j relates the tangential components of the incident electric and magnetic fields, respectively E_j and H_j, at a transition between layers j-1 and j to the subsequent transition between layers j and j+1

$$\begin{bmatrix} E_j \\ H_j \end{bmatrix} = \begin{bmatrix} \cos(\delta_j) & \dfrac{i \sin(\delta_j)}{\gamma_j} \\ i\,\gamma_j \sin(\delta_j) & \cos(\delta_j) \end{bmatrix} \cdot \begin{bmatrix} E_{j+1} \\ H_{j+1} \end{bmatrix}, \tag{7.2}$$

where the quantities δ_j, $\gamma_{s,j}$ for s-polarized light and $\gamma_{p,j}$ for p-polarized light of layer j are

$$\delta_j = \frac{2\pi\, l_j\, \sqrt{n_j - \sin^2(\theta)}}{\lambda}, \tag{7.3}$$

$$\gamma_{s,j} = \sqrt{n_j - \sin^2(\theta)}, \quad \gamma_{p,j} = \frac{n_j}{\sqrt{1 - \dfrac{\sin^2(\theta)}{{n_j}^2}}} \quad . \tag{7.4}$$

In these equations, θ is the complex angle of incidence and n_j the complex refractive index of layer j. The thickness l_j of the layers is normally a quarter wavelength for reflection coatings. A scaling factor can adjust this thickness for obliquely incident waves. For non-absorbing materials, the factor typically is the average of the cosine of the angle of incidence θ

$$l_j = \frac{\lambda}{4\,\mathrm{Re}(n_j)\,\overline{\cos(\theta)}}. \tag{7.5}$$

To evaluate the effect of an assembly of q layers on a plane wave we combine all individual characteristic matrices of the layers and of the substrate, with suffix m, in the vector

$$\left(\prod_{j=1}^{q} \begin{bmatrix} \cos(\delta_j) & \dfrac{i \sin(\delta_j)}{\gamma_j} \\ i\,\gamma_j \sin(\delta_j) & \cos(\delta_j) \end{bmatrix} \right) \cdot \begin{bmatrix} 1 \\ \gamma_m \end{bmatrix} = \begin{bmatrix} E_0 \\ H_0 \end{bmatrix}. \tag{7.6}$$

The complex reflectivity r is then given by

$$r = \frac{\gamma_0\, E_0 - H_0}{\gamma_0\, E_0 + H_0}. \tag{7.7}$$

Where the phase shift φ of the reflected field and the reflected intensity R are given by

$$\varphi = \arg(r), \tag{7.8}$$

$$R = r r^*. \tag{7.9}$$

7.2 Typical properties of multilayers

Layer or multilayer coatings are often used as an antireflection coating, an enhanced reflection coating or as a band-pass filter. The characteristics of the coatings are a function of the refractive indices and the thicknesses of the layers. A typical example of a reflective coating is a Bragg reflector. In a Bragg reflector, the optical thicknesses of each layer approximately are a quarter-wave, see Eq. (7.5). In extreme ultraviolet systems, the multilayers often consist of alternating layers of molybdenum and silicon. An alternative for relatively shorter wavelengths is the combination of molybdenum and beryllium. Although the reflectivity of a single transition from molybdenum to silicon or vice versa is very small, the addition of multiple reflections results in a saturation to a maximum reflectivity of up to 70% at about 40 layerpairs, see Fig. (7.2). The maximum intensity of the illumination is limited by the capacity of the system to dissipate the remaining 30% of

incident power by cooling of the mirrors. Otherwise heating of the mirror results in (accelerated) multilayer degradation and contamination.

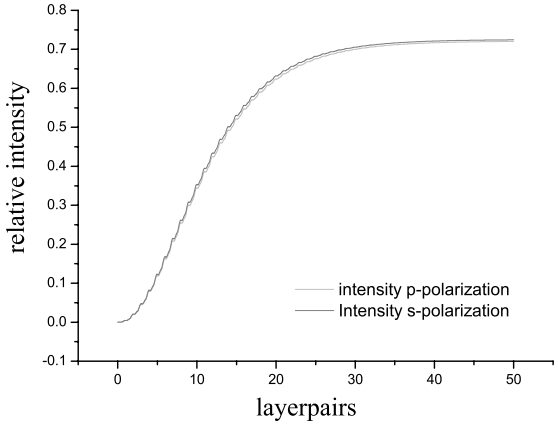

Fig. (7.2) The relative intensity of the field reflected by a multilayer is a continuous and increasing function of the number of layerpairs added at the rear side of a multilayer. In this example, the layerpairs consist of molybdenum and silicon. Note the extremely small difference in reflectivity between s- and p-polarization at near normal (5º) incidence.

A multilayer coating optimally reflects rays incident at a predetermined range of incidence angles. The allowed deviation from the average angle of incidence mainly decreases as a function of the average angle of incidence, see Fig. (7.3). In Fig. (7.4) the allowed variation in the incidence angle is shown as a function of the average incidence angle, when an arbitrarily selected minimum reflectivity of 60% is demanded. At an average incidence angle of 10˚, the reflectivity at normal incidence decreases below the minimum of 60%. The allowed variation shows a quick drop for average incidence angles larger than 10˚.

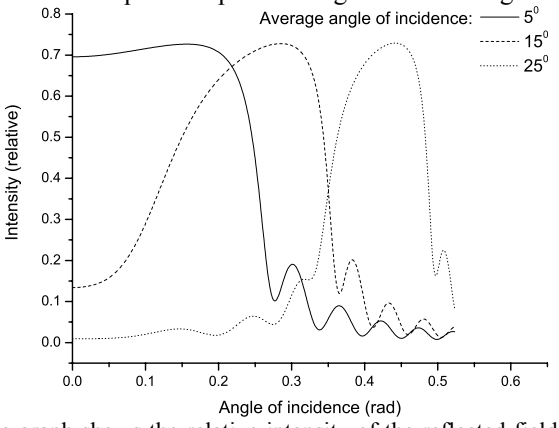

Fig. (7.3) The graph shows the relative intensity of the reflected field as a function of the angle of incidence for multilayers with different average angles of incidence. The usable variation of the angle of incidence decreases for larger average angles of incidence.

In extreme ultraviolet projection systems, the occurrence of both positive and negative incidence angles on the same mirror is very unlikely, as at least the rays with normal

incidence angles are clipped or obstructed. Otherwise, the allowed variation would double for the lower average incidence angles. The graph in Fig. (7.4) leads to the conclusion that the absolute optimal average angle of incidence is approximately 10˚ with a maximum range of variation of 16˚.

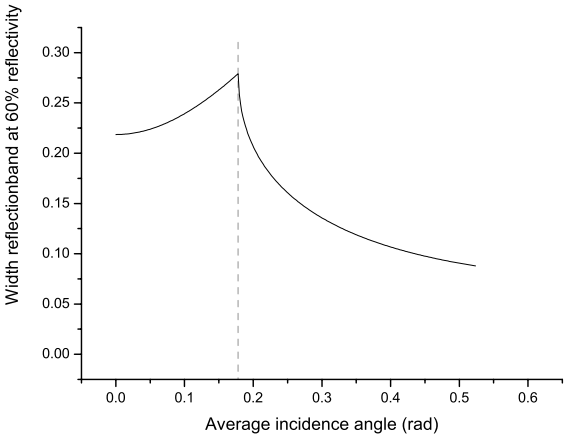

Fig. (7.4) The tolerable variation in the angle of incidence for a minimum reflectivity of 60% is a function of the average angle of incidence. The thicknesses of the layers are changed as a function of the average angle of incidence according to Eq. (7.5). The tolerable range first increases, after which it decreases strongly for average angles of incidence beyond 0.18 rad.

7.3 Enhanced reflectivity

In the previous section, the optical thicknesses of the layers are a quarter wavelength, according to Eq. (7.5). Typical designs of extreme ultraviolet multilayers have eighty layerpairs of molybdenum and beryllium with a maximum theoretical reflectivity of approximately R~0.78 or fifty periods of molybdenum and silicon with a maximum reflectivity of R~0.74. These materials have a high absorption, that is a relatively large complex part of the refractive index. The absorption affects the phase changes, because of which the ideal optical thickness of the layers is not exactly a quarter wave. The contributions from the different transitions in the multilayer are on a bent line, as is shown in Fig. (7.5). In the ideal case, all the contributing reflected fields have the same phase, that is the vectors are on a straight line. The optimization of the thicknesses are done with the aid of the thin-film program[2, 5].

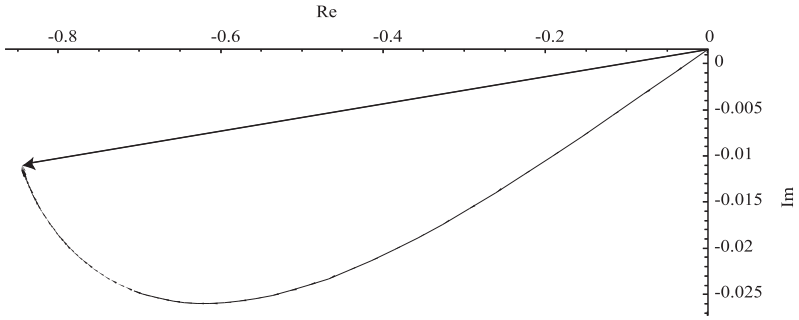

Fig. (7.5) The reflected field is a complex number, of which the real part is plotted along the horizontal axis and the imaginary part along the vertical axis. The vectors in the Figure represent the difference in the complex reflectivity r between a multilayer with N+1 layers and N layers, where the additional layer has been added between the substrate and the multilayer. In the Figure, all these vectors add up to the total reflectivity of a Mo / Si multilayer of 50 layerpairs. The angle of incidence is 0° and the thicknesses of the layers are given by Eq. (7.5).

An extreme ultraviolet lithographic system contains typically nine near-normal incidence mirrors: two reflectors in the illumination optics, six mirrors in the imaging optics plus the reflective reticle. These nine reflectors transmit R^9~0.10 of the intensity of the illumination and absorb the rest. The absorbed energy results in heating and degradation of the multilayers, thus limiting the maximum processing speed of the lithographic system and explaining the benefit of even a slight enhancement of the reflectivity of the multilayers.

Besides the modification of the thicknesses and materials of the layers, other methods to improve the reflectivity of extreme ultraviolet multilayers include the incorporation of additional materials with different complex refractive indices to the stack of the multilayer[6, 7]. The protective capping layers, in addition to protecting the mirrors from environmental attack, may also serve to improve the reflectivity characteristics. The addition of, for instance, boron (B), palladium (Pd), rubidium (Rb), rubidium-chloride (RbCl), rhodium (Rh), ruthenium (Ru), silicon-dioxide (SiO_2) or strontium (Sr) augments the peak reflectivity R with maximally ~5% for a single reflector and increases the optical throughput of nine mirrors by at most a factor of two, see Table (7.2).

Table (7.2) The peak reflectivity R of multilayer stacks based on molybdenum and silicon is enhanced for wavelengths of 13.4 nm by the incorporation of additional materials. The reflection factor after nine reflections is Rel. R^9, normalized with respect to the standard, non-optimized molybdenum / silicon multilayer. These data are taken from an article written by Singh et al.[6].

Multilayer	Capping Layer	R	Rel. R^9
Mo/Si	2 nm Si + 2 nm SiO_2	0.731	1.00
Mo/Si	1.7 nm Ru	0.758	1.39
Ru-Mo/Si	1.5 nm Ru	0.764	1.48
Ru-Mo/Sr-Si	1.5 nm Ru	0.781	1.81
Ru-Mo/Sr-Rb	1.5 nm Ru	0.814	2.63

7.4 Effective reflecting surface

In the case of reflection on a single transition between two different media, the reflectivity and phase shift depend on the angle of incidence θ and the ratio of the refractive indices n of the two media. The origin of the reflected ray is the point of intersection of the incident ray with the transition. The direction of the ray is given by the fact that the angle of reflection is equal to the angle of incidence, see Snel's law in Eq. (3.2). When we include layers or multilayers to an optical system, the reflected field should be approximated as good as possible. In an optical design program, rays can normally not be split into a number of rays, since the increase of the number of rays to evaluate quickly leads to an intractable problem in the case of multiple multilayer surfaces.

When doing thin-film calculations, one assumes an infinitely extending plane wave incident on a multilayer consisting of plane parallel surfaces. The calculations coherently sum all reflections coming from the different transitions in the multilayers. The total reflected field depends on the geometry of the normal to the multilayer surface **v** and the wave vector **k**. In the case of two-dimensional multilayer structures, the superposition of the infinite series of reflected (or transmitted) waves can be represented by a single plane wave with a certain phase φ, measured with respect to the top surface of the multilayer, see Fig. (7.6). The input field $E_{i,0}$ is related to the reflected field $E_{r,0}$ by

$$E_r = |r(k,v)|\, E_0\, e^{i\phi(k,v)}, \tag{7.10}$$

where the absolute value of r is the relative amplitude. The reflection coefficient $|r|$ is defined as

$$|r| = \left|\frac{E_r}{E_0}\right|. \tag{7.11}$$

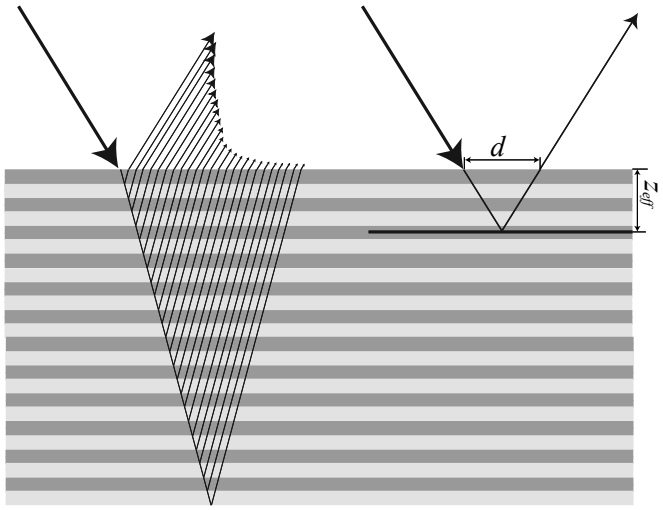

Fig. (7.6) Two approaches for representing the field reflected by a multilayer structure. On the left side, the approximation of reflected fields results in a large collection of rays with different phases and lateral displacements. On the right side the average reflected field has a lateral displacement d and reflects at the effective reflection depth z_{eff} in air.

Different methods exist to evaluate the absolute phase φ. We discuss three methods:

1. the examination of the reflection of a finite spectral wave,
2. the second order derivative of the phase of the reflected field to the angle of incidence,
3. the calculation of the weighted average of the depth of all interfaces.

Calculation of effective reflection depth: first method

The phase φ of the set of reflected fields depends on the group velocity v_g and the group delay t_g of a finite spectral wave package. The group velocity of a wave packet is given by

$$v_g = \frac{\partial \omega}{\partial k}, \tag{7.12}$$

and the group delay by

$$t_g = \frac{\partial \varphi}{\partial \omega}. \tag{7.13}$$

The distance traveled by the incident composite wave corresponds to the average optical path traversed, or, alternatively, to an average depth of reflection z_{eff}. The phase φ of the output field is

$$\varphi_g = \nu_g t_g \, \mathbf{k} \cdot \mathbf{v} = 2 \, z_{eff} \, \mathbf{k} \cdot \mathbf{v}. \tag{7.14}$$

In the general case of an arbitrary wave in three dimensions, the penetration depth or effective depth of reflection z_{eff} is given by

$$z_{eff}(\mathbf{k}, \mathbf{v}) = \frac{\partial}{\partial |\mathbf{k}|} \frac{\varphi(\mathbf{k}, \mathbf{v})}{2}. \tag{7.15}$$

In these equations, an important parameter is the angle of incidence θ between the wave vector **k** and the normal to the surface **v**. Brovelli et al. use the coupled mode equations and come to an estimate of the effective depth of reflection of an infinitely thick perfect Bragg reflector[8]

$$z_{eff}(\theta = 0) \approx \frac{\lambda(n_h + n_l)}{8 \, n_h n_l (n_h - n_l)}. \tag{7.16}$$

This relation leads to an effective depth of reflection of 45 nm in the case of n_h=n_{Si} and n_h=n_{mo}.

Calculation of effective reflection depth: second method

The effective reflecting surface is positioned at a distance z_{eff} from the top of the surface. In the case of a monochromatic wave, the phase shift of the reflected wave in Eq. (7.14) can be rewritten to

$$\varphi = -2 \, z_{eff} k \cos(\theta) \approx -2 \, k \, z_{eff} \left(\; 1 - \frac{\theta^2}{2} + O(4) \right). \tag{7.17}$$

All variables are known, except the exact unwrapped phase shift and the effective reflection depth. By taking the second order derivative of the phase shift as a function of the angle of incidence we find

$$\frac{\partial^2 \varphi}{\partial \theta^2} \approx 2 \, k \, z_{eff} = \frac{\varphi(\theta + \varepsilon) + \varphi(\theta - \varepsilon) - 2\varphi(\theta)}{\varepsilon^2}. \tag{7.18}$$

The resulting expression for the effective reflection depth is

$$z_{eff} = \frac{\varphi(\theta+\varepsilon)+\varphi(\theta-\varepsilon)-2\varphi(\theta)}{2k\varepsilon^2}. \tag{7.19}$$

The approximation of the cosine of the angle of incidence makes Eq. (7.19) only valid for small incidence angles. At normal incidence, the effective depth calculated with Eq. (7.19) is approximately 50 nm, that is less than the thickness of a layerpair difference with the result of Eq. (7.16), see Fig. (7.7).

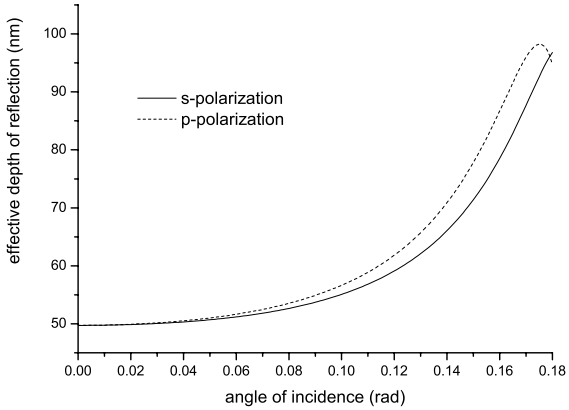

Fig. (7.7) The effective depth of reflection is in Eq. (7.19) and the Figure given as a function of the angle of incidence.

Calculation of effective reflection depth: third method

This method abandons the local summation principle. The separate axially and laterally displaced waves reflect from the multilayer structure and we add the waves in the image plane. A complication here is the infinite number of waves that formally occur after reflection at the multilayer structure. We propose to simplify this picture by concentrating on the *first order* reflected waves. The contribution D_j of an additional layer with number *j* between the substrate and the multilayer is a complex number represented in Fig. (7.5) by a vector

$$D_j = r_j - r_{j-1}. \tag{7.20}$$

The difference in the reflected field of a multilayer with *b* and *b*-1 layers represents the contribution of layer *b*. All these differential contributions add up to the total reflectivity of a multilayer. The relative contribution of a transition to the total reflectivity is defined as the inner product of the differential reflectivity and the total reflectivity of the entire multilayer

$$w_j = \frac{\begin{bmatrix}\mathrm{Re}(D_j)\\ \mathrm{Im}(D_j)\end{bmatrix}^T \cdot \begin{bmatrix}\mathrm{Re}(r_t)\\ \mathrm{Im}(r_t)\end{bmatrix}}{|r_t|^2}, \tag{7.21}$$

that is

$$w_j = \frac{\mathrm{Re}(D_j)\,\mathrm{Re}(r_t) + \mathrm{Im}(D_j)\,\mathrm{Im}(r_t)}{(\mathrm{Re}(r_t))^2 + (\mathrm{Im}(r_t))^2}. \tag{7.22}$$

The relative contribution can be used as the weighting factor w_b in the calculation of e.g. the average depth of reflection and the average lateral displacement. The summation of Eq. (7.21) over all transitions between media in the multilayer is unity. The weighted average of the depth of all interfaces leads to the effective depth of reflection of a multilayer

$$z_{eff} = \sum_j w_j\, z_j = \sum_j w_j\, j \frac{\lambda}{4\,\mathrm{Re}(n_j)\cos(\theta)}, \tag{7.23}$$

where the quantity z_j denotes the optical distance between the top of the multilayer to the transition with number j and it is derived from Eq. (7.5).

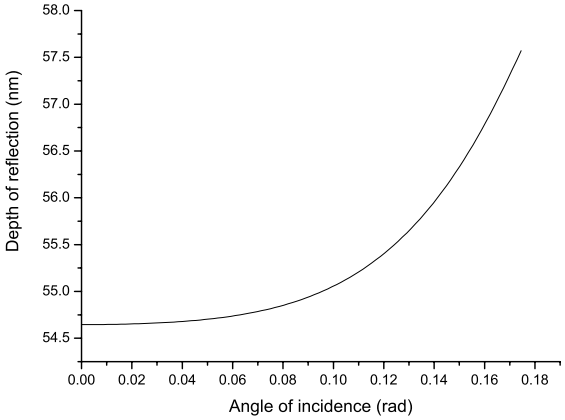

Fig. (7.8) The effective depth of reflection increases somewhat as a function of the angle of incidence.

This method predicts a less sensitive dependence of the effective depth of reflection on the angle of incidence than the previous method, compare Fig. (7.7) and Fig. (7.8). The lateral displacement, d, can easily be derived from the effective depth of reflection. A characterizing quantity of the lateral spread of the reflected beam is the variance of the lateral displacement, d_{var}

$$d_{\mathrm{var}} = \frac{\sum_j w_j \left(\left(z_j - z_{eff} \right) \tan(\theta) \right)^2}{\sum_j w_j}. \tag{7.24}$$

The lateral displacement of the rays increases quickly with larger incidence angles, see Fig. (7.9).

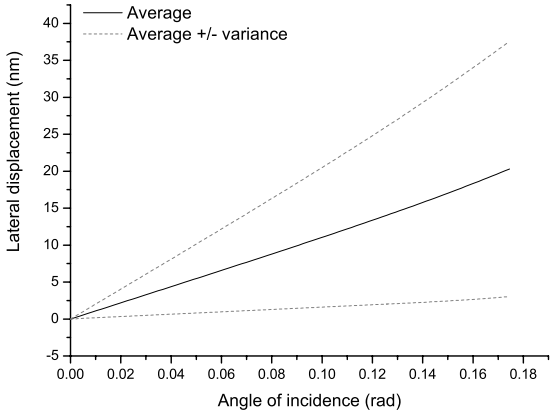

Fig. (7.9) The lateral displacement of the rays incident on a multilayer are strongly dependent on the angle of incidence. In this example, the dependence is almost linear with a gradient of 115 $^{nm}/_{rad}$.

The graphs illustrate the methods to calculate characteristics of the multilayers and in these examples, Eq. (7.5) determines the thicknesses of the layers. The resulting thicknesses are nevertheless not ideal, as the absorption of the materials is ignored in that equation. Normally the thicknesses are optimized to improve the reflectivity over a range of incidence angles, see Section 7.3. An example of a large difference between the ideal thicknesses and the thicknesses calculated with Eq. (7.5) is shown in Fig. (7.10) and Fig. (7.11). With ideal thicknesses, the phases of all contributions to the total reflected fields are equal, which is definitely not the case for large incidence angles.

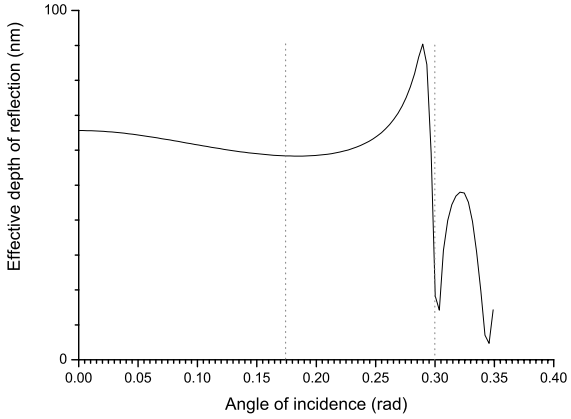

Fig. (7.10) The depth of the effective reflective surface as a function of the angle of incidence. The thicknesses of the layers are adjusted according to Eq. (7.5) for an average angle of incidence of 10°, see the first vertical dotted line from the left in the Figure. Note that the position of the effective reflecting surface starts to oscillate strongly at angles far away from the desired angle.

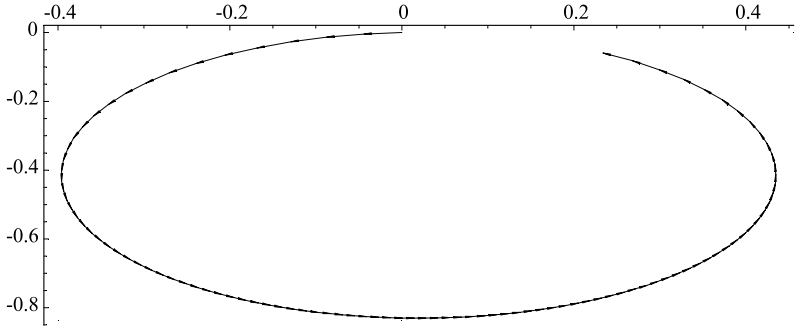

Fig. (7.11) The vectors in the Figure represent the difference in the complex reflectivity r between a multilayer with N+1 layers and N layers, where the additional layer is added between the substrate and the multilayer. In the Figure, all these vectors add up to the total reflectivity of a Mo / Si multilayer with 50 layerpairs of which the thicknesses are given by Eq. (7.5) with an average angle of incidence of 10°. The actual angle of incidence is 17.2°, see the second vertical dotted line in Fig. (7.10).

The evolution of the effective reflection as a function of the number of periods of a molybdenum / silicon multilayer at normal incidence is presented in Fig. (7.12). The effective reflection depth approaches a saturation value at around 45 to 50 layerpairs, the reflected intensity reaches saturation with less layerpairs (approximately 40 layerpairs, see Fig. (7.2)). Increasing the number of layerpairs after the point of saturation does not change the phase of the reflected field. This property can, together with the fact that the reflectivity is saturated, be used in a correction method for multilayer or substrate height errors.

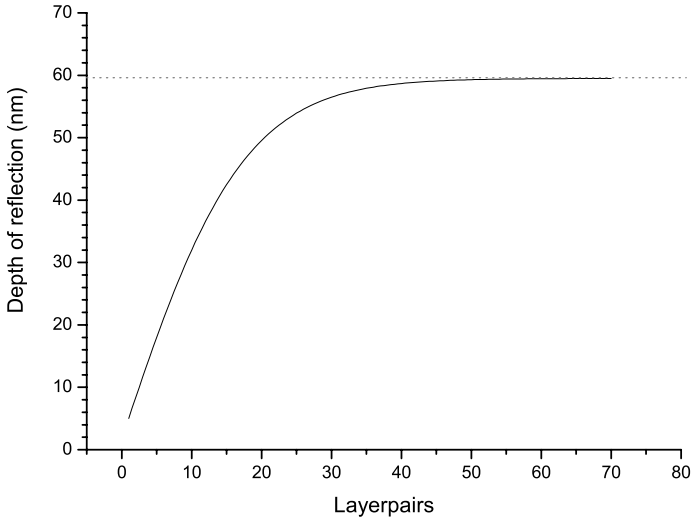

Fig. (7.12) The evolution of the penetration depth of a reflected field as a function of the number of periods of the multilayer. The horizontal dotted line indicates the saturation value of the effective depth of reflection.

7.5 Wavefront correction

The allowed figure errors of the reflecting surfaces in an extreme ultraviolet projection system are extremely small, see the subsection on **figure errors** (page 29). The deposition of the multilayers can warp the substrate for instance, which needs to be corrected. The low spatial frequency surface roughness (LSFR) errors of a multilayer may be corrected by adding a corrective layer. The difference in the phase φ of the field r_1 after reflection and the field r_0 before reflection is given by

$$\Delta\varphi = \varphi_1 - \varphi_0 = \arg(\frac{r_1}{r_0}) . \tag{7.25}$$

The addition of a corrective layer with varying thickness affects the pathlength of the rays in the optical system. The magnitude of the pathlength change depends on the difference in the refractive index of the corrective material, n_c, and that of vacuum, and on the thickness, l_c, of the corrective layer and can be approximated by the expression

$$\Delta\varphi = \arg\left(\frac{e^{i2l_c k\, n_c}}{e^{i2l_c k}} \right) = 2\, l_c\, k\, (n_c - 1) . \tag{7.26}$$

The complex refractive index of the corrective material accounts for an unwanted increase of the absorption. In Fig. (7.13) a molybdenum layer is deposited on a base multilayer of fifty periods of molybdenum and silicon. In this example the pathlength changes –0.15 nm per additional nm in the thickness of the corrective molybdenum layer. The curves in Fig. (7.13) have been obtained by using a complete thin-film analysis of the combination of the multilayer and the corrective layer.

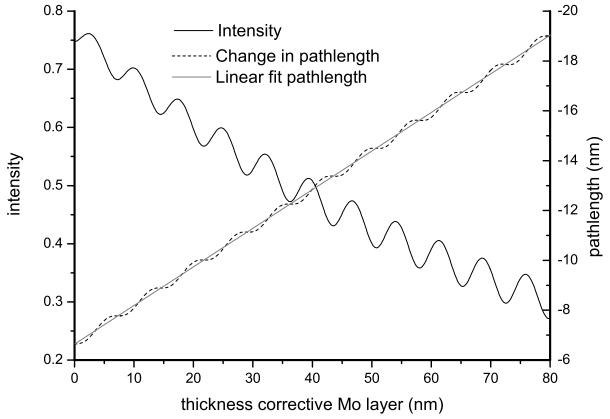

Fig. (7.13) The reflectivity and pathlength change as a function of the thickness of a corrective molybdenum layer. The left vertical axis specifies the scale of the decreasing intensity of the reflected field. The right vertical axis indicates the pathlength change. Since molybdenum has a refractive index smaller than unity, the pathlength of a ray decreases when vacuum is replaced by molybdenum, see Fig. (7.14).

In case the (mean) refractive index of the corrective layer or the added corrective (multi) layer is less than unity, valley-like figure errors are corrected, see Fig. (7.14). Hill-like errors demands refractive indices larger than unity, unfortunately in the spectral region of

interest these materials exhibit a large absorption. Other corrective methods include the removal[9] of layers of the base multilayer and the addition[10] of layers of various materials on top of the base multilayer, see Fig. (7.15).

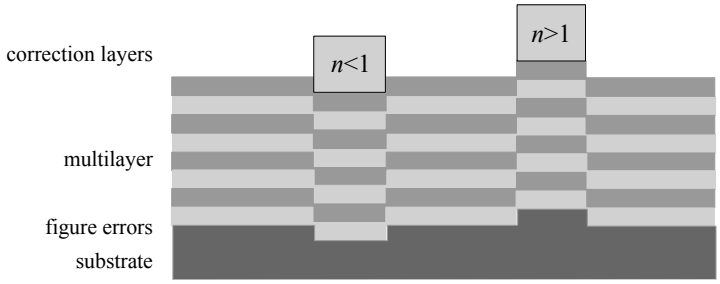

Fig. (7.14) The Figure shows a schematic drawing of a substrate with height errors, the multilayer structure and locally deposited correction layers. A lowering of the substrate is corrected with correction layers with an average refractive index smaller than unity. An average refractive index larger than unity corrects an elevation of the substrate.

The removal or deposition of a single period of the standard two-component molybdenum / silicon multilayer interference coating induces an effective phase change of $0.043\pi+2\pi$ compared to an identical thickness in vacuum, see Fig. (7.16). An important difference with the corrective layer method is that the origin of the reflected field moves up when layers are added to the front of the multilayer. In other words, the average field is reflected at a different level. This concept of the location of the reflection of the average field has been discussed in Section 7.4 and named the effective depth of reflection. When a corrective layer is added, the position of the effective depth of reflection remains virtually unchanged, relative to the substrate. In the case of a corrective multilayer stack, the effective depth of reflection roughly changes by the thickness of the corrective multilayer.

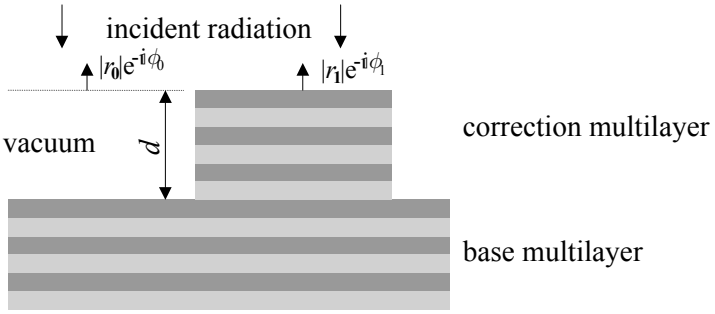

Fig. (7.15) A schematic representation of the geometry and the associated phase shifts due to additional or removed multilayer periods. The average depth of reflection alters when a corrective multilayer is added above the original saturated multilayer.

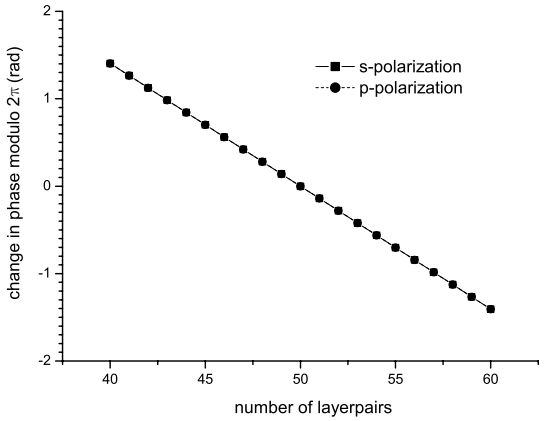

Fig. (7.16) The difference in average refractive index of a multilayer and vacuum produces the change in the phase modulo 2π when increasing the number of layerpairs in a multilayer. The graph shows that the induced phase shift is to a large extend linear with the number of removed or added layerpairs.

The presented phase change of the correction multilayer is larger than expected by only considering the refractive indices together with the thickness of the added corrective multilayer, see Eq. (7.26). The slight increase of the effective reflection depth just before saturation explains this discrepancy, see Fig. (7.12). For ten deposited periods of molybdenum / silicon for example the change of the effective reflection depth accounts for a phase change of 0.19π radians.

The magnitude of this wavefront shift can be enhanced using multi-component multilayers optimized for phase-change on reflection at the expense of reflectivity, see Table (7.3).

Table (7.3) The Table shows the phase sensitivity ($\Delta\varphi$) and the peak reflectivity of various correction multilayer designs. These correction multilayers are deposited on an optimized fifty-period molybdenum / silicon base multilayer tuned for λ=13.4 nm operation with a peak reflectivity of 0.748. The correction multilayers are of the form *x-y/z*.

Form x-y/z	Layer thi. (nm)			Peak Reflectivity		Δφ
	x	y	z	5 per.	10 per.	(units of π)
Mo/Si	-	2.68	4.17	0.749	0.749	0.043
Mo/Y	-	2.61	4.37	0.695	0.664	0.090
Ru-Mo/Y	2.00	2.00	3.06	0.685	0.639	0.109
Ru-Mo/Y	2.00	2.00	3.15	0.670	0.626	0.121
Ru-Mo/Si	2.00	2.00	3.10	0.658	0.608	0.111

7.6 Modeling of multilayers in optical design

To evaluate the optical imaging consequences of the incorporation of the multilayers in projection systems, the optical design programs Oslo[11] and CodeV[12] are used. The reflectivity and the phase shift of a ray reflected from a multilayer depend on the angle of incidence θ on a surface. The differences in phase-shift on a mirror have to be included

when evaluating the imaging quality of a projection system. Chen Liang et al. analytically looked at the aberrations introduced by multilayer coatings on a tilted flat mirror in the case of an oblique incident, finite aperture beam[13]. They found, up to the second order, a first-order field-dependent piston, a field-squared piston, defocus, field-independent tilt, field-independent astigmatism and anamorphic magnification. N.J. Duddles found that wave aberrations introduced by multilayer incorporation in extreme ultraviolet projection systems can be effectively described by a simple tilt and defocus of the wave front, allowing system performance to be recovered to a level close to that of the bare system[14]. F. Cerrina et al. report a non-negligible amount of spherical aberration introduced by the multilayers and a change of the apparent focal length of the optics[15].

In our analysis, we approximate a multilayer with an effective reflecting surface in optical design software, see Fig. (7.6). Secondly, we do not neglect the lateral displacement, i.e. the distance between the intersections of the top surface with the incoming and outgoing ray.

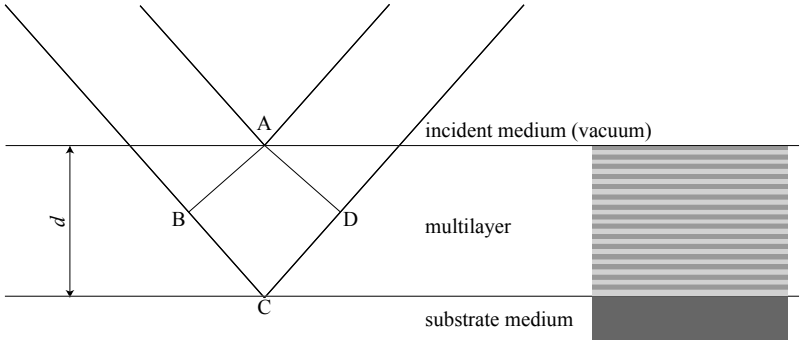

Fig. (7.17) The additional pathlength [*BCD*] a ray travels when reflecting on the substrate compared to a ray reflected on the top of the multilayer.

Optical design programs can model multilayers to some degree[11, 12]. The program traces the propagation of rays through an optical system; the rays start in the entrance pupil, pass the object point and intersect the successive surfaces in the system. A surface with a multilayer is first treated as a single transition between two media, that is a normal surface. The intersection of a ray with the surface is iteratively found, and, knowing the local surface normal, the angle of incidence is known. Using the refractive indices of the material on both sides of the transition, the angle of the outgoing ray is computed. Only when the new propagation direction of the ray is known, the multilayer on the surface is considered. With the angle of incidence and the location of the intersection of the ray with the surface as parameters, the phase and intensity changes of the ray are calculated.

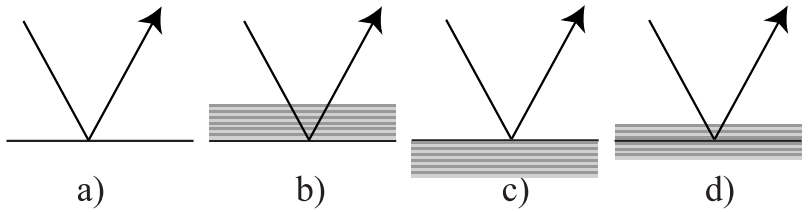

Fig. (7.18) The different ways to include multilayers on a reflective surface. In *a)* the surface is a single transition between two media. In *b)* the substrate is the original surface, the method applied by CodeV. The method in *c)* produces smaller alterations to the characteristics of the extreme ultraviolet projection systems, as the thickness of the multilayer stack is many times the wavelength. In *d)* the effective reflection depth represents the average location of the reflections.

Normally, coatings are employed for reflection enhancement, as an antireflection means or as a bandwidth filter coating. In most of these cases, it is convenient to put the coating on top of the original surface; the modeled surface is the substrate surface, see Fig. (7.18 part b). The optical design program reflects the rays from the substrate level, point *C* in Fig. (7.17). Using thin-film calculations, the optical design program, a thin-film program[2] or a home-made program computes the phase shift and reflectivity of the reflected field relative to the top of the coating as a function of the angle of incidence and the wavelength. In the optical design program, the phase associated with the traversal of the multilayer is subtracted from the phase change of the thin-film calculations to compensate for the thickness of the multilayer, see Fig. (7.17). The approach to position the substrate at the original surface becomes disadvantageous for thicker multilayers, as in the case of extreme ultraviolet multilayers. A typical value for the total thickness of an extreme ultraviolet multilayer is 300 nm, whereas the effective reflection occurs at approximately 50 nm below the top of the multilayer. In this approach, the difference in optical path length of a reflector with a single transition and a multilayer is mainly caused by the thickness of the multilayer, see Fig. (7.19).

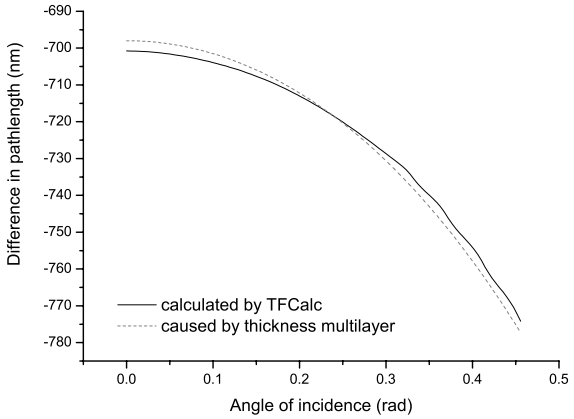

Fig. (7.19) The thickness of the multilayer is the most important originator of the optical pathlength difference between a reflector with a single transition and a multilayer. The program TFCalc has been used for the thin-film calculations.

An alternative approach is to locate the top of the multilayer on the original surface, see Fig. (7.18 c). To implement this approach, all surfaces can be repositioned in the optical design program by adding the thickness of the multilayer to the distances between the mirrors. Alternatively, a combination of a thin-film program with the optical design software can be used.

The effective reflection depth represents the location where the average field reflects. We propose to use the concept of an effective reflection depth and surface in optical design software, wherein the reflected field resulting from an incident ray must be represented by a single outgoing ray. The effective reflecting surface replaces the original surface, which would be the best representation of the multilayer, see Fig. (7.18 d). With this method, the average lateral displacement is correctly modeled.

Rays are perpendicular to a wavefront. The phase shift and intensity originating from the multilayers is a function of the angle of incidence and the thicknesses of the layers. The angles of incidence depend on the location of the intersection with the mirror and the position of the object point. Different thicknesses of the layers in the multilayer change the intensity and phase of the reflected field. The difference in phase shifts and intensity introduce a change of the wavefront and a directional change of the gradient of the wavefront, that is a ray, see Fig. (7.20). These changes are not considered in the approaches discussed in this section. The presence of multilayers is only manifest in the optical pathlength and the intensity of rays in optical design programs. Therefore, the rays could become oblique to the wavefront, which conflicts with the definition of a ray, an effect that increases with increasing average angle of incidence, see Fig. (7.21).

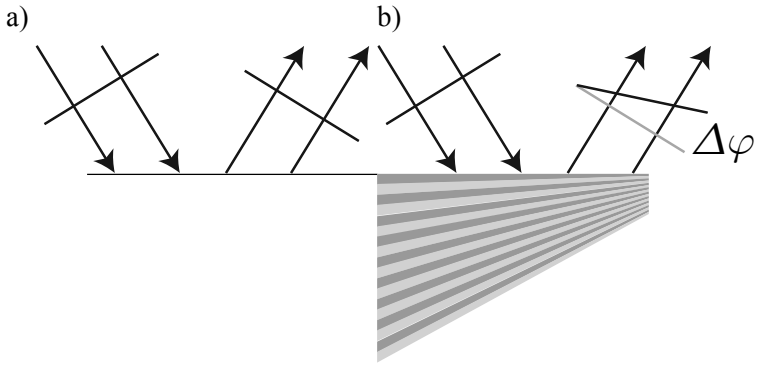

Fig. (7.20) The phase change originating from multilayers depends among others on the angle of incidence and the thicknesses of the layers. In the Figure two rays representing a plane wavefront reflect at a single transition (a) or at a graded multilayer (b). The phase variation originating from the multilayer is not equal for all rays in situation (b); consequently, there is no unique ray-wavefront combination in the reflected field. These types of layers demand a more thorough computation of the field resultant from the multilayer.

Another difficulty is that the thin-film calculations assume a plane wavefront incident on a plane surface with parallel multilayers, while our systems have curved wavefronts and curved surfaces with graded multilayers. The validity of the thin-film approximation is not indisputable. A weaker but similar condition, compared to the assumption made in the thin-film approximation, will be used in subsection 7.7.1 to calculate the grading of the multilayers.

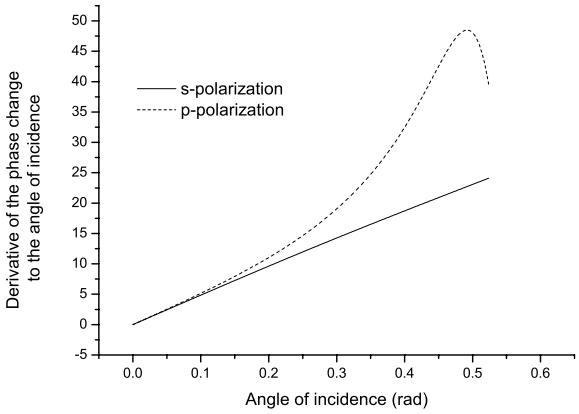

Fig. (7.21) The derivative of the phase change with respect to the angle of incidence demonstrates an increasing effect at larger average angles of incidence. The thicknesses in the multilayer are adjusted to the average angle of incidence (horizontal axis) using Eq. (7.5).

7.6.1 The implementation of multilayers in existing design software

The optical design program Code V[11] includes thin-film calculations, approached in a method similar to the thin-film calculations done in the program TFCalc[2] and our calculations in Mathematica[1]. Though the calculation methods are identical, the results are different.

Including the multilayer in an optical system produces a difference in the optical path length of the rays. The substrate corresponds to the original single transition[16]. In a combination of molybdenum and silicon, the total thickness of the multilayer is typically 350 nm thick at zero incidence angle. The expected difference in optical pathlength between a system with and without a multilayer is approximately twice the thickness of the multilayer (700 nm). In contrast with what we expect, the optical design program comes up with an almost zero pathlength difference. A more annoying phenomenon are the jumps in the optical path length, see Fig. (7.22). The size of these jumps is exactly the wavelength, 13.4 nm, which suggests phase jumps of 2π. To distinguish between the s- and p-polarization the optical design program also tracks the phases of the s- and p-polarization. These phases are not included in Fig. (7.22), as these introduce even more jumps.

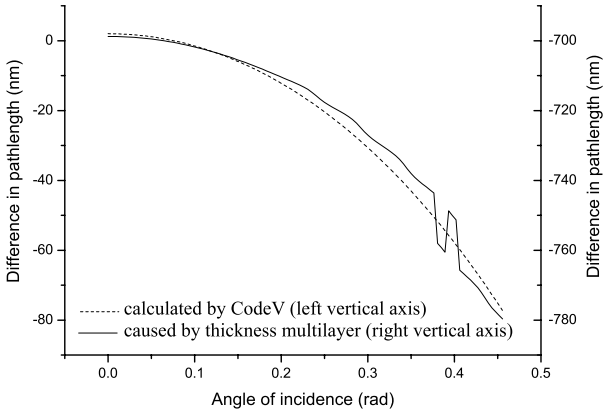

Fig. (7.22) The difference in optical pathlength between a system with and without a multilayer calculated in the optical design program as a function of the angle of incidence. The jumps are exactly a wavelength high and suggest discontinuities of 2π in the phase.

The disagreement of the optical design program Code V with other thin-film calculations was presented to its manufacturer, but could not be solved. A comparison of TFCalc and of Code V is shown in Fig. (7.23). Our calculations done in Mathematica exactly match the calculations done by the thin-film program TFCalc.

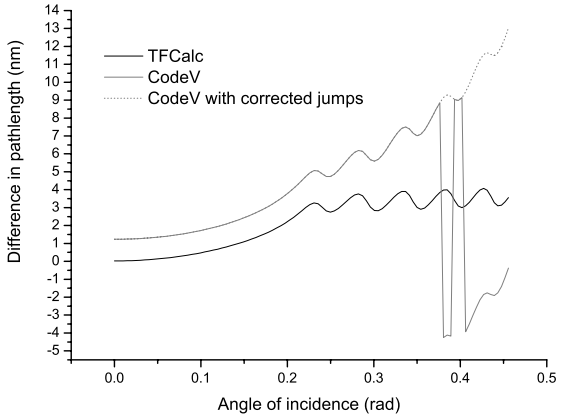

Fig. (7.23) The comparison of the incorporation of multilayers in an optical system with the thin-film calculations done by TFCalc or CodeV. Even when the jumps are corrected, the developments of the optical path length calculated in the two programs do not match.

7.6.2 Optimizing graded multilayers

In a typical extreme ultraviolet projection system, the angles of incidence of rays are not constant on a particular surface. The characteristics of the whole multilayer must be optimized for the domain of the angles of incidence, see Section 7.3. The thicknesses in the multilayer can also vary as a function of the position on the mirror. Multilayers with varying thicknesses of the layers are named *graded* multilayers.

The optical design program Code V can model graded multilayers. The coefficients of the grading can however not be included in the set of variables to be optimized. In the previous section, some disagreements in the thin-film calculations of the optical design program were discussed. Therefore we prefer to use a combination of a thin-film program and the optical design program. The combination enables the optimization of the grading of the multilayers simultaneously with the common variables of the optical system. The transmission of the rays is included in the merit function of the optimization. We use user-defined surfaces, e.g. (translated) Chebyshev surfaces (see the subsection on **chebyshev polynomials**, page 42), with an additional coefficient for the grading of the multilayer. All thicknesses in a multilayer vary with a factor that is only dependent on the position on the surface.

The sequence in which the optical design program evaluates the rays in a system is shown in Table (7.4). First, the program iteratively calculates the intersections of a surface with a set of rays, and the changes to the optical pathlengths and intensities are included. The Table shows the parameters given by the program, the parameters needed to do the calculations, where and how the principal calculations are done and the output of the user-defined functions.

Table (7.4) The incorporation of graded multilayers in CodeV needs the use of user-defined surfaces and user-defined interferogram functions. The coefficients of the user-defined surface include the coefficients of the graded multilayer. A Dynamic Link Library (DLL) is a separate file that contains functions and resources for Windows-based programs.

Intersection surface	
input	surface number
	coefficients surface
	coefficients grading surface
	x and *y* coordinates
needed	coefficients surface
	x and *y* coordinates
calculations	done in CodeV DLL
output	*z* of the surface function
Influence multilayer on optical path length	
input	angle of incidence
	x and *y* coordinates of surface intersection
	Wavelength
	surface number
needed	coefficients grading surface
	surface number
	Wavelength
calculations	done in TFCalc DLL via CodeV DLL
output	relative path length change

Influence multilayer on transmission	
input	angle of incidence
	x and y coordinates of surface intersection
	Wavelength
	surface number
needed	coefficients grading surface
	surface number
	Wavelength
calculations	done in TFCalc DLL via CodeV DLL
output	relative path length change

7.7 The optical performance of systems with multilayers

In this section, we estimate the consequences of the presence of multilayers on the optical performance. The point spread functions (PSF) of a perfectly focused beam returning from a reflecting surface with a graded or with a non-graded multilayer are compared. In our example, the perfectly focused beam comes from a parabolic reflector focusing a collimated beam as shown in Fig. (7.24). The intensity of the point spread function of a single reflecting surface in a system with a numerical aperture of 0.3 in the image plane is shown on the left side in Fig. (7.25). Each transition between two different media in the multilayer leads to a displaced focus. The contribution of each transition is the contribution of the differential field of a layer to the total reflected field, see Eq. (7.21).

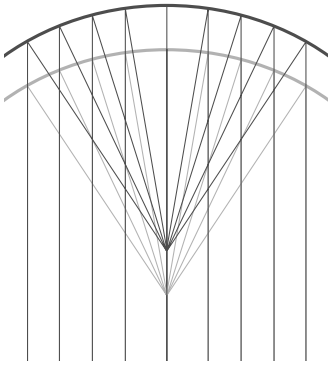

Fig. (7.24) A parabolic reflector perfectly focuses a collimated beam. When the reflector is covered with a multilayer each transition between layers leads to a focus displaced by the equivalent distance in vacuum between the layer transitions. We suppose that the refractive indices of the multilayer materials are close to unity.

Due to the high absorption in the materials at extreme ultraviolet wavelengths, the contributions of the transitions further away from the top surface rapidly decrease. This decrease causes the asymmetry of the point-spread function when a multilayer is deposited on a reflector, as shown on the right side of Fig. (7.25). When using non-graded multilayers consisting of materials with a refractive index close to unity, of which the thickness along the optical axis is constant, the distance between the different foci equals the thickness of a layer.

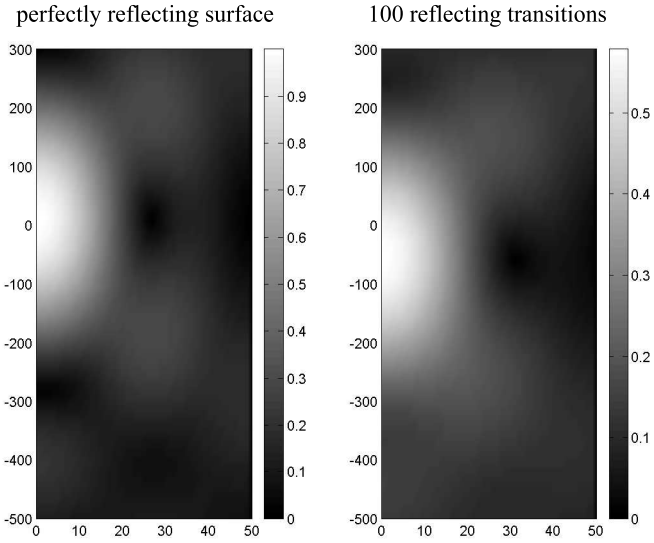

Fig. (7.25) The intensity of the point spread function of a perfect system with a numerical aperture of 0.3 in the image plane with a graded (left side) and with a non-graded (right side) multilayer of fifty layerpairs. The defocus along the optical axis is plotted along the vertical axis, in a range from -500 to 300 nm. The radial distance r to the optical axis is plotted along the horizontal axis, in a range from 0 to 50 nm. The optical axis coincides with r=0. Note the different absolute scales in the axial and radial directions.

In Fig. (7.25) the numerical aperture in the image plane is 0.3, a typical value for six-mirror extreme ultraviolet projection systems. The numerical aperture of eight-mirror systems is in the order of 0.4. In Fig. (7.26) we show the intensity of the point-spread function of a perfect system with a numerical aperture of 0.5 to exaggerate the influence of the multilayer. Here again, the left side shows the intensity of the point spread function of a perfect parabolic mirror with a single reflective surface and the right side sum the contribution of the interfaces of fifty layerpairs. In the case of the perfect mirror with a single reflective surface, the point spread function at a numerical aperture of 0.5 shows a hyperbolic decrease of the feature size compared to an aperture of 0.3, according to Eq. (2.2). The depth of focus depends on the inverse square of the numerical aperture, see Eq. (2.4). Because of the use of a multilayer as a reflector, the depth of focus increases, the feature size increases and the intensity decreases. The main conclusion is that the incorporation of non-graded multilayers severely degrades the optical performance, including the point-spread function. The thickness of the layers in the multilayer should be optimized to make the different foci coincide. In a graded multilayer the thicknesses change as a function of the location on the mirror or, if preferred, as a function of the incidence angle, see Eq. (7.5), in order to counteract the degradation of the point-spread function.

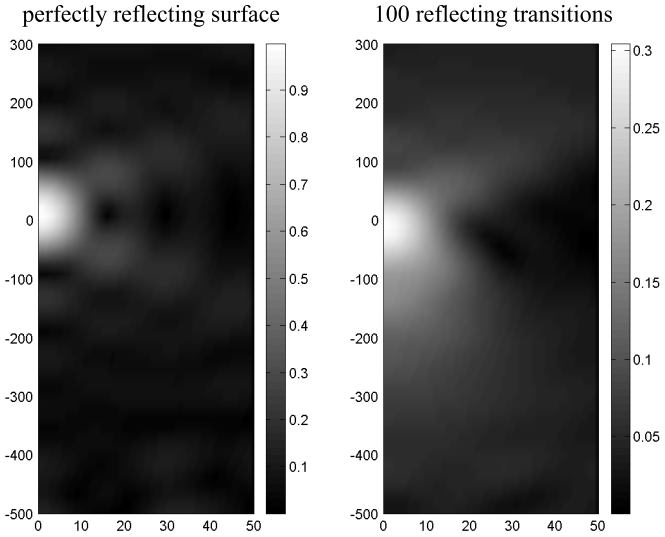

Fig. (7.26) The intensity of the point spread function of a perfect lens with a numerical aperture of 0.5 in the image plane with a graded (left side) or with a non-graded (right side) multilayer of fifty layerpairs. The ranges on both the horizontal and vertical axis are the same as in Fig. (7.25).

7.7.1 Calculation of the grading of multilayers

A conventional method to calculate the grading of a multilayer is to evaluate the average (and the standard deviation) of the angles of incidence in a set of different positions on a surface. With the average of the angle of incidence, the rescaling factor of the thicknesses at the different positions can be computed, with for instance, Eq. (7.5). The rescaling factor only depends on the radial distance r to the optical axis, to retain the rotation symmetry in the system.

The thin-film calculations coherently sum all reflections coming from the different transitions in the multilayers, which assumes an infinitely extending plane wave incident on a multilayer consisting of plane parallel surfaces. This condition is only satisfied with plane reflectors and parallel layers in the multilayer. A less strict but similar condition demands that the foci of all fields originating from different transitions coincide, in order to achieve good optical imaging performance with a multilayer. The latter condition requires that all reflected fields are co-directional and have the same phase modulo two pi. The reflected fields are then in resonance. We propose to use this last condition to calculate the grading of multilayers.

The condition that the different foci of the reflected fields of a single object point should coincide in the image plane implies that the propagation direction of rays on every position in the interface of the multilayer and vacuum is unique. This is the case for a reflecting surface with a single transition. In the case of multilayers, an important benefit of this condition is the disappearance of the necessity to evaluate a set of rays per location on the surface. In the typical case of a molybdenum / silicon extreme ultraviolet multilayer that set

comprises at least a hundred different rays with different directions, intensities, origins, and phase shifts for each position on the multilayer surface. In a six-mirror system, each ray in the object plane would lead to $100^6=10^{12}$ rays to evaluate in the image plane.

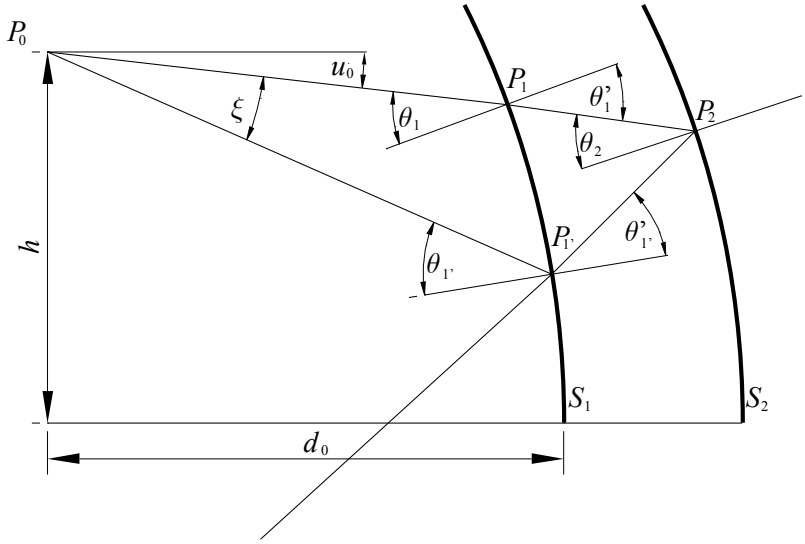

Fig. (7.27) The Figure shows the distributed reflection on a surface with a set of rays coming from an object in the left-top of the picture. We enforce the shape of the layer S_2 to be such that the rays reflecting on the front coincide with rays reflected on the back of the surface.

In Fig. (7.27) two rays originate from an object point P_0 located in the left-top corner of the picture at a height h and at an axial distance d_0 to the first reflective surface. The thick lines in the Figure represent the two almost conformal surfaces S_1 and S_2. Two rays originate from the object and hit the first surface. Of the first ray, we only consider the transmitted ray that is reflected on the second surface in point P_2. Of the second ray, we consider the reflected ray at the first surface in point $P_{1'}$.

The angle of the first ray relative to the optical axis at the object point is u_0. For this ray we find, using Snel's law

$$n_0 \sin(\theta_1) = \sin(u_0 + \alpha_1) = n_1 \sin(u_1 + \alpha_1) = n_1 \sin(\theta_1'), \tag{7.27}$$

where α_i is the slope of a surface in point P_i relative to the optical axis and n_i the refractive index of the material after surface i. The slope in an intersection point is given by

$$\alpha_i = \tan^{-1}\left(\frac{\partial z_i}{\partial r}\right)_{r=r_i}, \tag{7.28}$$

where z_i is the surface sag of surface i.

The angle of the second ray is u_0 plus a small deviation ξ in the object space

$$\theta_2 = u_0 + \xi + \alpha_2, \tag{7.29}$$

$$n_0 \sin(u_0 + \xi + \alpha_{1'}) = -n_1 \sin(u_{1'} - \alpha_{1'}). \tag{7.30}$$

The intersection of the first ray with the second surface is in point P_2 in Fig. (7.27). The slope of the second mirror in point P_2 is given by

$$\alpha_2 = -\frac{u_1 + u_{1'}}{2}. \tag{7.31}$$

The deviation ξ should be such that the first and second ray both pass point $P_{1'}$ and propagate in the same direction. In order to get maximum positive interference, the phase difference between the two rays in point $P_{1'}$ should be a multiple of 2π.

The angle ξ is in general very small and can be found for a known object distance, object height, first surface, and angle u_0. With the angle ξ, the position and slope of the point P_2 on the second surface is found. By varying the angle u_0 within the proper domain, a set of locations and slopes of the second reflective surface are found. These points and slopes of the surface can be fitted with a polynomial. An example of the calculation of the grading of a layer is shown in Fig. (7.28) and Fig. (7.29).

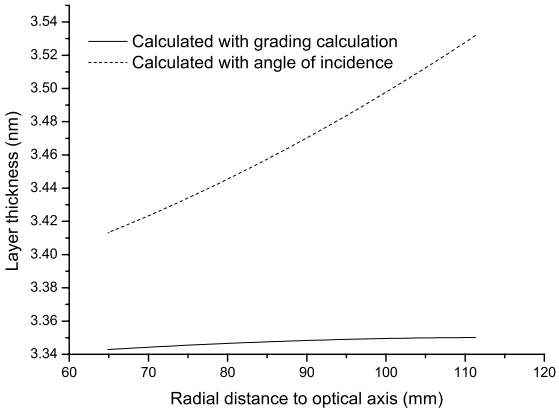

Fig. (7.28) The Figure shows the thickness of a single layer of Silicon on the first reflective surface of the four-mirror design shown in Fig. (6.3), calculated with two methods. The first method is to calculate the thickness of the multilayer as a function of the angle of incidence, see Eq. (7.5). The second method is discussed in this subsection.

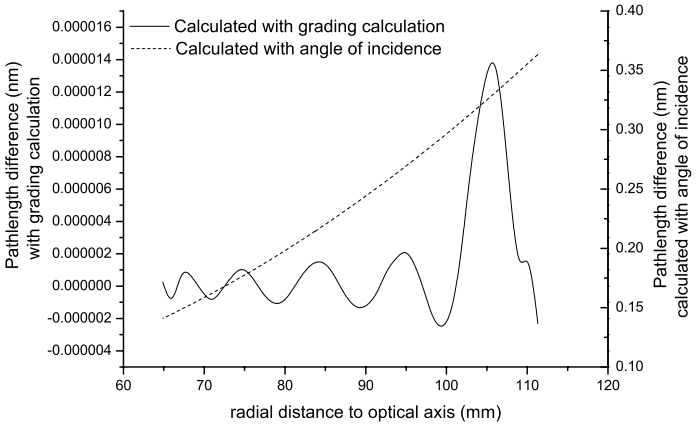

Fig. (7.29) The intended difference in the pathlength between a ray reflected on the first transition and the second transition is half a wavelength. In the Figure, the difference between the intended and the actual pathlength difference is shown as a function of the radial distance to the optical axis. Note the extremely different vertical axis scales in the Figure.

7.8 Conclusion

The thin-film approach of multilayers assumes an infinitely extending plane wave incident on an infinitely extending plane multilayer with parallel transitions. The field reflected by a multilayer relates to the input field by an absolute relative amplitude and an absolute phase, see Eq. (7.10). The absolute value of the relative amplitude results from thin-film calculations. The absolute phase can be determined with several methods, of which we prefer the method of calculating the weighted average of the depth of all interfaces. A multilayer can be replaced by a surface located at an *effective reflective surface* in optical design software. The absolute phase and the absolute relative amplitude give the position and reflectivity of the effective reflective surface.

The allowed low spatial frequency roughness (LSFR) errors of each multilayer are extremely small. Two correction methods of these errors are the addition of a corrective layer and the addition of a corrective multilayer. The removal or deposition of a single standard molybdenum / silicon multilayer induces an effective phase change of 0.043 π and an absolute phase change of 2.043 π.

The multilayers and wavefronts in an extreme ultraviolet projection system are not flat. This results in a degradation of the point-spread function when multilayers with uniform thicknesses of the different layers are used. The rescaling factor of the thicknesses at the different positions can be calculated with the average cosine of the angle of incidence. To avoid the degradation of the point-spread function, the propagation direction of rays on each position on the mirrors must be unique and the difference in phase an integer times 2π. With these condition we calculate the grading of multilayers.

7.9 References

1. Mathematica, Wolfram Research Inc., Champaign, 1999.
2. Thin Film Design Software for Windows, Software Spectra Inc., Portland, 1999.
3. H. A. Macleod, *Thin-film optical filters,* 2 ed. (Adam Hilger Ltd, Bristol, 1986).
4. B. Henke, E. Gullikson, and J. Davis, "X-ray interactions: photoabsorption, scattering, transmission and reflection at E = 50-30,000 eV, Z = 1-92"," At. Data Nucl. **54**, 181-342 (1993).
5. A. G. Michette and Z. S. Wang, "Optimisation of depth-graded multilayer coatings for broadband reflectivity in the soft X-ray and EUV regions," Opt. Commun. **177**(1-6), 47-55 (2000).
6. M. Singh and J. J. M. Braat, "Design of multilayer extreme-ultraviolet mirrors for enhanced reflectivity," Appl. Opt. **39**(13), 2189-2197 (2000).
7. J. I. Larruquert, "Reflectance enhancement in the extreme ultraviolet and soft x rays by means of multilayers with more than two materials," Journal of Optical Society of America A **19**(2), 391-397 (2002).
8. L. R. Brovelli and U. Keller, "Simple Analytical Expressions for the Reflectivity and the Penetration Depth of a Bragg Mirror between Arbitrary Media," Opt. Commun. **116**(4-6), 343-350 (1995).
9. M. Yamamoto, "Sub-nm figure error correction of an extreme ultraviolet multilayer mirror by its surface milling," Nucl. Instrum. Methods Phys. Res. Sect. A **467**, 1282-1285 (2001).
10. M. Singh, M. Bal, J. Braat, D. Joyeux, and U. Dinger, "Wavefront correction methods for extreme ultraviolet multilayer reflectors," to be published (2002).
11. OSLO, Lambda Research Corporation, Littleton, 2001.
12. Code V, Optical Research Associates, Pasadena, 2001.
13. C. Liang, M. R. Descour, J. M. Sasian, and S. A. Lerner, "Multilayer-coating-induced aberrations in extreme-ultraviolet lithography optics," Appl. Opt. **40**(1), 129-135 (2001).
14. N. J. Duddles, "Effects of Mo/Si multilayer coatings on the imaging characteristics of an extreme-ultraviolet lithography system," Appl. Opt. **37**(16), 3533-3538 (1998).
15. F. Cerrina, S. Bollepalli, M. Khan, H. Solak, W. Li, and D. He, "Image formation in EUV lithography: Multilayer and resist properties," Microelectron. Eng. **53**(1-4), 13-20 (2000).
16. "Incorporation of multilayer coatings into raytrace: phase shift correction terms for transmitted and reflected rays," in *Optical Research Associates, internal software development document,* (Pasadena, 2001).

8 Conclusions

This thesis describes the design of an extreme ultraviolet projection system. We emphasised on the initial design of a projection system and the optical characteristics of the multilayers in a projection system.

For the design of extreme ultraviolet projection systems, the large number of parameters makes the search for a starting configuration for subsequent optimization difficult. We have developed an efficient method based on the paraxial approximation to find starting points which can subsequently be optimised. The paraxial approximation does not use the aspherical coefficients and enables the possibility to impose several conditions to the projection system. Because of this, the number of variables that describe a reflective projection system drastically decrease to $2N$-1, where N is the number of reflective surfaces.

An essential condition a projection system has to fulfil is the absence of obstruction. Particularly this requirement limits the solution spaces of systems to small domains. With the relative small number of parameters in the paraxial model, we do exhaustive searches for possible configurations for which the absence of obstruction is the principal criterion. Comparing paraxial and finite ray-tracing results by means of two-dimensional analyses has proven the validity of our paraxial method. With the two-dimensional analyses, we compare for instance the evaluation of obstruction with paraxial or finite ray tracing. The paraxial obstruction-free domains include the corresponding finite obstruction-free domains. This property guarantees that our exhaustive paraxial search method is able to detect all obstruction-free domains of interest for starting configurations.

We have developed a new method based on the relative arrangement of mirrors to classify extreme ultraviolet systems. Unlike earlier classifications, the class of a system does not change, when the system is optimized. We have presented the systems found in the different classes with the exhaustive search method for four-, six- and eight-mirror projection systems. Systems in the useful four- and six-mirror classes are already known from the patent literature. An exception is a four-mirror system in class 6-, which has a good optical performance and a relatively large numerical aperture. Some systems show surprising mirror arrangements, although factors other than the absence of obscuration make some of them less attractive. The avoidance of large angles of incidence is one of such factors. Large incidence angles make aberration correction difficult and multilayer compatibility questionable. However, some of the eight-mirror classes with good optical performances are new and promising for further development.

A typical multilayer in extreme ultraviolet consists of fifty layerpairs of molybdenum and silicon. The total thickness of these multilayers is approximately twenty-five wavelengths. We have shown that the average reflection occurs at approximately 3.7 wavelengths below the top of the multilayer. The average or effective reflection depth determines the imaging properties of the multilayer, e.g. the average lateral ray displacement and phase change on reflection. The average value of the cosines of the incidence angles should remain close to unity. The optimum average angle of incidence is of the order of 10°, with a small variation. Apart from the reflectivity, the relative phase shift also changes as a function of the incidence angle, effectively adding aberration to the wavefront.

In optical design programs, phase changes that occur at reflection due to the presence of multilayers are summed at the exit pupil and are combined with the optical path length of

the ray. At the same time, reflectance changes are multiplied. These phase and intensity variations result from thin film calculations, which assume a coherent plane wavefront incident on a plane multilayer. The reflected fields are summed up into a single outgoing field. On the other hand, the optical performance deteriorates when the condition that all rays at a given position on the reflecting multilayer have a unique direction is not fulfilled. We showed that the latter condition together with the condition for maximum positive interference (i.e. the phase differences of the rays on a given position differ an integer times 2π), can be used to calculate the local optimum thickness of a multilayer (*grading*).

9 Abbreviation and symbol list

The following subsections list the abbreviations, symbols and mathematical notations used. The numbers refer to the page where the abbreviation or symbol are explained and defined.

9.1 Abbreviations

ASIC	Application-Specific Integrated Circuit	13
B	boron	102
CD	Critical Dimension	16
CVAL	Curvilinear Variable Axis Lens	20
DLL	Dynamic Link Library	118
DRAM	Dynamic Random Access Memory	13, 15
DUV	Deep UltraViolet	18
EBDW	Electron Beam Direct Write	19
EM	ElectroMagnetic	35
EPL	Electron projection lithography	19
ETS	Engineering Test Stand	21, 22, 87
EUCLIDES	Extreme Ultraviolet Concept Lithography Development Systems	21
EUV	Extreme UltraViolet	20, 22
HSFR	High Spatial Frequency Roughness	30
ITRS	International Technology Roadmap for Semiconductors	15
LBNL	Lawrence Berkeley National Laboratory	21
LEEPL	Low Energy Electron Proximity Lithography	20
LLC	Limited Liability Company	21
LLNL	Lawrence Livermore National Laboratory	21
LPP	laser-produced plasma	22
LSFR	Low Spatial Frequency Roughness	29, 109
MAPPER	Multi-Aperture Pixel-by-Pixel Enhancement of Resolution	20
MEMS	Micro Electro-Mechanical Systems	13
MHz	Millions of Instructions Per Second	15
MIPS	Millions of Instructions Per Second	15
MPU	MicroProcessor Units	13, 15
MSFR	Mid Spatial Frequency Roughness	30
MTF	Modulation Transfer Function	38
NGL	Next-Generation Lithography	19
OPD	Optical Path Difference	38
OTF	Optical Transfer Function	38
Pd	palladium	102
PO	Projection Optics	21
PREVAIL	Projection Exposure with Variable Axis Immersion Lenses	20
PSF	Point Spread Function	16, 119
PXL	Proximity X-Ray Lithography	20
Rb	rubidium	102
RbCl	rubidium-chloride	102
Rh	rhodium	102
rms	root-mean-squared	24

Ru	ruthenium	102
SCALPEL	Scattering with Angular Limitation Projection Electron-Beam Litho.	20
Sematech	Semiconductor Manufacturing TECHnology	15
SiO_2	silicon-dioxide	102
SLM	Spatial Light Modulator	69
SNL	Sandia National Laboratories	21
Sr	strontium	102
VNL	Virtual National Laboratory	21
YAG	Yttrium Aluminum Garnet	22

9.2 Symbols

$\bar{}$		refers to the principal ray	36
$R_n^m(r)$		the radial part of the Zernike polynomials of order *n*	42
$\|r\|$		the reflection coefficient	103
A		the matrix with derivatives of operands with respect to variables	75
a_i		the sign of the angle of incidence of the principal ray at surface *i*	46
b		the number of values for each parameter	64
B	[m^2]	the total difference between finite and paraxial ray tracing	60
C		the class number of a system	46
c_i	[m^{-1}]	the curvature at the pole of surface *i*	36, 40, 52
d	[m]	the lateral displacement of a ray	106
d_i	[m]	the distance between surface *i* and *i*+1	36, 122
D_j		the contribution of layer *j* to the total reflected field	105
d_{var}	[m^2]	the variance of the lateral displacement	106
$E_{i,j}$	[V/m]	the incident electric field at transition *j* in a multilayer	97
E_j	[V/m]	the electric field at transition *j* in a multilayer	97, 98
$E_{r,j}$	[V/m]	the reflected electric field at transition *j* in a multilayer	97
$E_{t,j}$	[V/m]	the transmitted electric field at transition *j* in a multilayer	97
f		the set of operands	73
f_i		operand number *i*	73
g_i		coefficients of basis functions	41
H		the Lagrange invariant	37, 52
h	[m]	the object height	36, 122
H_i	[T]	the magnetic field at transition *i* in a multilayer	98
k	[m^{-1}]	the (circular) wave number	35
k	[m^{-1}]	the wave vector	103, 104
k_1		a process-related factor for the feature size	16
k_2		a process-related factor for the depth of focus	17
l_j	[m]	the thickness of a coating in a multilayer	99
M		the matrix that paraxially relates rays from object to image plane	36
M_a		the angular paraxial magnification	37, 47, 53
M_t		the transverse paraxial magnification	36, 47, 53
N		the number of reflective surfaces in a system	35, 48, 52, 64, 127
NA		the numerical aperture at the image side	16, 37, 53
N_c		the number of times the principal ray crosses the optical axis	48
n_i		refractive index of material between surface *i* and *i*+1	36, 52, 98, 122

N_i		the number of intermediate images	48
n_j		the refractive index of the material between layer j and j+1	99, 103
NO		the numerical aperture at the object side	37, 53
O		the presence of obstruction in a system	26
$O(i)$		the order of the neglected terms	36
p_i	[m^{-1}]	the optical power of surface i	44
$P_{i,f}$		the intersection of mirror i with ray f	25
q		the number of constraints imposed on the system	64
r		the complex amplitude of the reflected field	99, 109
R		the intensity reflection coefficient of a multilayer	99, 102
Rel. R^9		the reflection factor after 9 reflections	102
$\mathbf{R}_i$		the paraxial refraction matrix of surface i	36
r_i	[m]	perpendicular distance to the optical axis	40, 42
s		the number of the stop surface	37
S_i		surface number i in an optical system	122
$s_{i,j}$	[m^{2j}]	the j^{th} order aspherical coefficient of surface i	40
t_g	[s]	the group delay	104
T_i		the i^{th} polynomial of a set of polynomials	41, 43
$\mathbf{T}_i$		the paraxial transfer matrix of surface i	36
U		the amplitude of a wave	35
u_i	[rad]	angle of ray after surface i relative to optical axis	36, 37, 53, 65, 122
$\mathbf{v}$		the normal of the multilayer	103
v_g	[m/s]	the group velocity	104
$\mathbf{w}$		the set of weighting factors	73
$w(t)$		a weighting function	41
w_i		weighting factor number i	73, 106
$\mathbf{x}$		the set of optimization variables of a system	73
y_s	[m]	the y-coordinate of the intersection of a ray with surface s	37
$y_{s,f}$	[m]	the y-coordinate of the intersection of ray f with surface s	25
$Z(r,\psi)$		the Zernike polynomials in radial and azimuthal coordinates	42
z_{eff}	[m]	the effective depth of reflection in a multilayer	104
z_i	[m]	the surface sag of surface i	40, 122
z_j	[m]	the optical distance between top of multilayer to transition j	106
$z_{s,f}$		the surface sag at the intersection of ray f with surface s	25
α_1	[m^2]	rescale factors for the Chebyshev parameter	42
α_2	[m^2]	rescale factors for the Chebyshev parameter	42
α_i	[rad]	the slope of a surface in point P_i relative to the optical axis	122
$\beta_{i,1}$		the height of the escape function for the i^{th} local minimum	77
$\beta_{i,2}$		the width of the escape function for the i^{th} local minimum	77
γ_p		a dimensionless quantity for p-polarized light	99
γ_s		a dimensionless quantity for s-polarized light	99
δ	[m]	the feature size	16
$\delta_{i,j}$		the Kronecker delta function	41
δ_j		complex phase factor in the multilayer calculations	99
δ_R	[m]	the Rayleigh resolution	16
$\mathbf{\Delta x}$		the change vector of the variables	75
ε_i		real constants	41
θ_i	[rad]	the angle of incidence on surface i	35, 40, 46, 99, 103, 104, 111
κ		the conic constant	40

λ	[m]	the wavelength	16
μ		the damping factor of optimization	76
$\boldsymbol{\mu}$		the diagonal damping matrix	76
$\boldsymbol{\nu}$		the normal to the surface	104
τ_i		the Chebyshev coefficients	43
υ		the number of configurations to be evaluated	64
Φ		the error function	73
φ	[rad]	the phase of a field	99, 103, 104, 109
χ	[m]	the effective depth of focus	17, 26
χ_D	[m]	the diffraction determined depth of focus	16
ψ	[rad]	azimuthal polar coordinate	42

9.3 Mathematical notation

a	a column vector
A	a matrix
i	the square root of –1
[*b*,*c*]	a domain
∇	the gradient or first derivative
$a \in [b,c]$	the logical evaluation if *a* is an element of the domain [*b*,*c*]
$\lvert\ \rvert$	the absolute value
*	the conjugate of a complex number
^{T}	the transpose of a matrix
·	the inner product
×	the vector product
∨	the logical summation, *or*
∧	the logical multiplication, *and*
⊕	the logical exclusive or, *xor*
$\bigvee_m$	the logical summation (*or*) of all integer values of *m*
$\bigwedge_m$	the logical multiplication (*and*) of all integer values of *m*
1*e*-5	10^{-5}
$\bar{}$	referring to the principal ray
arg()	the argument of a complex number
Re()	the real part of a complex number
Im()	the imaginary part of a complex number
sign()	the sign of a number, a logical *true* if the number is positive
even()	the logical evaluation if a number is even

10 Appendix

In this Appendix, we present some arbitrary configurations of 6- and mainly 8-mirror systems that are the result of a paraxial to finite ray conversion, but have not been fully optimized yet. Some of these systems have a potential application in extreme ultraviolet lithography. Others, for evident reasons, will not pass a first selection based on e.g. an angle of incidence criterion.

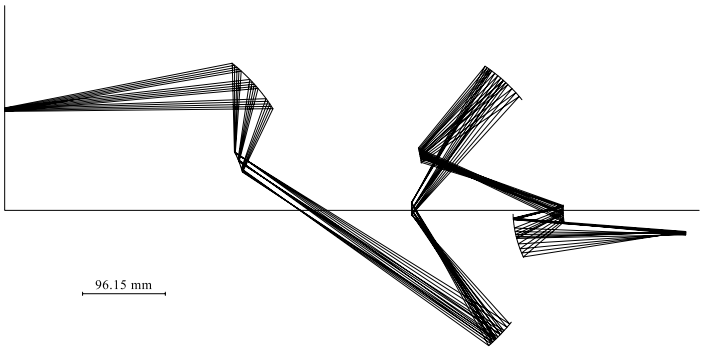

Fig. (10.1) A negative eight-mirror system in class 146-. The root-mean-squared wavefront error is of the order of 26 λ, for a wavelength of 13 nm, a numerical aperture of 0.4, and object heights between 114 and 118 mm.

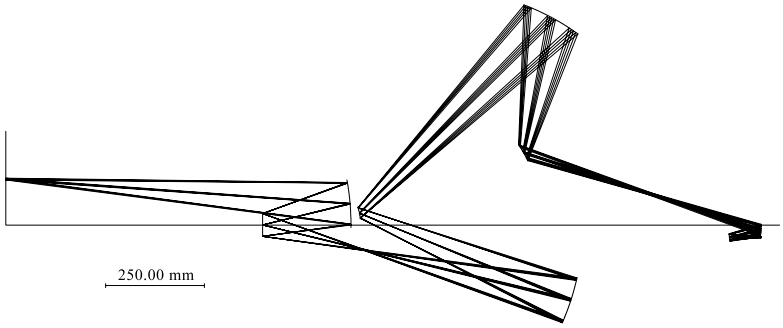

Fig. (10.2) This negative eight-mirror system in class 154-, has a root-mean-squared wavefront error smaller than 0.73 λ when the object heights are between 114 and 118 mm, and the numerical aperture is 0.4.

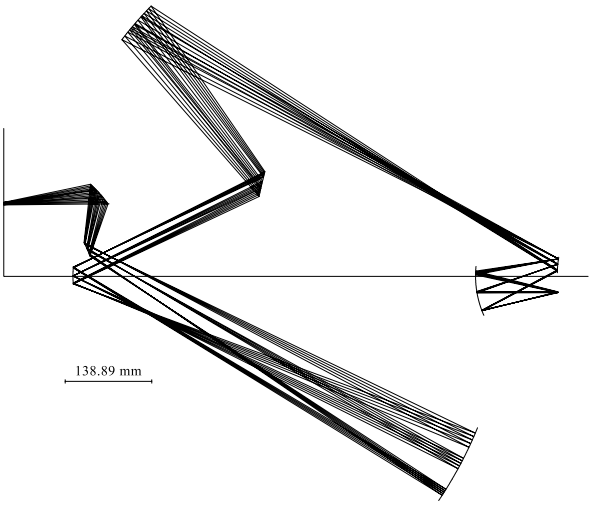

Fig. (10.3) The Strehl ratio of this negative eight-mirror system in the class 182- is larger than 0.924 when the object heights are between 114 and 118 mm.

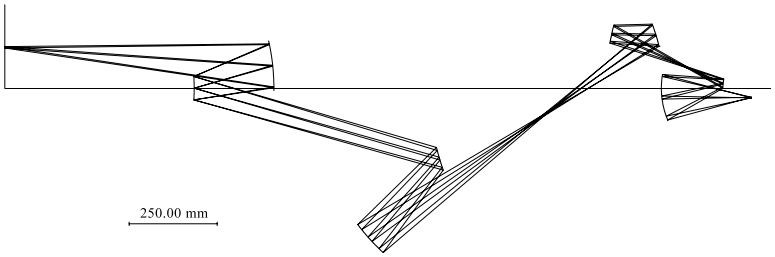

Fig. (10.4) The object heights are between 114 and 118 mm, and the numerical aperture is 0.4, as a result of which this negative eight-mirror system in class 162- has a root-mean-squared wavefront error smaller than 30 λ.

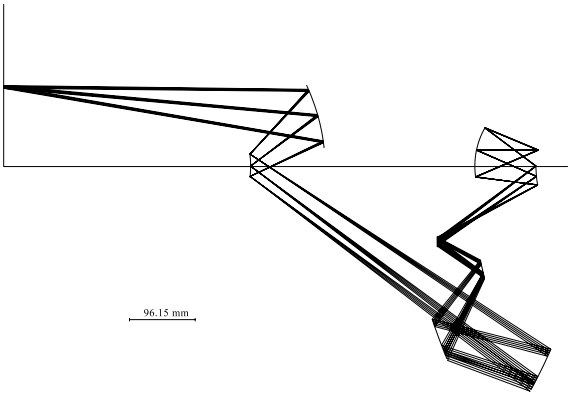

Fig. (10.5) Positive eight-mirror system in the class 181+, with a distortion smaller than 2 nm, object heights between 114 and 118 mm, numerical aperture of 0.4, and a root-mean-squared wavefront error smaller than 0.5 λ.

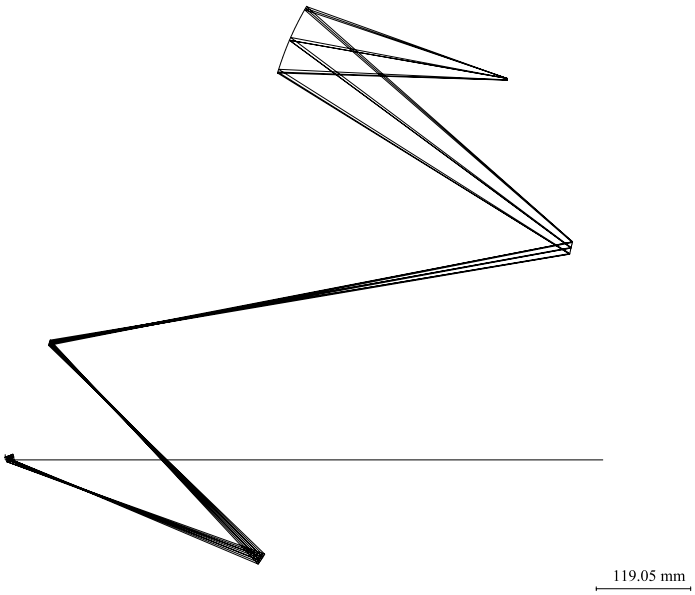

Fig. (10.6) A positive six-mirror system in the class 46+, which could be used in a maskless extreme ultraviolet lithographic system. The magnification of the system is $^{1}/_{160}$, with image heights between 2.95 and 3.05 mm, a numerical aperture of 0.3, and a root-mean-squared wavefront error smaller than 11.2 λ.

10.1 Acknowledgements

For this work, we gratefully acknowledge the use on special conditions of the optical design software programs Oslo and CodeV. CodeV is developed by Optical Research Associates in Pasadena, CA, Oslo by Lambda Research Corporation in Littleton, MA, both in the USA.

11 Biography

Matthieu Frédéric Bal was born on January 6, 1974 in Utrecht. He attended the Lorentz Lyceum in Eindhoven and began his studies in applied physics at the University of Groningen in 1992. He did an internship at the University VII of Paris in the hospital of Broussais under the supervision of prof. P. Flaud, with as subject the propagation of pressure waves in flexible tubes which simulate an artery. The subject of his master's thesis was the measurement of the aggregation of erythrocytes with ultrasound backscattering done under supervision of Dr. J. Lubbers. He started his research on the optical design of extreme ultraviolet lithographic projection systems as a Ph.D. student at the Delft University of Technology in January 1999.

11.1 Publication list

M. Bal, F. Bociort, and J. Braat, "Analysis, search and classification for reflective ring-field projection systems," to be published (2002).

M. Singh, M. Bal, J. Braat, D. Joyeux, and U. Dinger, "Wavefront correction methods for extreme-ultraviolet multilayer reflectors," accepted for publication in: *Applied Optics*.

M. Bal, M. Singh, and J. Braat, "Optimizing multilayer coatings for Extreme UV projection systems", in *3rd International Conference on Optics-photonics Design & Fabrication,* Y. Ichoka ed., (The Optical Society of Japan, Tokyo, 2002), pp. 169-170.

M. Bal, F. Bociort, and J. Braat, "The Influence of Multilayers on the optical Performance of Extreme UltraViolet Projection Systems," in *International Optical Design Conference 2002 Proceedings, Vol. 4832*, J. M. Sasian and P. K. Manhart, eds., SPIE **4832**, to be published Dec. 2002).

M. Bal, F. Bociort, and J. Braat, "Lithographic apparatus and device manufacturing method," EP patent 1 20 95 03 A2 (May 29, 2002).

M. Bal, "EUV projection systems, new eight-mirror systems and the modeling of multilayers," (Technical University of Delft, Delft, 2001).

M. Bal, F. Bociort, and J. Braat, "The effect of obstructions on the design of reflective ring-field projection systems," in *Soft X-Ray and EUV Imaging Systems II*, D. Tichenor and J. Folta, eds. (SPIE, San Diego, 2001), pp. 25-31.

M. Bal, "EUV projection systems, a systematic approach for obstruction free design," (Technical University of Delft, Delft, 2001).

F. Bociort, M. Bal, and J. Braat, "Systematic analysis of unobscured mirror systems for microlithography," in *2nd International Conference on Optical Design and Fabrication (ODF2000)* (The Optical Society of Japan, Tokyo, 2000), pp. 339-342.