University of California
Los Angeles

Analysis and Modeling of Photomask Near-Fields in Sub-wavelength Deep Ultraviolet Lithography with Optical Proximity Corrections

A dissertation submitted in partial satisfaction
of the requirements for the degree of
Doctor of Philosophy in Electrical Engineering

by

Jaione Tirapu Azpiroz

2004

The dissertation of Jaione Tirapu Azpiroz is approved.

Tatsuo Itoh

Stanley Osher

Yahya Rahmat-Samii

Eli Yablonovitch, Committee Chair

University of California, Los Angeles

2004

Table of Contents

List of Figures

List of Tables

Acknowledgments

(Acknowledgments omitted for brevity.)

Vita

1975	Born, Pamplona, Navarre, Spain.
1996–1997	Erasmus Academic Exchange Student, Electric and Electronic Engineering Department, University of Surrey, United Kingdom.
1999	M.S. (Telecommunications Engineering), Public University of Navarre, Spain.
1999–2000	Trainee, European Space research and TEchnology Centre (ESTEC), European Space Agency (ESA), The Netherlands.
2000–2002	"La Caixa" Graduate Fellowship Awardee, Electrical Engineering Department, UCLA.
2002–2003	Research Assistant, Electrical Engineering Department, UCLA.
2003–2004	Recipient of the "Dissertation Year Fellowship" by the UCLA Graduate Division.

Publications

"Fast evaluation of Photomask Near-Fields in Sub-Wavelength 193nm Lithography". J. Tirapu-Azpiroz and E. Yablonovitch. *Proceedings of the SPIE, Optical Microlithography XVI*, Vol. 5377, pp:00-00 (22-27 February 2004).

"Modeling of Near-Field effects in Sub-Wavelength Deep Ultraviolet Lithography". J. Tirapu-Azpiroz and E. Yablonovitch. To Appear in *Future Trends of Microelectronics 2003*. Wiley Interscience.

"Boundary Layer Model to Account for Thick Mask Effects in Photolithography". J. Tirapu-Azpiroz, P. Burchard and E. Yablonovitch. *Proceedings of the SPIE, Optical Microlithography XVI*, Vol. 5040, pp:1611-1619 (23-28 February 2003).

"Method of Moments Analysis of 2 Dimensional Photonic Bandgap Structures fabricated via Dielectric Cylinders". Tirapu-Azpiroz, Jaione. ESTEC Working Paper No. 2104. TOS-EEA June 2000. European Space Agency.

"Study of the Delay Characteristics of 1D Photonic Bandgap Microstrip Structures". Tirapu J., Lopetegi T., Laso M.A.G., Erro M. J., Falcone F. and Sorolla M. *Microwave and Optical Technology Letters*, vol.23, (no.6), Wiley, 20 Dec. 1999. p.346-349.

"Dispersion cancellation using 1-D Photonic Bandgap Microstrip Structures in Transmission". Tirapu J., Lopetegi T., Falcone F. and Sorolla M. 24th *International Conference on Infrared and Millimeter Waves*, September 5-10, 1999, Monterey, California.

Abstract of the Dissertation

Analysis and Modeling of Photomask Near-Fields in Sub-wavelength Deep Ultraviolet Lithography with Optical Proximity Corrections

by

Jaione Tirapu Azpiroz

Doctor of Philosophy in Electrical Engineering

University of California, Los Angeles, 2004

Professor Eli Yablonovitch, Chair

The utilization of 193nm wavelength lithography with a 0.85 NA to print 65nm wafer features translates into a k_1 factor approaching values around 0.3 and mask features of the order of the wavelength for $4X$ magnification. In addition, Alternating Phase-Shifting Masks (Alt. PSM) employ etching profiles with abrupt discontinuities and trench depths also in the order of the wavelength for 180^o phase-shifting openings. As a consequence of wavelength sized and high aspect ratio mask features, mask topography effects are becoming an increasing source of simulation errors, which are particularly critical for Alternating Phase-Shifting Masks, and demand rigorous resource-consuming 3D electromagnetic field simulations in the sub-wavelength regime.

Conventional application of Kirchhoff's Boundary Conditions on the mask surface provides the so-called "Thin Mask" approximation of the object field. Sub-wavelength lithography, however, place a serious limitation on this approximation that fails to account for the topographical or "Thick Mask" effects. A new simulation model is proposed, which is based on a comparison of the fields

produced by both the thick and ideal thin masks on the wafer. The key result of our simulations is that the thick mask effects can be interpreted, to a good approximation, as an intrinsic edge property, and modeled with just two fixed parameters: width and transmission coefficient of a locally-determined boundary layer, applied to each chrome edge.

The Boundary Layer model (BL model) is theoretically founded on the well-established Physical Theory of Diffraction (PTD). According to this theory, the total scattered field by a metallic object is constructed by adding a "fringe" field, generated by electric and magnetic equivalent edge currents along the edges of the scatterer, to the Kirchhoff's or Physical Optics approximation. We observed how the relative errors of the field real and imaginary components on the wafer follow an inverse law on the opening mean size and height, respectively, what allowed us to reduce the model to a simple boundary layer of fixed width and transmission coefficient. The proposed model, therefore, consists of a sophisticated version of Kirchhoff approximation, simply adding a boundary layer to every edge.

The BL model accurately accounts for thick mask effects of the fields on the mask, incorporating effects of electromagnetic coupling due to the high numerical aperture ≥ 0.7, and accurately compensates for phase errors even at planes out of focus. This greatly improves the accuracy of aerial image computation in photolithography simulations at a reasonable computational cost.

CHAPTER 1

Introduction

Optical Lithography represents the main manufacturing technology of today integrated circuits. Over the last three decades, lithography has been instrumental to the historical trend, better known as Moore's Law [1], of doubling chip density roughly every eighteen months while maintaining a nearly constant chip price. Nowadays *photolithography* continues to enable this steady miniaturization of the wafer Critical Dimensions below the illumination wavelength through the utilization of ingenious techniques of resolution enhancement [2, 3].

Transistors form the basic units of integrated circuits fabricated on silicon wafers. They are connected together to implement more complex functional units, such as inverters or adders which, once interconnected according to an optimized design, can perform complicated tasks. The number of such transistors on a circuit is heading toward 1 billion and beyond [4], and its critical dimension is, according to the 2003 International Technology RoadMap for Semiconductors (ITRS) [5], about 90nm at the time this thesis was written.

The schematic representation of the circuit elements must be translated into the set of geometrical shapes that need to be deposited onto the silicon substrate, distributed in several material levels, to create and physically connect these devices. The physical pattern of each separate level is etched on a photomask using either electron or optical beam writers, and then transferred repeatedly to the wafer by lithography.

Prior to being exposed with the mask pattern, the silicon substrate is coated with a layer of photoresist. Photoresists are materials, usually organic polymers, that undergo photochemical reactions when exposed to light. After the mask pattern has been projected onto the resist by optical lithography, the selected areas are removed with a developer solution. Hence the desired circuit pattern has been replicated on the resist film and the wafer can now undergo etch or ion implantation operations. This fabrication cycle may be repeated as many as 30 times to shape the different layers that comprise an integrated circuit.

1.1 The Lithography Process

In the fabrication process briefly described above, photolithography represents the most critical step in the determination of a circuit smallest dimension. Wafer steppers are the primary tools to perform the imaging process. They can be found in two configurations: "step and repeat" and "step and scan". Figure 1.1 shows one advanced stepper model from ASM Lithography, the 193nm Step & Scan ASML 5500/950B, which incorporates the most advanced imaging technologies [6].

In a "Step and Repeat" configuration, the entire mask field is imaged at once, although only a portion of the wafer is being exposed. Once the exposure dose is reached, the wafer is moved and the operation repeated until the total wafer area has been exposed. The image field size is therefore limited by the largest lens field size of sufficient imaging quality. In a "Step and Scan" configuration, the lens field size does not cover the entire reticle area and only part of the mask pattern is exposed when the light shutter is opened. The mask and wafer are then scanned with accurate synchronization until the entire mask has been projected onto the wafer. This technique allows to increase the image field size with the same lens

field size, thus eliminating the need for larger and more expensive lenses.

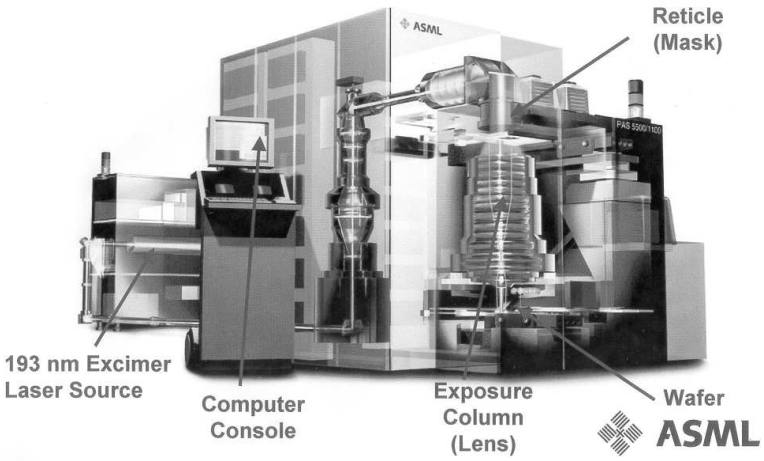


Figure 1.1: 193nm Step & Scan ASML 5500/950B from ASM Lithography (Courtesy of ASML[6].)

The lithography process can be divided in the four modules illustrated in figure 1.2. Those are the illumination system, comprising a light source and the condenser optics, the photomask, the projection optics and the resist-coated wafer. In his monograph, Levinson [7] covers in detail the theory and practical aspects of every step of the lithographic process.

1.1.1 Illumination Configuration

The illumination system supplies a highly monochromatic beam of light of high and uniform intensity. Monochromatic light is important because refractive lenses can be designed to operate nearly aberration-free, but only over a very narrow bandwidth of wavelengths. Moreover, intense illumination guaranties high throughput since the necessary dose of exposure per wafer is reached faster. Possible light sources for optical lithography must be very intense and of very narrow bandwidth, what determines the available wavelengths of operation. Table 1.1

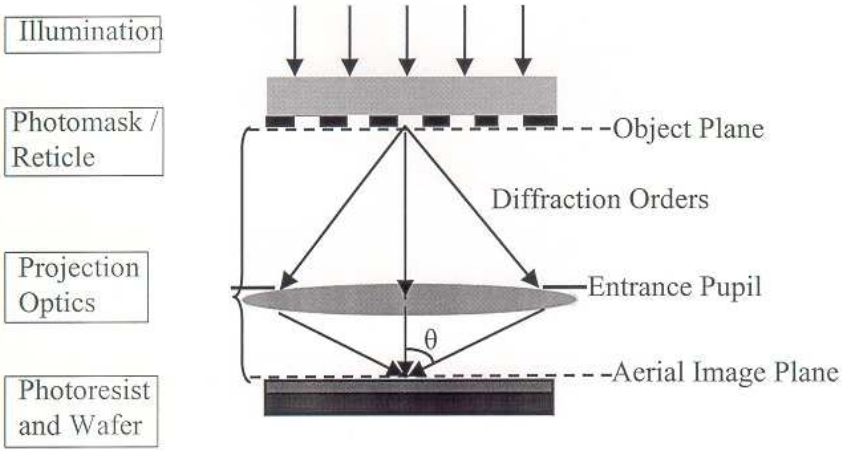


Figure 1.2: General optical lithography process diagram.

provides the list of light sources with applications in lithography. In the mid 90s, lithography steppers started using Deep UltraViolet light at 248nm, and continue today at both 248nm and 193nm. 157nm lithography (Vacuum Ultraviolet) is nowadays under development.

Table 1.1:

Common light sources at different lithographic wavelengths

WAVELENGTH	LIGHT SOURCE	YEAR OF INTRODUCTION
436 nm	Hg arc lamp (g-line)	1970
365 nm	Hg arc lamp (i-line)	1984
248 nm	Krf excimer laser	1989
193 nm	ArF excimer laser	1999
157 nm	F_2 excimer laser	after 2004

Uniformity of the illumination intensity on the object is engineered by the condenser optics configuration, which also establishes the amount of spatial coherence of the light and performs various forms of spectral filtering. Köhler

configuration [8] predominates in lithography because it provides uniform illumination from a source that in general is non-uniform, provided well-corrected condenser lenses are employed [9]. In Köhler illumination, each point of the source originates a coherent, linearly polarized plane wave emerging from the lens with an angle determined by the source point location relative to the optical axis [8]. This configuration is further discussed in section 4.2.

1.1.2 Reticle

Although the terms Mask and Reticle named two different things in the past, they are nowadays used interchangeably. Both names refer to a substrate layer of glass, usually fused silica at DUV wavelengths, covered by a film of chromium of thickness between 50nm and 110nm to provide good absorption of the incident light. It carries the pattern that will be transferred to thousand of wafers, thus its quality and durability are of critical importance since defects on the mask will be reproduced on the wafer.

Common mask patterns are two dimensional contacts, as well as arrays of lines and spaces that alternate transparent areas with opaque regions covered with chrome. Depending on the type of photoresist, contact holes on the wafer will be produced by either apertures on the chrome layer or opaque chrome features surrounded by glass. Diffraction at the chrome edges degrades the opacity of the chrome feature in the later case, and bright holes on the reticle are usually preferred to produce contacts.

1.1.3 Projection Optics

As can be seen in figure 1.1, the imaging system consists of a complex setup of 25-40 glass elements providing a reduction factor of $4X$ or $5X$. Most stepper

lenses are refractive, of up to one meter length and some 500Kg weight, firmly held in a concentric manner inside the stepper. Higher reduction factors would mitigate the sensitivity to mask defects, but it would also shrink the size of the field illuminating the wafer below the chip size (24.6mm x 24.6mm).

Due to the wave nature of the light illuminating the reticle, diffraction effects at the chrome edges are inevitable. Based on Huygens principle [8], diffraction theory shows that the angular distribution of the field diffracted by the photomask is, after propagating a distance equivalent to a few wavelengths, proportional to the spatial Fourier transform of the mask field distribution. This means that low-spatial-frequency components, which arise from large mask features, propagate with small diffraction angles with respect to the reticle normal while high-spatial-frequencies, corresponding to small mask features, propagate at large angles relative to the reticle normal. Only those spatial frequencies collected by the entrance pupil will be recombined by the projection optics and imaged onto the resist surface. Higher frequency components are filtered out by the lens. The final aerial image is therefore a partial reproduction of the original pattern. As a consequence, diffraction effects limit the ultimate resolution of the imaging system.

High resolution image formation then relies on nearly diffraction-limited imaging characteristics of refractive lenses at the illumination wavelength. This requires very high quality fused silica and strict design specifications, what is raising the fraction of the cost represented by the lens relative to the total stepper.

1.1.4 Photoresist

Photoresists are available in two main flavors, positive or negative. Positive resists become soluble in developer solution upon exposure of light, while negative resists

lose their solubility on those areas exposed to light. This means that in order to produce contact holes on positive resists, the mask feature consists of an aperture on the chrome layer. On the other hand, in order to produce a contact hole on a negative resist, one needs a mask consisting of an opaque chrome feature surrounded by glass. Due to diffraction effects, the aerial image produced by openings on the mask surface is of better quality than the image produced by opaque chrome features and, therefore positive resists are preferred in practice [7]. These are the type of mask patterns analyzed in most of this thesis.

Regardless of the quality of the aerial image, poor resist contrast can degrade the resolution attainable with a particular lithography system. Resist contrast depends on the resist material as well as on the resist process parameters, which need to be carefully monitored.

1.1.5 Technology Node

Since first predicted in 1965, Moore's Law [1] has driven the semiconductor manufacturing industry into a pace of one new technology node every two years. Historically technology nodes have been associated with the introduction of Dinamic Random Access Memory (DRAM) chips with the smallest metal half-pitch. Recent advances on Microprocessor units with gate lengths smaller than the DRAM half-pitch provide new parameters to characterize the technology node. In an attempt to set industry standards, the International Technology RoadMap for Semiconductors (ITRS) [5] defines technology node as "the minimum metal pitch used on any product, either DRAM half-pitch or Metal 1 (M1) half-pitch in Logic/MPU (MicroProcessor Unit)". As of 2003, technology nodes continue to be associated with the introduction of DRAM chips with the smallest metal half pitch.

Roughly every two years, this technology roadmap is released as a consensual guidance for the research and development efforts of the semiconductor industry in the next 15 years. Table 1.2 shows a small sample of the technology requirements and introduction time estimates until 2009 as predicted by the 2003 edition of the ITRS.

Table 1.2:

International Technology Roadmap for Semiconductors (ITRS), 2003 Edition

TECHNOLOGY NODE	2003	2004	2005	2006	2007	2008	2009
DRAM							
Half Pitch (nm)	100	90	80	70	65	57	50
MPU							
Half Pitch (nm)	120	107	95	85	76	67	60
Printed Gate length(nm)	65	53	45	40	35	32	28
Physical Gate length (nm)	45	37	32	28	25	22	20
MASK MINIMUM FEATURES							
Nominal image size (nm)	260	212	180	160	140	128	112
OPC clear feature (nm)	200	180	160	140	130	114	100
OPC opaque feature (nm)	130	106	90	80	70	64	56

1.2 Process Parameters

1.2.1 Numerical Aperture

The numerical aperture (NA) of a lens is defined as:

$$NA = nsin\theta \tag{1.1}$$

where n is the diffraction index of the medium surrounding the lens and θ is the angle of the lens acceptance cone illustrated in figure 1.2. The space at the entrance side of the lens is called object space, and the space at the exit side of the lens is called image space. The numerical apertures on both spaces are related through the magnification of the system, M:

$$M = \frac{NA_o}{NA_i} \tag{1.2}$$

which usually takes values $M = \frac{1}{4}$ or $M = \frac{1}{5}$ for $4X$ or $5X$ reduction systems, respectively. Values of NA_i larger of 0.7 are nowadays of common use in optical lithography

1.2.2 Resolution

Resolution of an optical projection system is determined by the size of the minimum resolvable feature and it is limited by the finite numerical aperture of the imaging lens entrance pupil. Theoretically an isolated feature produces a continuous spectrum of spatial frequencies such that some of them always pass through the entrance pupil to the image space. Therefore, isolated features can always be resolved regardless of their arbitrarily small size. Periodic patterns such as gratings diffract a finite and discrete set of spatial frequencies at intervals $\triangle k = \frac{2\pi}{p}$, where k denotes the spatial wavevector and p the grating period, also

known as the grating pitch. These spatial frequencies are commonly known as the reticle diffraction orders, and they diverge from the object plane at discrete angles determined by $sin\theta_m = m\frac{\lambda}{p}$, with $m = 0, \pm1, \pm2...$ The interference of at least two diffraction orders is needed to create enough spatial variation to resolve the grating image. Thus the grating period p must be large enough to allow for two diffraction orders to be collected by the lens entrance pupil. The smallest resolvable grating pitch is therefore determined by the ratio of the illumination wavelength to the numerical aperture of the projection optics:

$$p_{min} = \frac{\lambda}{NA} \tag{1.3}$$

The above discussion assumed spatial coherent illumination. However, as indicated in section 1.1.1 and further discussed in section 4.2, the illumination system establishes certain amount of partial coherence determined by the factor σ, which will be defined in section 1.2.4. Under partial coherent illumination the minimum resolvable grating pitch is given by [2]:

$$p_{min} = \left| \begin{array}{lll} \frac{\lambda}{NA} & for\ \sigma = 0 & \text{coherent illumination} \\ \frac{1}{1+\sigma}\frac{\lambda}{NA} & for\ 0 < \sigma < 1 & \text{partial coherent illumination} \\ \frac{1}{2}\frac{\lambda}{NA} & for\ \sigma = \infty & \text{incoherent illumination} \end{array} \right. \tag{1.4}$$

Rayleigh's resolution limit was derived as an arbitrary criterion for the minimum distance between two stars resolvable by a telescope. This criterion can be associated to the imaging of contact holes. By considering two mutually incoherent point sources, this distance is limited by the finite size of the entrance pupil in a similar fashion to equation (1.3):

$$d_{min} = 0.61\frac{\lambda}{NA} \tag{1.5}$$

Equations (1.3) to (1.5) all represent theoretical limits since they were derived assuming that the resolution is limited only by diffraction. In order to incorporate

the effect of the resist contrast and other process parameters, a k_1 factor is defined for each process to describe its resolution capabilities:

$$CD = k_1 \frac{\lambda}{NA} \tag{1.6}$$

where CD represents the system Critical Dimension printable by the system. For a grating formed by equal lines and spaces, the CD refers to the half-pitch, $\frac{p_{min}}{2}$, with p_{min} defined by equation (1.4). According to this equation, the minimum theoretical value of k1 factor is 0.25, although it is limited in practice to values of 0.3 or larger.

1.2.3 Depth of Focus

Images observed on planes out of focus become blurred and loose definition. The range of distances about the focal plane over which the image is adequately sharp according to certain specifications is defined as the Depth of Focus (DoF). This quantity can be seen to be governed by the following expression [8]:

$$Depth\ of\ Focus = DoF = \pm\ k_2 \frac{\lambda}{NA^2} \tag{1.7}$$

commonly referred to as the Rayleigh Depth of Focus.

Values of DoF encountered in lithography are of the order of a $\pm 0.4 \mu m$. Increasing the NA in an attempt to improve resolution results in a rapid degradation of the Rayleigh DoF. Furthermore, the photoresist has a finite thickness of the order of the DoF $(0.3 - 0.8 \mu m)$, thus the position of the plane of best focus must be controlled accurately to provide good image quality throughout the resist thickness. Depth of focus is another important factor in the resolution equation.

1.2.4 Partial Coherent Factor

The set of plane waves incident on the mask under Köhler illumination correspond to points in the spatial frequency space or k-space and can be conveniently visualized as diagrams. Figure 1.3(a) shows the diagram of a circular source of partially coherent light, modelled as point sources with direction cosines (p_x, p_y) propagating within the condenser Numerical Aperture. The plane wave incident on the mask due to each of these point sources is expressed in its phasor form as:

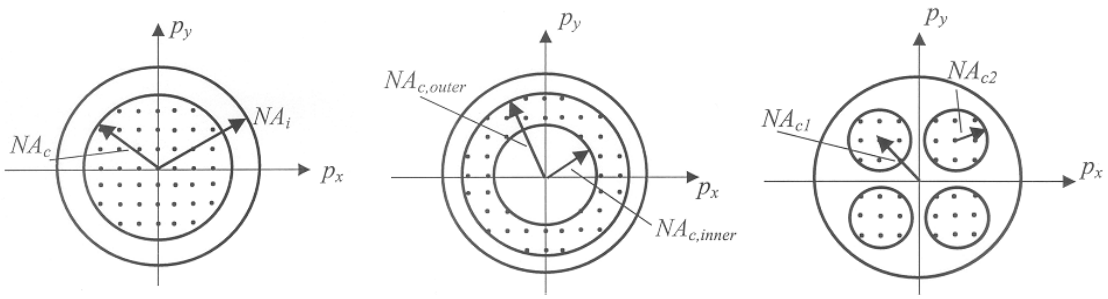
$$\mathbf{E}(x, y, z) = \mathbf{E_0} e^{-jk(p_x x + p_y y + p_z z)} \tag{1.8}$$

where $\mathbf{k} = (k_x, k_y, k_z) = k(p_x, p_y, \sqrt{1 - p_x^2 - p_y^2})$ is the wavevector.

The partial coherence factor for circular sources, σ, is a measure of the spatial extension of the light source and is defined as the ratio of the condenser lens numerical aperture, NA_c, to the imaging lens numerical aperture in the object space, NA_o:

$$\sigma = \frac{NA_c}{NA_o} \tag{1.9}$$

This factor can be adjusted to enhance the resolution of specific mask patterns. Dense periodic patterns benefit from large values of σ, while small values of σ provides better image quality with sparse or nearly isolated features [2, 7]. Large partial coherent factors also help reduce proximity effects although at the expense of image contrast, and common values of sigma range from 0.3 to 0.8. Circular light sources provide directional uniformity and guaranty the same printing quality for features in all orientations. More advanced illumination schemes, however, such as the annular or the quadrupole, with source diagrams as sketched in figure 1.3(b) and (c), respectively, can further improve image fidelity of mask patterns with specific symmetry at the expense of some directional uniformity.



(a) (b) (c)

Figure 1.3: Common modified source schemes for advanced illumination used in optical lithography. Source discretization is performed following a Cartesian distribution of grid points. (a) Circular illumination. (b) Annular illumination. (c) Quadrupole illumination.

1.3 Resolution Enhancement Techniques

1.3.1 Sub-Wavelength Lithography

Advances on optical lithography equipment and technology, as well as a progressive shortening of the exposure wavelength, as indicated by table 1.1, have been simultaneously pursued in an attempt to reduce the minimum printable feature. Refractive lenses with the required transparency and quality at 248nm wavelength can only be made of fused silica. Calcium fluoride (CaF_2) can also be used in combination with fused silica at 193nm, what can improve the operating bandwidth, and it represents the principal candidate at 157nm. The lack of transparent optical components at shorter wavelengths limits the available wavelengths in Deep Ultraviolet lithography, which is now being performed within the sub-wavelength regime. This means that the minimum feature of the wafer circuit is smaller than the wavelength of the light source used to print it.

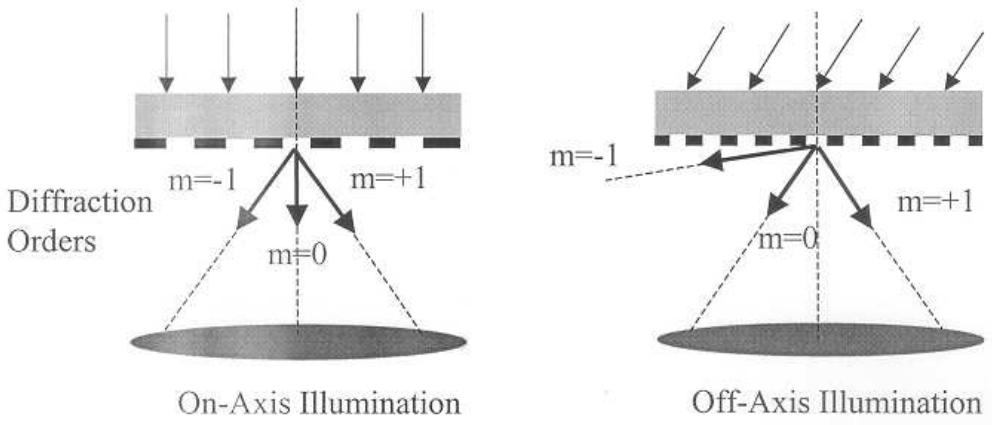
According to the International Technology Roadmap for Semiconductors (ITRS) as of its 2003 edition in table 1.2, feature sizes of 90nm half pitch were expected to

start being manufactured by 2004. Semiconductor industry, however, is running ahead of schedule and started producing 90nm wafer features by January 2003 [4], employing 193nm lithography and numerical apertures as high as 0.75, and plans to start high volume manufacturing of the 65 nm node in 2005 [10].

The utilization of 193nm wavelength lithography with a 0.85 NA to print 65nm wafer features translates into a k_1 factor approaching values around 0.3 and mask features of the order of the wavelength for $4X$ magnification. Several techniques of resolution enhancement (RETs) have been developed and are being increasingly employed, together with imaging systems of higher numerical aperture (NA), to overcome the limits of optical lithography. A thoughtful description of these techniques can be found in Wong's monograph [2], and selected papers on the topic have been collected in a volume by Schellenberg [3]. A brief description of some of these techniques is provided in sections 1.3.2 to 1.3.4.

1.3.2 Off-Axis illumination

When on-axis illumination is used on periodic gratings, figure 1.4(a), both diffraction orders with indexes +1 and -1 need to be collected by the lens to produce interference. In addition the 0^{th}-order is also collected, which carries no frequency information and contributes only with DC background. By illuminating at an off-axis angle, figure 1.4(b), the image can be formed by interference of the 0^{th} diffraction order and either one of the ± 1 orders. This increases resolution by allowing more separation distance between the first two diffraction orders of periodic gratings, that is, between the orders 0^{th} and ± 1. Some examples of advanced illumination schemes were introduced in section 1.2.4 and plotted in figure 1.3. They can be achieved by introducing an aperture between the light source and the condenser optics.



(a) (b)

Figure 1.4: Schematic of (a) On-axis illumination and (b) Off-axis illumination.

1.3.3 Optical Proximity Correction

Even though not all corrections applied to the mask shapes are necessarily to account for effects of the proximity of other features, it is customary to refer to all adjustments of the mask features compensating for low k_1 effects as Optical Proximity Corrections (OPC).

One simple type of optical proximity correction consist of varying the size of the mask feature depending on its initial dimension and the position of the nearest features. This technique is known as line biasing and is illustrated in figure 1.5(a).

Off-axis illumination schemes help enhance the resolution of dense features, but do not improve the imaging of isolated features. Another type of OPC can be applied to sparse features to simulate a dense environment. These are called Scattering Bars or Assist Features and they are sub-resolution geometries that do not print on the wafer because all the high spatial frequencies are filtered out by the lens. Assist features as those of figure 1.5(b), in combination with off-axis

illumination, can improve the imaging of sparse geometries.

Modifications of the original mask pattern with the addition of geometries such as Hammer Heads, illustrated in figure 1.5(c), and Serifs, illustrated in figure 1.5(d), can compensate for corner rounding and line shortening.

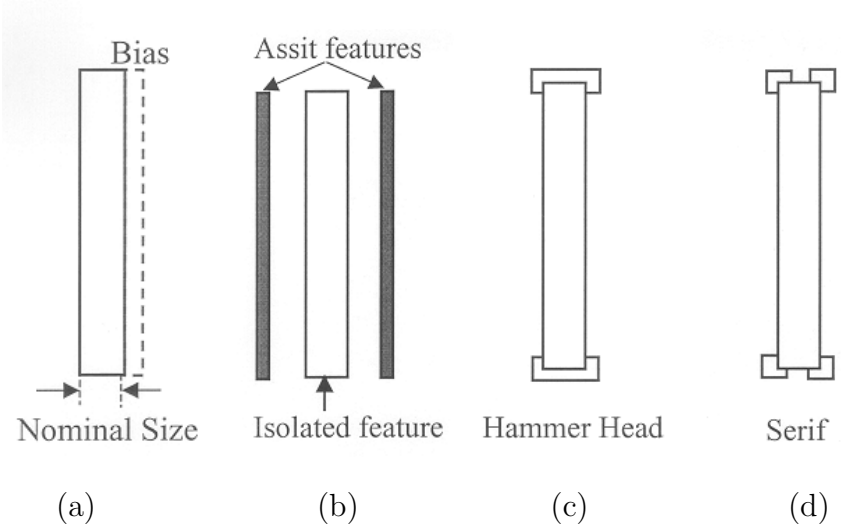


Figure 1.5: (a) Line Biasing, (b) Scattering Bars, (c) Hammer Heads and (d) Serifs.

1.3.4 Phase-Shifting Masks

Phase-shifting masks (PSM) modulate both amplitude and phase of the electromagnetic field propagating through them and improve image contrast by inducing destructive interference of the fields with opposite phases. Masks are commonly classified into Binary and Phase-Shifting depending on whether the transmission coefficient through the mask takes only the values 0 and 1, or a certain amount of phase shift is introduced. Several versions of Phase-Shifting masks exist, among them the Alternating PSM and the Attenuated PSM.

Levenson's Alternating PSM [11] introduce 180^o phase difference through two contiguous apertures on the mask by etching the glass behind one or both open-

ings with a depth difference equivalent to 180^o:

$$d_{180^o} = \frac{\lambda}{2(n_{glass}(\lambda) - n_{air}(\lambda))} \tag{1.10}$$

The index or refraction of the glass at DUV wavelengths is about $n_{glass} = 1.5$, and that of air is $n_{air} = 1$, what results in etching depths of the order of the wavelength.

The principle of operation of both binary and Alternating Phase-Shifting masks is compared in figure 1.6(a) and (b), respectively. Phase-Shifting masks concentrate light diffracted by dense patterns into the most oblique components within the lens NA, enhancing fine features of the image and reducing the DC background.

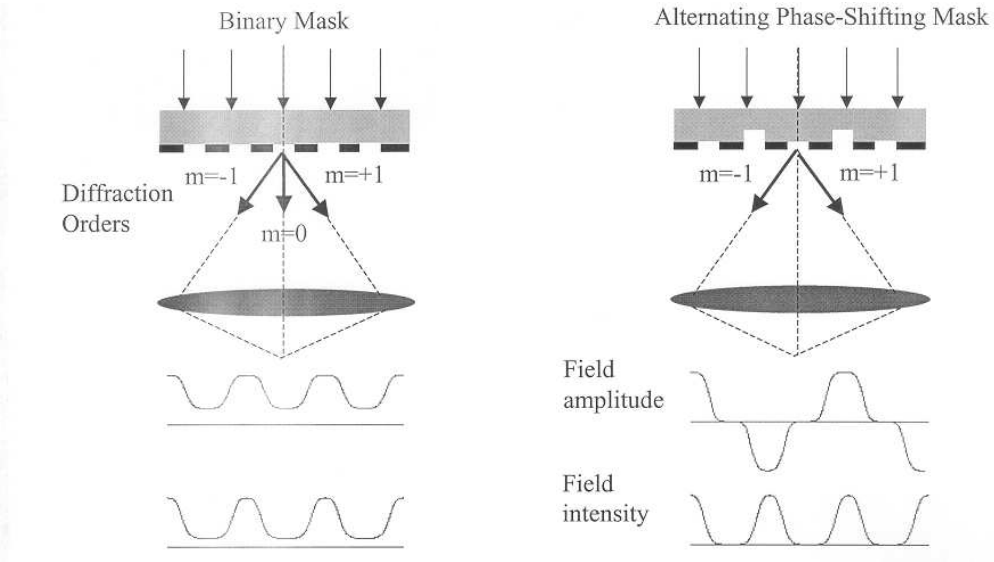


Figure 1.6: (a) Binary Mask operation and (b) Phase-Shifting Mask operation.

Attenuated PSM [12] allows partial transmission through the chrome with a phase difference of 180^o with respect to the transmission through the clear

openings. Destructive interference between fields enhances the image contrast on the contours.

1.3.5 Immersion Lithography

Despite the acceleration of the technology node roadmap, development of the 157nm lithography technology replacement is being delayed due to technical difficulties regarding the quality of the lens calcium fluoride material and challenges with the 157nm resists [13]. It is uncertain whether it will be ready to support the 65nm requirements. As a consequence, the 193nm tools will be used for the critical layers of the 90nm and 65nm generations, and may be extended for the 45nm node by means of strong Resolution Enhancement Techniques such as OPC and Alt. PSM, and systems of higher numerical aperture (up to or larger than 0.85).

Furthermore, there is and increased interest in the so-called "immersion lithography" as the enabling technology to further extend the limits of the existing lithography [14, 15]. In immersion lithography a liquid medium, in principle water, fills the space between the front lens at the exit pupil and the photoresist. The index of refraction of the medium surrounding the lens is therefore larger than 1 and, according to equation (1.1), the effective numerical aperture of the imaging system can be made larger than unity. This technology can provide the necessary resolution enhancement without the reduction of the DoF associated with any increase of the physical NA, however it presents numerous challenges that yet need to be addressed.

1.4 Modeling of the Lithography Process

Accurate modeling and efficient simulation of the lithography process are critical parts of the Integrated Circuit fabrication cycle.

1.4.1 Modeling of the Illumination System

Two equivalent methods, both based on the spatial discretization of the source into a number of spatially incoherent point sources, are usually utilized in lithography to model imaging with partially coherent illumination. The light source is generally engineered to guarantee that the illumination produced by two distinct source points is mutually incoherent. (Since a laser is used in the most advanced steppers, rather than a thermal source, special methods are needed to guarantee spatial incoherence.) In the *Source Integration or Abbe's Method* [16, 17], the coherent images generated by each source point are incoherently added together to produce the final partial coherent image. In the equivalent *Transfer Cross Coefficient or Hopkins Method* [8, 18], the integration over the source is carried out first and the result provides directly the aerial image intensity distribution generated by the partially coherent light. These two methods are further analyzed in section 4.2.

1.4.2 Reticle Electromagnetic Field Evaluation

In addition to mask features of the order of the wavelength, alternating Phase-Shifting Masks (Alt. PSM) employ etching profiles with abrupt discontinuities and trench depths also in the order of the wavelength for 180^o phase-shifting openings. As a consequence, mask topography effects are becoming an increasing source of simulation errors, which are particularly critical for Alternating

Phase-Shifting masks [10, 19, 20, 21]. In practice, however, the computational cost of evaluating Maxwell equations on even small mask areas is too high and Kirchhoff Boundary Conditions have been traditionally assumed. The Kirchhoff approximation replaces the mask by an ideal binary transmission function, the Thin Mask model, which neglects polarization and transmission errors of the real mask.

It is the focus of chapters 3 and 4 of this thesis to introduce and validate an advanced modeling philosophy, capable of incorporating topographical effects and polarization dependencies of the field transmitted by the photomask, but retaining the efficiency of Kirchhoff's approximation. This model replaces the thick mask with the customary thin mask, adding only a fixed-width, locally-determined boundary layer to every edge. Boundary layers are already employed in industry to account for the losses in peak intensity of the field traveling through small apertures in the chromium mask, but always in the form of a bias, that is, an opaque boundary layer. In contrast, our imaginary boundary layer model added to the Kirchhoff approximation allows modeling of thick mask effects, different polarizations, and accounts for phase errors on the aerial image by permitting a complex transmission coefficient in the boundary layer area.

Alternative modeling methods have also been studied in literature [22, 23, 24]. Lam and Neureuther's recent "Domain Decomposition Method" [22] employs pre-calculated diffracted fields from isolated edges that are added afterwards according to the diffracting patterns.

Yan's approach [23] shares with our boundary layer model the possibility of locally modeling topographical mask effects with a boundary band of different transmission coefficient at the thin mask edges. Yan's approximates the diffraction effects on the edges of Extreme Ultraviolet Lithography (EUVL) infinite

lines of width 2.23λ, by adding a strip to the thin mask model. However in his approach, the width and transmission coefficients of the boundary layer were obtained by matching the diffraction ripples of the near field evaluated on the mask surface. For the complex DUVL transmission masks analyzed in this thesis, we performed a systematic study of rectangles of different aspect ratios and sizes, and selected the boundary layer parameters to optimize the fidelity of the central field amplitude on the wafer, not the mask.

In the DUV regime, Adam and Neureuther [24] followed an approach that replaces the rigorous em field on the aperture by a *"scalar complex mask transmission function"* that best matches the complex diffraction pattern of the near field in its lowest spatial frequencies. Their procedure speeds up the calculations while providing good accuracy, but it needs to be re-evaluated for square or rectangular features of different dimensions and aspect ratios.

In references [22, 23, 24] the emphasis was placed on matching the near field of the electromagnetic waves, as they propagate past the edges of the mask. Most of those detailed near field features never survive transmission through the lens, nor do they have any effect on the final image in the photo-resist. Thus in [22, 23, 24] design freedom is expended un-necessarily on matching the exact near field diffraction ripples, that may be of no consequence. In our approach we adjust the OPC corrections to match the final pattern after it has propagated through the lens. In addition we deal with the issue of the mutual interaction between mask edges, by verifying that the boundary model is reasonably successful for rectangular openings of all different sizes and aspect ratios. In references [22, 23, 24] only isolated edges, or a single pair of edges was considered

Full 3D electromagnetic simulations can be performed following different methods. The Finite-Different Time-Domain method based on Yee's algorithm [25, 26]

was employed throughout this thesis to rigorously evaluate the fields on the mask surface. In particular, we utilized the computer program TEMPEST 6.0, developed at the Advanced Lithography Group in the Electrical Engineering and Computer Science Department of the University of California, Berkeley [27].

An integral equation method, based on the Dyadic Green's function for stratified media [28], has also been proposed to calculate the field transmitted by the photomask [29], and even a Wavelet-based method for "fast and rigorous" calculations [30]. Any of these 3D methods, however, are extremely resource consuming and difficult to apply to large areas of the photomask.

1.4.3 Formulation of the Imaging System

Scalar Diffraction Theory, valid for NA up to about 0.4 [31] assumes separable field components, all parallel to one another and to the polarization of the light source (paraxial approximation). It also ignores polarization effects and coupling between electromagnetic components.

By eliminating the paraxial approximation, thus accounting for oblique propagation of the diffraction orders, scalar diffraction theory can be extended to NA about 0.7 [31, 32]. It does not include, however, polarization effects and electromagnetic coupling.

Rigorous electromagnetic diffraction theory is needed for an accurate description of imaging at higher NA than 0.7. Vector Diffraction Theory includes electromagnetic coupling between field components and takes into account the polarization direction of the electric field, which is not necessarily parallel to the polarization of the source. Vector Diffraction Theory is the subject of chapter 2 and Scalar Diffraction Theory is introduced in appendix A.

1.4.4 Resist Modeling

The term "Latent Image" refers to the field amplitude profile inside the photoresist that results from the exposure to an aerial image impinging on the top surface. Common calculations of the fields inside the resist follow a thin film formulation where each plane wave exiting the lens is used as the input to a matrix routine [33, 34, 35, 36, 37].

Evaluation of the formation of the latent image intensity in the photoresist relies also in several models raging from full vector to scalar diffraction theory, where several approximations are applied to speed up the calculations [38]. Just as the full vector models of the imaging process, vector models of the photoresist latent image formation incorporate oblique direction of propagation, polarization effects and electromagnetic coupling between electric field components. Mack et. al. showed [38] that for numerical apertures up to 0.7, scalar models, which ignore the electromagnetic polarization and coupling effects, yielded results that were comparable to those of full vector theory.

Nevertheless, most of these models ignore the photochemical reaction occurring in the resist during exposure, also known as bleaching [33, 36, 38, 39]. As a consequence of exposure, the resist absorption coefficient decreases with time modifying its refractive index and, in effect, the field distribution. Yeung [34] included the effect of bleaching by calculating the refractive index at every time step until the total exposure dose had been delivered. In his method he employed Dill's model of positive photoresist behavior under exposure [40]. Incorporating Dill's model of the bleaching process into full 3D electromagnetic simulation of the exposure can accurately predict the formation of the latent image profile within the resit, however the resultant computational burden is very high.

1.5 In This Thesis

The principal scope of this thesis is the analysis of the errors on aerial image simulations due to electromagnetic diffraction on the reticle topography, and their accurate and computationally efficient modeling. Chapter 3 introduces the Physical Theory of Diffraction foundation of the proposed modeling methodology as well as a historical perspective of its origins. Our model consists of a slight sophistication of the conventional, and rather inaccurate, Thin Mask model, commonly used in lithographic simulations, but with the capability of incorporating topographical effects and polarization dependencies of the field transmitted by the photomask. It is the focus of chapter 4 of this thesis to validate this advanced modeling philosophy for coherent as well and partial coherent illumination, and isolated as well as dense mask features.

Chapter 2 provides a detailed and rigorous description of the aerial image formation for coherent normal illumination, necessary for optical systems of numerical apertures larger than 0.7. Polarization effects and coupling between electromagnetic components becomes noticeable a high numerical apertures and can only be rigorously evaluated through the utilization of vector diffraction theory.

Finally, three appendices are included in this thesis that constitute a good complement to some of the theory presented here. Appendix A covers the scalar counterpart of the vector theory of diffraction. Rigorous electromagnetic derivation of Stratton-Chu and Franz formulas, described in chapter 2, is covered by appendix B. And as a complement to chapter 3, the application of the Physical Theory of Diffraction to the imaging of a rectangular aperture is detailed in appendix C.

CHAPTER 2

Vector Formulation of the Imaging System

Optical projection printing is the technique employed for image patterning on the wafer surface in current optical lithography. Focusing of electromagnetic waves with a lens has long been subject of study in relation to microscope operation. Hence a powerful mathematical and physical framework exists to describe the different phenomena involved in optical lithography. This imaging theory was specially developed in its scalar form [8, 41, 42], known as *Scalar Diffraction Theory*, which is described in more detail in Appendix A.

Scalar diffraction theory [8] yields accurate results of the field in the image space when numerical apertures up to 0.7 are employed [7, 31, 32, 39] (given that the paraxial approximation is not applied), but it fails to account for the polarization and oblique direction of propagation of the vector components, as well as coupling between the various electromagnetic components of the em field, with higher numerical aperture (NA). Rigorous vector diffraction theory was first applied to optical imaging and exposure process for optical lithography by Yeung [34] based in the work by E. Wolf [43], and subsequently analyzed by several other authors in more recent articles [33, 36, 39].

2.1 Wolf's Formulation of Debye's Integral

The phase transformation induced in the propagating wave by a lens composed of spherical surfaces has the property of mapping incident waves into spherical waves, which converge towards the focal point if the lens is a converging one. Until 1909, most theoretical treatments of wave focusing were based on Huygens-Fresnel principle which utilizes spherical-wave representations of the fields. In 1909 Debye reformulated the scalar focusing problem using plane waves rather than spherical waves. Scalar theory of the imaging process is described in appendix A. The scalar focusing formulation was extended to electromagnetic fields by Wolf in his vector generalization of Debye's representation [43].

In the determination of the vector representation of the electromagnetic fields in the image space, Wolf's generalization of Debye's integral formulation [43] was applied which, based on the notation of figure 2.1, takes the form:

$$\mathbf{E}_{image}(x', y', z') = \frac{j}{\lambda} \iint\limits_{{s_x}^2+{s_y}^2 \leq NA^2} \frac{\mathbf{a}(s_x, s_y)}{s_z}\, e^{-jk\left[C+\Phi(s_x,s_y)+\hat{\mathbf{s}}\cdot\mathbf{r}'\right]}\, ds_x ds_y, \tag{2.1}$$

where the temporal term $e^{j\omega t}$ of the monochromatic and coherent wave has been dropped. The phase term e^{-jkC} is constant representing the phase accumulated while propagating through the lens and can also be dropped. The phase term $\Phi(s_x, s_y)$ denotes the aberration function with respect to the ideal spherical wavefront converging towards the focal point.

As indicated in appendix A, in Debye's original derivation, Kirchhoff Boundary Conditions were applied on the exit pupil surface, and each point of the exit pupil was assumed to lay on a spherical wavefront converging towards the focal point. Provided that the linear dimensions of the exit pupil are large compared to λ, hence neglecting the effects due to the edges, the field emerging from the

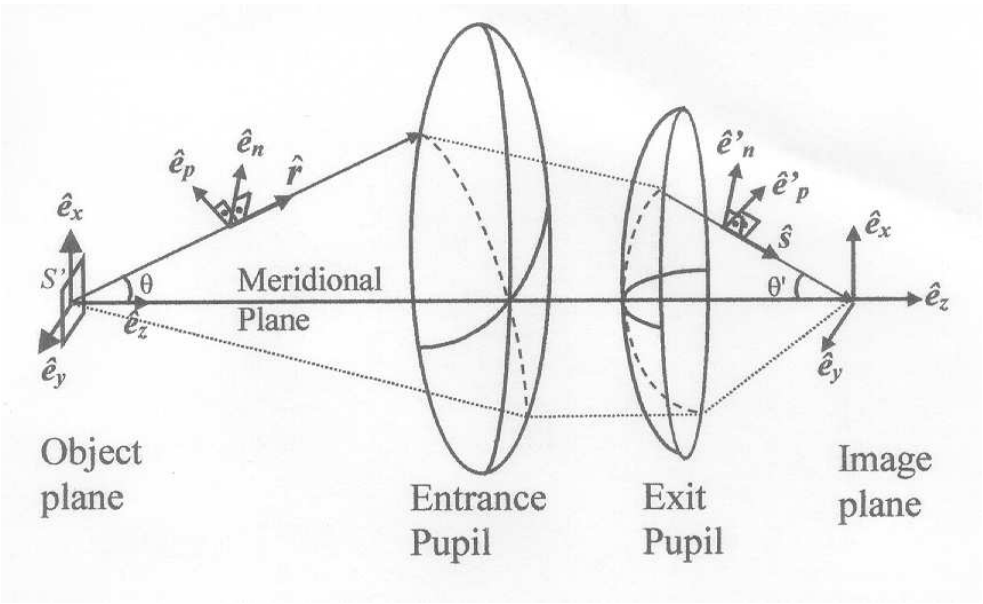


Figure 2.1: Schematic optical projection system setup. Object space coordinates are denoted (x, y, z), while image space coordinates are denoted (x', y', z'). Each direction $\hat{\mathbf{r}}$ of propagation of the waves diffracted by the object form, together with the optical axis, $\hat{\mathbf{e}}_z$, the meridional plane. The corresponding propagation direction $\hat{\mathbf{s}}$ of the wave converging towards the image plane from the exit pupil, lies on the same meridional plane. Field components along the normal and parallel directions to this plane, $\mathbf{E}_n$ and $\mathbf{E}_p$, maintain the same amplitude as the wave vector $k_o\hat{\mathbf{r}}$ is rotated into $k_o\hat{\mathbf{s}}$, that is, $\mathbf{E'}_n$ and $\mathbf{E'}_p$.

exit pupil can be expressed as:

$$\mathbf{E}_{ExitPupil}(\mathbf{r}') = \mathbf{a}(s_x, s_y)\frac{e^{-jk\left[C+\Phi(s_x,s_y)-R'\right]}}{R'} \tag{2.2}$$

Due to the application of Kirchhoff boundary conditions in the derivation of equation (2.1), the approximation will be accurate at distances from the pupil plane that satisfy certain condition. Conventionally, the condition that the Fresnel Number of the focusing geometry as given by $N = a^2/\lambda f$, where a represents

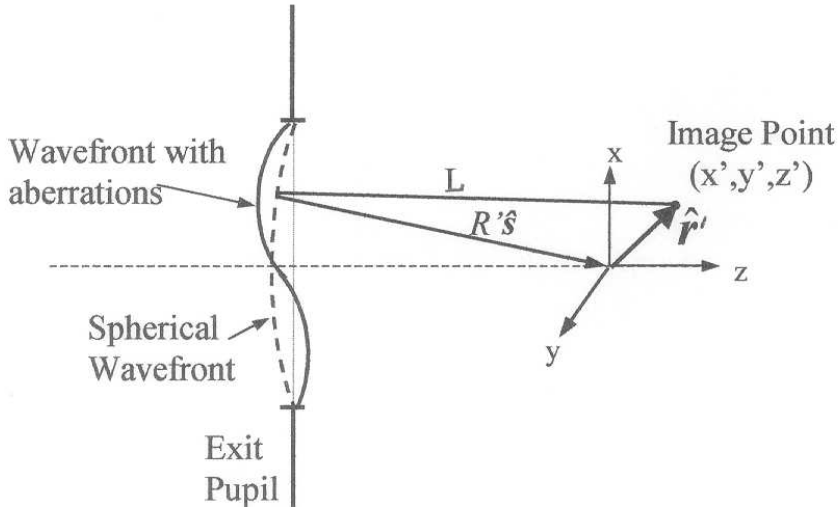


Figure 2.2: Notation for Debye's integral formulation. Converging spherical wave to the gaussian image point on the optical axis from a circular aperture. Also shown is the effect of aberrations.

the aperture radius and f the geometrical focal distance, is much greater than unity or, equivalently, $NA >> \sqrt{\lambda/f}$ [44], is a measure of the validity of Debye's representation. It cannot be applied, however, when the angular aperture is larger than 45^o as for high NA imaging systems. In that case, the general form of the validity condition, as stated by Wolf & Li [44], must be employed:

$$kf >> \frac{\pi}{\sin^2(\theta'/2)} \tag{2.3}$$

With focusing distances of the order of a few mm to $500\mu m$ [7], and numerical apertures as high as 0.85 for 193nm wavelength lithography, this condition guaranties that the Debye integral representation yields essentially the same results for the fields in the focal region as techniques based on Huygens-Fresnel principle [41, 44], Kirchhoff formulation [45] or plane-wave decompositions [36], the difference being in a constant phase factor.

2.1.1 Electromagnetic Fields at the Entrance Pupil

Rigorous electromagnetic theory should be utilized to derive the fields at the entrance pupil. Yeung pioneered the application of rigorous electromagnetic diffraction theory to evaluate the fields diffracted by the photomask in optical lithography [33, 34]. In his extension of the Hopkins theory for partially coherent imaging [33], he evaluated the far field diffraction due to rigorously calculated electromagnetic fields on the mask surface by means of the well-known Stratton-Chu formula [46]. This procedure, which has been followed by other authors in recent publications [39], suffers from a deficiency in its description of the far-fields that can be resolved by using the equivalent formulation due to Franz [47].

According to Stratton-Chu formulation, the diffraction of fields in a volume V limited by a surface S obeys the following equation:

$$\begin{aligned}\mathbf{E}^{SC}(\mathbf{r}) = T \iint_{S'} \Big\{ &- j\omega\mu\big(\hat{\mathbf{n}}' \times \mathbf{H_o}(\mathbf{r}')\big)G(\mathbf{r},\mathbf{r}') + \big(\hat{\mathbf{n}}' \times \mathbf{E_o}(\mathbf{r}')\big) \times \nabla' G(\mathbf{r},\mathbf{r}') \\ &+ \big(\hat{\mathbf{n}}' \cdot \mathbf{E_o}(\mathbf{r}')\big)\nabla' G(\mathbf{r},\mathbf{r}')\Big\} ds', \\ \textit{with } T = &\begin{cases} 1 \text{ if } \mathbf{r} \notin S \\ 2 \text{ if } \mathbf{r} \in S \end{cases}\end{aligned} \tag{2.4}$$

where $G(\mathbf{r},\mathbf{r}')$ represents the free-space Green's function as given by (2.5) for an observation point $\mathbf{r}$ due to a source point $\mathbf{r}'$:

$$G(\mathbf{r},\mathbf{r}') = \frac{e^{-jk|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} \tag{2.5}$$

$\mathbf{E_o}$ and $\mathbf{H_o}$ refer to the electric and magnetic fields on the exit surface of the photomask, respectively, and the unit vector $\hat{\mathbf{n}}'$, normal to the integration surface towards the scattering space, is for our calculations coincident with the unit vector $\hat{\mathbf{e}}_z$ as can be observed from figure 2.3. Equation (2.4) is valid only for fields that are continuous on the surface S. In our calculations these fields on

the mask are either rigorous 3D electromagnetic solutions, or stepwise discontinuous surface field distributions resulting from the direct application of Kirchhoff boundary conditions on the mask surface. The discontinuity can be reconciled with Stratton-Chu equation (2.4) by considering a line distribution of charge on the contour C of the discontinuity:

$$\mathbf{E}^{diff}(\mathbf{r}) = \mathbf{E}^{SC}(\mathbf{r}) + \frac{1}{j\omega\varepsilon} \oint_{C'} \left[(\hat{\mathbf{l}} \cdot \mathbf{H_o}) \nabla' G \right] dl' \tag{2.6}$$

where $\hat{\mathbf{l}}$ represents a unit vector tangent to C at each integration point and dl' is a differential length. This line integral takes into account the effect of a line density of electric charge due to the discontinuity in the tangential components of the magnetic fields on C, and only when added to (2.4) does the resultant expression satisfy Maxwell's equations [46]. A similar term due to the discontinuity of the tangential electric field on C should be added to the magnetic field.

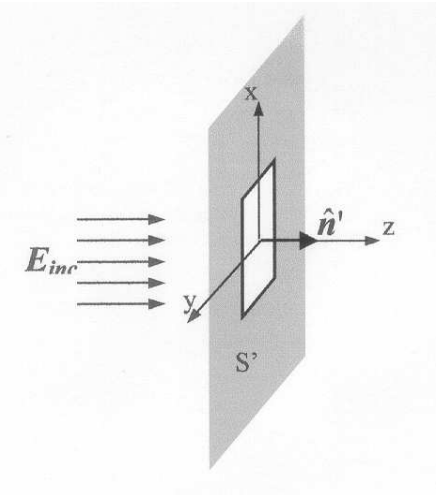


Figure 2.3: Vector notation for Stratton-Chu Integral formulation applied to the fields diffracted by a photomask

On the other hand, the Continuity condition [48] provides the following rela-

tion between the fields:

$$\hat{\mathbf{n}}' \cdot \mathbf{E_o}(\mathbf{r}') = -\frac{1}{j\omega\varepsilon}\nabla' \cdot \big(\hat{\mathbf{n}}' \times \mathbf{H_o}(\mathbf{r}')\big) \tag{2.7}$$

Inserting equation (2.7) into the last term inside the brackets of (2.4) and applying Gauss theorem in 2D, also known as Surface Divergence Theorem [49, 50], it takes form (see appendix A for details):

$$\iint_{S'} \nabla' \cdot (\hat{\mathbf{n}}' \times \mathbf{H_o})\nabla' G \, ds' = -\iint_{S'} \Big((\hat{\mathbf{n}}' \times \mathbf{H_o}) \cdot \nabla'\Big)\nabla' G \, ds' - \oint_{C'} \Big[(\hat{\mathbf{n}}' \times \mathbf{H_o})\nabla' G\Big] \cdot \hat{\mathbf{n}}_\mathbf{s} dl' \tag{2.8}$$

where the contour C bounds the integration surface S where the field is non-zero, and the unit vector $\hat{\mathbf{n}}_\mathbf{s} = \hat{\mathbf{l}} \times \hat{\mathbf{n}}'$ is the contour normal unit vector pointing outwards on S as indicated by figure 2.4.

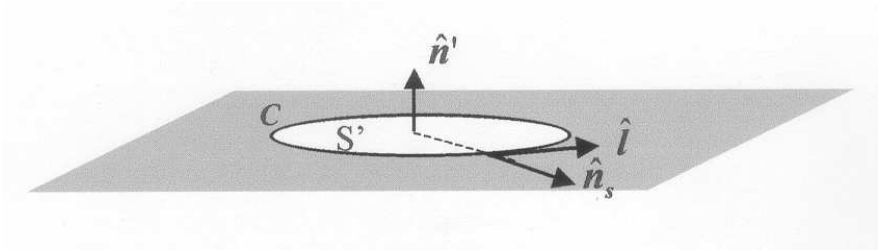


Figure 2.4: Vector notation for the transformation between Stratton-Chu and Franz formula.

Plugging equation (2.8) into (2.4) and making use of the vector identity $(\hat{\mathbf{n}}' \times \mathbf{H_o}) \cdot (\hat{\mathbf{l}} \times \hat{\mathbf{n}}') = (\hat{\mathbf{l}} \cdot \hat{\mathbf{n}}')(\hat{\mathbf{n}}' \cdot \mathbf{H_o}) - (\hat{\mathbf{l}} \cdot \mathbf{H_o})(\hat{\mathbf{n}}' \cdot \hat{\mathbf{n}}') = -(\hat{\mathbf{l}} \cdot \mathbf{H_o})$, equation (2.4) then transforms into:

$$\begin{aligned}\mathbf{E}^{SC}(\mathbf{r}) = \iint_{S'} \Big\{ &-j\omega\mu\big(\hat{\mathbf{n}}' \times \mathbf{H_o}(\mathbf{r}')\big)G(\mathbf{r},\mathbf{r}') + \big(\hat{\mathbf{n}}' \times \mathbf{E_o}(\mathbf{r}')\big) \times \nabla' G(\mathbf{r},\mathbf{r}') \\ &+ \frac{1}{j\omega\varepsilon}\big((\hat{\mathbf{n}}' \times \mathbf{H_o}(\mathbf{r}')) \cdot \nabla'\big)\nabla' G(\mathbf{r},\mathbf{r}')\Big\} ds' - \frac{1}{j\omega\varepsilon}\oint_{C'} \Big[(\hat{\mathbf{l}} \cdot \mathbf{H_o})\nabla' G\Big] dl'\end{aligned} \tag{2.9}$$

The surface integral in equation (2.9) can be identified as Franz's formula for the scattering of fields [47, 51], while the line integral is exactly the contribution due to the field discontinuity about the contour C, which had to be added to

Stratton-Chu formula in equation (2.6) in order to account for the line distribution of sources at the discontinuity. The relation between these two formulations can be expressed as:

$$\mathbf{E}^F(\mathbf{r}) = \mathbf{E}^{SC}(\mathbf{r}) + \frac{1}{j\omega\varepsilon} \oint_{C'} \Big[(\hat{\mathbf{l}} \cdot \mathbf{H_o}) \nabla' G \Big] dl' \tag{2.10}$$

with

$$\mathbf{E}^F(\mathbf{r}) = \iint_{S'} \Big\{ -j\omega\mu \big(\hat{\mathbf{n}}' \times \mathbf{H_o}(\mathbf{r}')\big) G(\mathbf{r}, \mathbf{r}') + \frac{1}{j\omega\varepsilon} \big((\hat{\mathbf{n}}' \times \mathbf{H_o}(\mathbf{r}')) \cdot \nabla'\big) \nabla' G(\mathbf{r}, \mathbf{r}') + \big(\hat{\mathbf{n}}' \times \mathbf{E_o}(\mathbf{r}')\big) \times \nabla' G(\mathbf{r}, \mathbf{r}') \Big\} ds' \tag{2.11}$$

Both formulations can derived from the Vector-Dyadic Green's theorem [50] or from the Vector Green's Theorem [46]. Stratton-Chu's formula is obtained when the Vector Green's theorem is applied to the electric field vector with the aid of the free-space scalar Green's function, while the magnetic field and a modified free-space Green function are employed to derive Franz's formulas (see appendix B for details). Comparing both Franz's and Stratton-Chu's formulas, it is observed that the line integrals added by Stratton and Chu are contained inherently in Franz's formulas. A superior formulation of the vectorial Huygens principle is therefore due to Franz, since it does satisfy Maxwell equations for both continuous and discontinuous electromagnetic fields. A detailed comparison between both formulations was recently published by Tai [47], who proved that when physical optics approximation is applied to an aperture on a metallic surface, then the line integrals must be added to Stratton-Chu's formulas while Franz's formulas cover them automatically. When there is no such discontinuity of the fields on C, then both formulations are equivalent.

The entrance pupil is located at the far-field region of the diffracted field, such

that the Green's function can be approximated by:

$$G(\mathbf{r},\mathbf{r}') = \frac{e^{-jk|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} \approx \frac{e^{-jkr}}{4\pi r} \quad e^{jk\hat{\mathbf{r}}\cdot\mathbf{r}'} \tag{2.12}$$

given that $|\mathbf{r}-\mathbf{r}'|$ can be replaced by its binomial expansion, $r-(\mathbf{r}\cdot\mathbf{r}')/r$, as the observation point $\mathbf{r}$ recedes to the far-field zone. Using:

$$\nabla' G(\mathbf{r},\mathbf{r}') = (jk+\frac{1}{r})G(\mathbf{r},\mathbf{r}')\hat{\mathbf{r}} \approx jkG(\mathbf{r},\mathbf{r}')\hat{\mathbf{r}}, \tag{2.13}$$

and $\mathbf{n}' = \hat{\mathbf{e}}_z$, then:

$$(\hat{\mathbf{e}}_z \times \mathbf{H})\cdot\nabla'\big(\nabla' G(\mathbf{r},\mathbf{r}')\big) \approx -k^2\big[(\hat{\mathbf{e}}_z \times \mathbf{H_o})\cdot\hat{\mathbf{r}}\big]G(\mathbf{r},\mathbf{r}')\hat{\mathbf{r}} \tag{2.14}$$

Substitution of equation (2.13) and (2.14) into (2.11) yields the final integral representation of the fields diffracted by the photomask at the entrance pupil:

$$\begin{aligned}
&\mathbf{E}_{EntrancePupil}(\mathbf{r}) = \\
&= -jk\frac{e^{-jkr}}{4\pi r}\iint_{S'}\big\{\eta(\hat{\mathbf{e}}_z \times \mathbf{H_o}(\mathbf{r}')) - \eta\big[(\hat{\mathbf{e}}_z \times \mathbf{H_o}(\mathbf{r}'))\cdot\hat{\mathbf{r}}\big]\hat{\mathbf{r}} - (\hat{\mathbf{e}}_z \times \mathbf{E_o}(\mathbf{r}'))\times\hat{\mathbf{r}}\big\}e^{jk\hat{\mathbf{r}}\cdot\mathbf{r}'}ds' = \\
&= \frac{1}{j2\lambda}\frac{e^{-jkr}}{r}\boldsymbol{F}\big\{[\eta(\hat{\mathbf{e}}_z \times \mathbf{H_o}(\mathbf{r}')) - \eta\big[(\hat{\mathbf{e}}_z \times \mathbf{H_o}(\mathbf{r}'))\cdot\hat{\mathbf{r}}\big]\hat{\mathbf{r}} - (\hat{\mathbf{e}}_z \times \mathbf{E_o}(\mathbf{r}'))\times\hat{\mathbf{r}}];\frac{r_x}{\lambda},\frac{r_y}{\lambda}\big\}
\end{aligned} \tag{2.15}$$

where $\eta = \sqrt{\frac{\mu}{\varepsilon}}$ represents the intrinsic impedance of the propagation medium and $\boldsymbol{F}$ denotes de Fourier transform evaluated at the spatial frequencies $(\frac{r_x}{\lambda},\frac{r_x}{\lambda})$. The unit vector $\hat{\mathbf{r}}$ indicates the direction from a point on the object plane pointing towards the observation point $\mathbf{r}$ on the entrance pupil.

Only image points of small linear dimensions around the optical axis, $\hat{\mathbf{e}}_z$, as compared to the distance r, are of interest in this analysis, what means that, according to the method of stationary phase [8], only points about the optical axis will contribute significantly to the diffraction integral. Therefore the origin

of the unit vectors $\hat{\mathbf{r}}$ and $\hat{\mathbf{s}}$ can be taken as the center of coordinates on the object plane, and they can be expressed as:

$$\hat{\mathbf{r}} = \frac{\mathbf{r}}{r} = r_x\,\hat{\mathbf{e}}_x + r_y\,\hat{\mathbf{e}}_y + r_z\,\hat{\mathbf{e}}_z = sin\theta cos\varphi\,\hat{\mathbf{e}}_x + sin\theta sin\varphi\,\hat{\mathbf{e}}_y + cos\theta\,\hat{\mathbf{e}}_z \tag{2.16}$$

$$\hat{\mathbf{s}} = s_x\,\hat{\mathbf{e}}_x + s_y\,\hat{\mathbf{e}}_y + s_z\,\hat{\mathbf{e}}_z = -sin\theta' cos\varphi\,\hat{\mathbf{e}}_x - sin\theta' sin\varphi\,\hat{\mathbf{e}}_y + cos\theta'\,\hat{\mathbf{e}}_z \tag{2.17}$$

For each direction $\hat{\mathbf{r}}$, inspection of equation (2.15) shows that the fields on the entrance pupil behave locally as plane waves, satisfying the condition (2.18), and are transversal to the ray direction of propagation. The boundary line charges have the effect of cancelling the longitudinal field components on the far-zone field. Stratton-Chu's formula (2.4), on the other hand, does not predict transverse electromagnetic waves in the far field region of the aperture, what represents another deficiency in favor of Franz's formulation.

$$\mathbf{H}_{EntrancePupil}(\mathbf{r}) = \sqrt{\frac{\varepsilon}{\mu}}\,\hat{\mathbf{r}} \times \mathbf{E}_{EntrancePupil}(\mathbf{r}) \tag{2.18}$$

2.1.2 Electromagnetic Fields at the Exit Pupil

High resolution image formation relies on nearly diffraction-limited imaging characteristics of refractive lenses at the illumination wavelength, what imposes tight design specifications to lens manufactures and allows us to model the imaging system without considering the propagation details through it. As a consequence, the fraction of the cost represented by the lens relative to the total stepper continues to increase due to the high quality fused silica glass and stringent surface finish requirements. Provided that most of the imaging lenses have surfaces that are spherical, the projection lens is therefore represented by a spherical entrance

pupil surface with center at the object and a spherical exit sphere with center at the image, as sketched in figure 2.1.

Turning our attention to equation (2.1) of the aerial image, Wolf applied Kirchhoff boundary conditions on the plane of the exit pupil such that the field at each point on the pupil aperture is represented by equation (2.2). The amplitude function $\mathbf{a}(s_x, s_y)$ can be related to each diffracted ray at the entrance pupil, equation (2.15), by tracing the geometrical ray propagation and polarization state through the optical system. A detailed knowledge of the optical design is therefore needed. High quality, nearly diffraction-limited imaging lenses will be assumed instead, modelled as lossless, isoplanatic (phase invariant) and nearly free of aberrations. The isoplanarity of the imaging system means that the image of an object point changes only in location but not in form as the source point moves through the object plane [41]. Under these circumstances and according to Fresnel refraction formula [52], each ray traces a path that lies on its meridional plane (plane formed by the ray direction, $\hat{\mathbf{r}}$, and the optical axis, $\hat{\mathbf{e}}_\mathbf{z}$) and its polarization vector will, to a good approximation, maintain a constant angle with the meridional plane along the entire path traced by the ray if the angles of incidence at the various surfaces in the system are small [34, 53].

For each source point, the fields at the entrance pupil in the object space are linearly polarized on the plane normal to the propagation direction of each diffracted ray. Each of these diffracted rays is incident on the entrance pupil at nearly normal incidence. Thus the electric field vector forms a very small angle with the glass surface of the first lens. The polarization direction of the fields obtained after refraction at the first lens surface will, according to Fresnel refraction, be also linearly polarized and its polarization direction will be effectively unchanged. Repeating this argument for each lens surface encountered throughout the imaging system, the polarization direction of the electric field vector

maintains an approximately constant angle with the meridional plane, while the ray direction $\hat{\mathbf{r}}$, equation (A.11), will be rotated onto $\hat{\mathbf{s}}$, equation (A.14), as it propagates. Assuming negligible losses due to reflection and absorption, the magnitude of the field vectors at the image region must satisfy the conservation of energy:

$$|\mathbf{E}_{ExitPupil}(s_x, s_y)|^2 \; da' = |\mathbf{E}_{EntrancePupil}(r_x, r_y)|^2 \; da \tag{2.19}$$

where, based on the notation of figures 2.1 and 2.2, $da = r^2 sin\theta d\theta d\varphi$ and $da' = R'^2 sin\theta' d\theta' d\varphi'$.

Further, the angle θ between the incident ray and the entrance pupil, and the angle θ' between the outgoing ray and the exit pupil, must satisfy the Abbe's sine condition [8], $sin(\theta) = Msin(\theta')$ where M denotes the demagnification of the lens (usually or $\frac{1}{4}$ or $\frac{1}{5}$). Assuming an index of refraction equal to unity on both object and image spaces and noting that $d\varphi = d\varphi'$, these two conditions result in the following obliquity factor for the field magnitude at the exit pupil:

$$R'|\mathbf{E}_{ExitPupil}(s_x, s_y)| = r|\mathbf{E}_{EntrancePupil}(r_x, r_y)| \; M \; \sqrt{\frac{cos\theta'}{cos\theta}} \tag{2.20}$$

The direction of polarization will be, in general, different for different diffraction directions. For each of these diffraction orders the electric field can be decomposed into its projection onto the meridional plane, that is, along the direction $\hat{\mathbf{e}}_\mathbf{p}$ parallel to the meridional plane and normal to the ray direction, and its component along the direction $\hat{\mathbf{e}}_\mathbf{n}$ normal to both the meridional plane and to the ray direction. The field amplitudes along these normal and parallel directions to the meridional plane, E_p and E_n, will remain unchanged by refraction during propagation through the optical system, except for the geometrical factor arising from the conservation of energy (2.19). The ray direction $\hat{\mathbf{r}}$, however, is rotated onto $\hat{\mathbf{s}}$ as it propagates through the system such that the projection of the electric field polarization onto the global cartesian axes $(\hat{\mathbf{e}}_\mathbf{x}, \hat{\mathbf{e}}_\mathbf{y}, \hat{\mathbf{e}}_\mathbf{z})$ will be changed and some field coupling between cartesian components will take place.

This polarization rotation is accounted for by the Polarization Rotation Tensor $\mathbf{T}$ derived in section 2.2. Moreover, the spatial frequencies in the image space are related to those in the object space according to $s_x = -\frac{r_x}{M}$ and $s_y = -\frac{r_y}{M}$, and the final relationship between the field amplitude at the exit and entrance pupils is provided by:

$$\mathbf{a}(s_x, s_y) =$$
$$= \frac{1}{j2\lambda} M \sqrt{\frac{cos\theta'}{cos\theta}} \mathbf{T} \boldsymbol{F} \{ [\eta(\hat{\mathbf{e}}_z \times \mathbf{H_o}(\mathbf{r}')) - \eta\big[(\hat{\mathbf{e}}_z \times \mathbf{H_o}(\mathbf{r}')) \cdot \hat{\mathbf{r}}\big]\hat{\mathbf{r}} - (\hat{\mathbf{e}}_z \times \mathbf{E_o}(\mathbf{r}')) \times \hat{\mathbf{r}}]; \frac{Ms_x}{\lambda}, \frac{Ms_y}{\lambda} \} \tag{2.21}$$

for each of the rays at the exit pupil with direction cosines (s_x, s_y, s_z). Note that only the amplitude is considered in equation (2.21) and all phase factors arising from the propagation through the object space and the imaging system as well as the term due to aberrations, enter Debye's formula (2.1) as $e^{-jk\left[C+\Phi(s_x,s_y)\right]}$.

2.2 Polarization Tensor

The cartesian components of each plane wave at the exit pupil can be determined from those at the entrance pupil. For each ray direction (s_x, s_y, s_z), polarization rotation can be expressed as a tensor $\mathbf{T}$, obtained after decomposing the fields into their local projections along the directions normal and parallel to its meridional plane, E_p and E_n, and by applying the condition that the polarization angle with respect to this plane remains approximately constant throughout the projection system.

The directions $\hat{\mathbf{e}}_\mathbf{n}$ and $\hat{\mathbf{e}}_\mathbf{p}$ are given by:

$$\hat{\mathbf{e}}_n = \hat{\mathbf{r}} \times \hat{\mathbf{e}}_z = \frac{\left[r_y \hat{\mathbf{e}}_x - r_x \hat{\mathbf{e}}_y\right]}{\sqrt{r_x^2 + r_y^2}} \tag{2.22}$$

and

$$\hat{\mathbf{e}}_p = \hat{\mathbf{e}}_n \times \hat{\mathbf{r}} = -\frac{\left[\sqrt{1 - r_x^2 - r_y^2}(r_x \hat{\mathbf{e}}_x + r_y \hat{\mathbf{e}}_y) - (r_x^2 + r_y^2)\hat{\mathbf{e}}_z\right]}{\sqrt{r_x^2 + r_y^2}} \tag{2.23}$$

where

$$\begin{aligned} &\hat{\mathbf{r}} = r_x\hat{\mathbf{e}}_x + r_y\hat{\mathbf{e}}_y + r_z\hat{\mathbf{e}}_z \\ &with\ \ r_z = \sqrt{1 - r_x^2 - r_y^2} \end{aligned} \tag{2.24}$$

The transformation matrix between both coordinate systems is as follows:

$$\begin{pmatrix} \hat{\mathbf{e}}_p \\ \hat{\mathbf{e}}_n \\ \hat{\mathbf{r}} \end{pmatrix} = \begin{pmatrix} -\frac{r_z r_x}{\sqrt{r_x^2+r_y^2}} & -\frac{r_z r_y}{\sqrt{r_x^2+r_y^2}} & \frac{(r_x^2+r_y^2)}{\sqrt{r_x^2+r_y^2}} \\ \frac{r_y}{\sqrt{r_x^2+r_y^2}} & -\frac{r_x}{\sqrt{r_x^2+r_y^2}} & 0 \\ r_x & r_y & r_z \end{pmatrix} \cdot \begin{pmatrix} \hat{\mathbf{e}}_x \\ \hat{\mathbf{e}}_y \\ \hat{\mathbf{e}}_z \end{pmatrix} = \mathbf{T}_o \cdot \begin{pmatrix} \hat{\mathbf{e}}_x \\ \hat{\mathbf{e}}_y \\ \hat{\mathbf{e}}_z \end{pmatrix} \tag{2.25}$$

such that on the entrance pupil:

$$\begin{aligned} \boldsymbol{E}_{EntrancePupil}(\mathbf{r}) &= \begin{pmatrix} E_x & E_y & E_z \end{pmatrix} \cdot \begin{pmatrix} \hat{\mathbf{e}}_x \\ \hat{\mathbf{e}}_y \\ \hat{\mathbf{e}}_z \end{pmatrix} = \begin{pmatrix} E_x & E_y & E_z \end{pmatrix} \cdot \mathbf{T}_o^{-1} \begin{pmatrix} \hat{\mathbf{e}}_p \\ \hat{\mathbf{e}}_n \\ \hat{\mathbf{r}} \end{pmatrix} = \\ &= \begin{pmatrix} E_p & E_n & 0 \end{pmatrix} \cdot \begin{pmatrix} \hat{\mathbf{e}}_p \\ \hat{\mathbf{e}}_n \\ \hat{\mathbf{r}} \end{pmatrix} \end{aligned} \tag{2.26}$$

given that the fields are orthogonal to the ray direction of propagation in the far field region of the reticle as observed in equation (2.15).

Repeating this procedure on the image side, we have:

$$\hat{\mathbf{e}}'_n = \hat{\mathbf{e}}_n = \frac{\left[-s_y\hat{\mathbf{e}}_x + s_x\hat{\mathbf{e}}_y \right]}{\sqrt{s_x^2 + s_y^2}} \tag{2.27}$$

and

$$\hat{\mathbf{e}}'_p = \hat{\mathbf{e}}'_n \times \hat{\mathbf{s}} = \frac{\left[\sqrt{1 - s_x^2 - s_y^2}(s_x\hat{\mathbf{e}}_x + s_y\hat{\mathbf{e}}_y) - (s_x^2 + s_y^2)\hat{\mathbf{e}}_z \right]}{\sqrt{s_x^2 + s_y^2}} \tag{2.28}$$

where

$$\begin{aligned} &\hat{\mathbf{s}} = s_x\hat{\mathbf{e}}_x + s_y\hat{\mathbf{e}}_y + s_z\hat{\mathbf{e}}_z \\ &with \quad s_x = -\tfrac{r_x}{M} \quad s_y = -\tfrac{r_y}{M} \quad and\ \ s_z = \sqrt{1 - s_x^2 - s_y^2} \end{aligned} \tag{2.29}$$

and the following transformation matrix:

$$\begin{pmatrix} \hat{\mathbf{e}}'_p \\ \hat{\mathbf{e}}'_n \\ \hat{\mathbf{s}} \end{pmatrix} = \begin{pmatrix} \frac{s_z s_x}{\sqrt{s_x^2+s_y^2}} & \frac{s_z s_y}{\sqrt{s_x^2+s_y^2}} & -\frac{(s_x^2+s_y^2)}{\sqrt{s_x^2+s_y^2}} \\ -\frac{s_y}{\sqrt{s_x^2+s_y^2}} & \frac{s_x}{\sqrt{s_x^2+s_y^2}} & 0 \\ s_x & s_y & s_z \end{pmatrix} \cdot \begin{pmatrix} \hat{\mathbf{e}}_x \\ \hat{\mathbf{e}}_y \\ \hat{\mathbf{e}}_z \end{pmatrix} = \mathbf{T}_i \cdot \begin{pmatrix} \hat{\mathbf{e}}_x \\ \hat{\mathbf{e}}_y \\ \hat{\mathbf{e}}_z \end{pmatrix} \tag{2.30}$$

According to the meridional plane approximation, the field components along the directions normal, $\hat{\mathbf{e}}_\mathbf{n}$, and parallel, $\hat{\mathbf{e}}_\mathbf{p}$, to the meridional plane remain approximately the same as they propagate through the optical projection system, such that at the exit pupil we can assume: $E'_p = E_p$ and $E'_n = E_n$. Thus on the image side:

$$\boldsymbol{E}'_{ExitPupil}(\mathbf{r}) = \begin{pmatrix} E'_p & E'_n & 0 \end{pmatrix} \cdot \begin{pmatrix} \hat{\mathbf{e}}'_p \\ \hat{\mathbf{e}}'_n \\ \hat{\mathbf{s}} \end{pmatrix} = \begin{pmatrix} E_p & E_n & 0 \end{pmatrix} \cdot \begin{pmatrix} \hat{\mathbf{e}}'_p \\ \hat{\mathbf{e}}'_n \\ \hat{\mathbf{s}} \end{pmatrix} =$$

$$= \begin{pmatrix} E_p & E_n & 0 \end{pmatrix} \cdot \mathbf{T}_i \cdot \begin{pmatrix} \hat{\mathbf{e}}_x \\ \hat{\mathbf{e}}_y \\ \hat{\mathbf{e}}_z \end{pmatrix} = \begin{pmatrix} E_x & E_y & E_z \end{pmatrix} \cdot \mathbf{T}_o^{-1} \cdot \mathbf{T}_i \cdot \begin{pmatrix} \hat{\mathbf{e}}_x \\ \hat{\mathbf{e}}_y \\ \hat{\mathbf{e}}_z \end{pmatrix} \tag{2.31}$$

with

$$\begin{pmatrix} E'_x & E'_y & E'_z \end{pmatrix} = \begin{pmatrix} E_x & E_y & E_z \end{pmatrix} \cdot \mathbf{T}_o^{-1} \cdot \mathbf{T}_i = \begin{pmatrix} E_x & E_y & E_z \end{pmatrix} \cdot \mathbf{T} \tag{2.32}$$

and

$$\mathbf{T}_o^{-1} \cdot \mathbf{T}_i = \mathbf{T} = \begin{pmatrix} T_{xx} & T_{yx} & T_{zx} \\ T_{yx} & T_{yy} & T_{zy} \\ T_{zx} & T_{zy} & T_{zz} \end{pmatrix} \tag{2.33}$$

$$T_{xx} = \frac{s_y^2+s_x^2\left(s_z r_z - M(s_x^2+s_y^2)\right)}{(s_x^2+s_y^2)} \qquad T_{yy} = \frac{s_x^2+s_y^2\left(s_z r_z - M(s_x^2+s_y^2)\right)}{(s_x^2+s_y^2)}$$
$$T_{yx} = \frac{-s_x s_y\left(1 - s_z r_z + M(s_x^2+s_y^2)\right)}{(s_x^2+s_y^2)} \qquad T_{zz} = -(s_x^2+s_y^2)M + r_z s_z$$
$$T_{zx} = -s_x(r_z + M s_z) \qquad T_{zy} = -s_y(r_z + M s_z)$$

These results are consistent with the so-called "polarization state matrix" based in the trigonometrically determined polarization transformation by Mansuripur [54], as long as the scalar approximation is applied in the object space, that is, $\frac{M(s_x^2+s_y^2)}{s_z}$, $\frac{M s_z}{s_z} << r_z \sim 1$.

2.3 Simulations

Simulations have been performed for polarized illumination along the x-axis and a $4X$ reduction ratio. As indicated by equation (2.15), only those field components tangential to the mask surface intervene in the expression for the field diffracted by the photomask in the object space. Assuming the XY plane laying on the mask surface, only E_{o_x}, E_{o_y}, H_{o_x} and H_{o_y} need to be considered. For a horizontally polarized illumination, the electric and magnetic field amplitudes on the exit surface of a square aperture on the chrome layer of side length 1.4λ are plotted in figure 2.5. Field amplitude along the main polarization direction, E_{o_x} and H_{o_y} components, exhibit thick mask effects due to the polarization dependence of the boundary conditions and diffraction effects. The cross-polarized field components on the mask surface, E_{o_y} and H_{ox}, present high peaks on the chrome edges of amplitude only one order of magnitude smaller that the main component. These high amplitude edge diffraction peaks produce high order diffraction orders that propagate with large divergence angles towards the entrance pupil. This effect can be observed in figure 2.6 where the spatial frequency spectra for E_{o_x} and H_{o_y} spreads over a broader area than that of E_{o_y} or H_{o_x}.

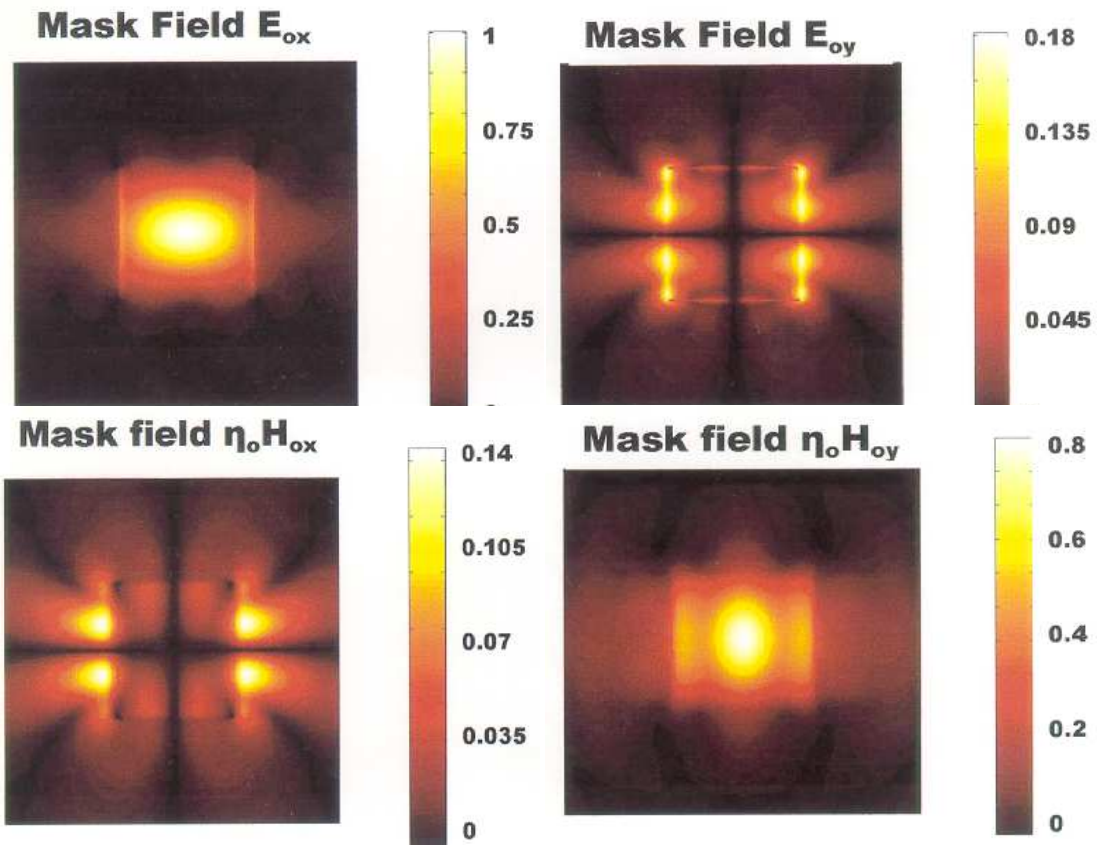


Figure 2.5: Electric and magnetic normalized field amplitudes at the exit surface of a 1.4λ side square opening on a 95nm thick chrome layer with horizontally polarized $\lambda = 193nm$ illumination.

The optical imaging system provides a $4X$ reduction ratio such that, even with high numerical apertures (0.7 to 0.9) of the imaging lens, the angle described by the collected waves on the entrance pupil is small. Under these circumstances, the high order spatial frequencies of the cross-polarized components of the field on the mask fall outside the entrance pupil area and are filtered out by the optical imaging system. Only a small fraction of the energy, several orders of magnitude smaller than the main polarized component, passes through to the image space. Figure 2.7 is a plot of the diffracted field from the 1.4λ square opening on the chrome layer illuminated by $\lambda = 193nm$ x-polarized coherent light

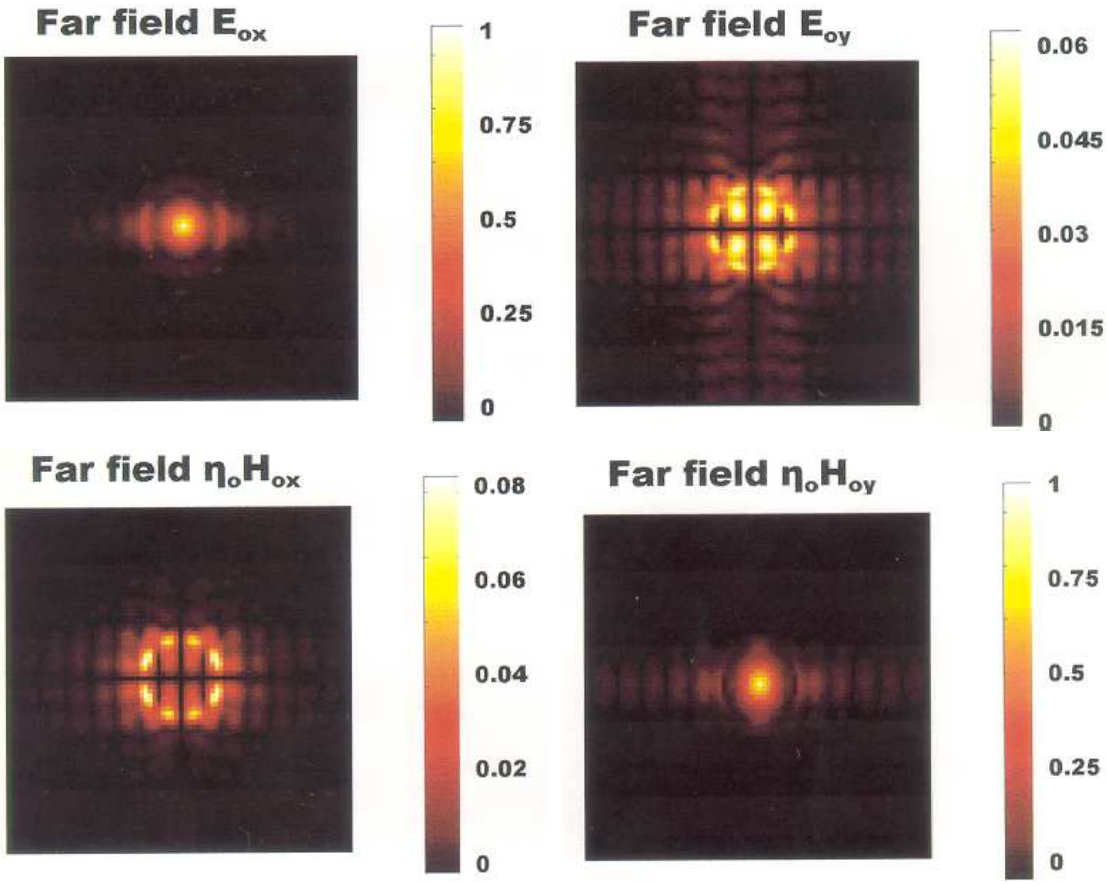


Figure 2.6: Electric and Magnetic normalized field amplitudes at the far field region of a square aperture of 1.4λ side length on a chrome layer of 95nm thickness with $\lambda = 193nm$ horizontally polarized illumination.

that is collected by the entrance pupil of a $NA = 0.85$ imaging system. Hence E_{o_y} and H_{o_x} contribution in the diffracted field expression (2.15) is negligible and can be eliminated from the equation without introducing any noticeable error.

On the basis of the results in figure 2.7, one further approximation can be applied to equation (2.15) that will henceforth be called *Plane Wave Approximation*. The field components along the main polarization direction, E_{o_x} and H_{o_y}, satisfy, with an accuracy of about a 92% based on TEMPEST rigorous

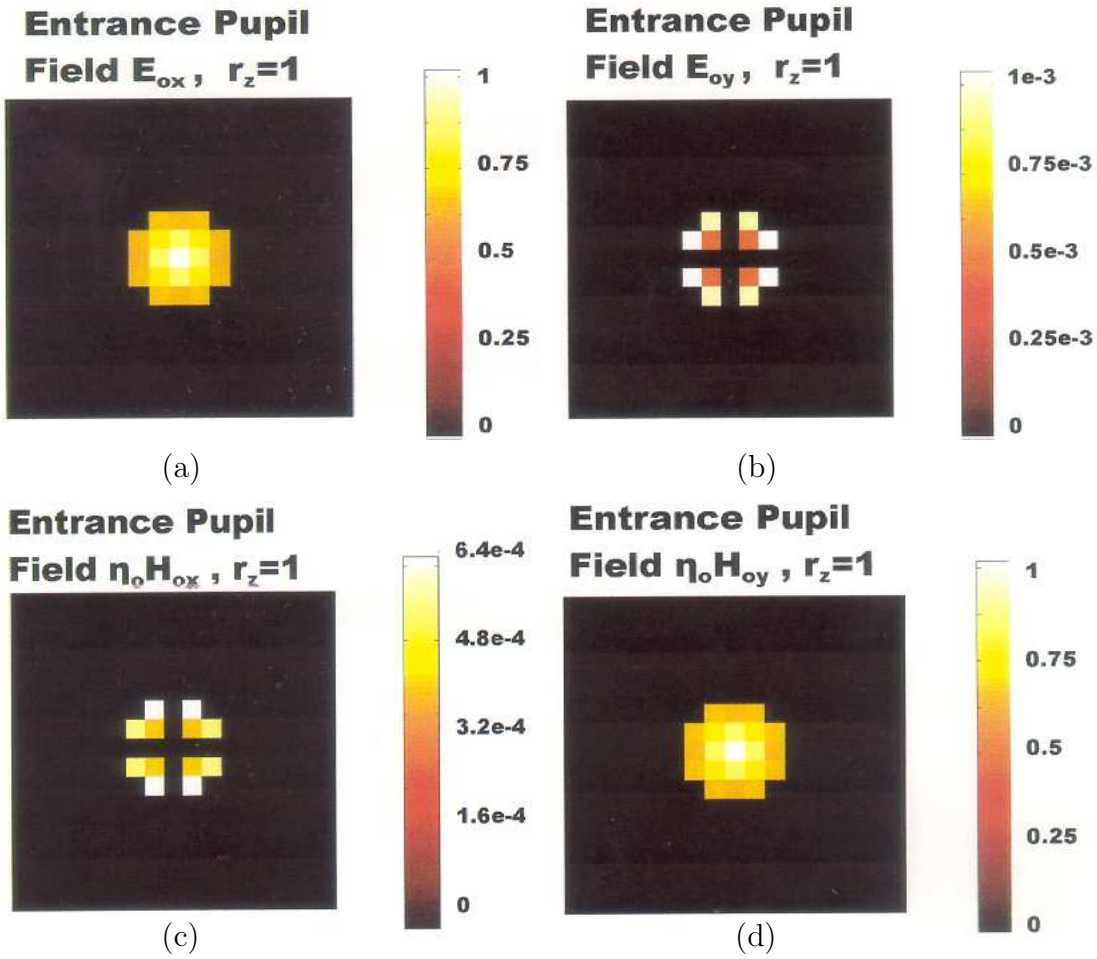


Figure 2.7: Electric and Magnetic diffracted orders collected by the diameter of the imaging system Entrance Pupil with a $NA_o = \frac{0.85}{4}$ on the object space.

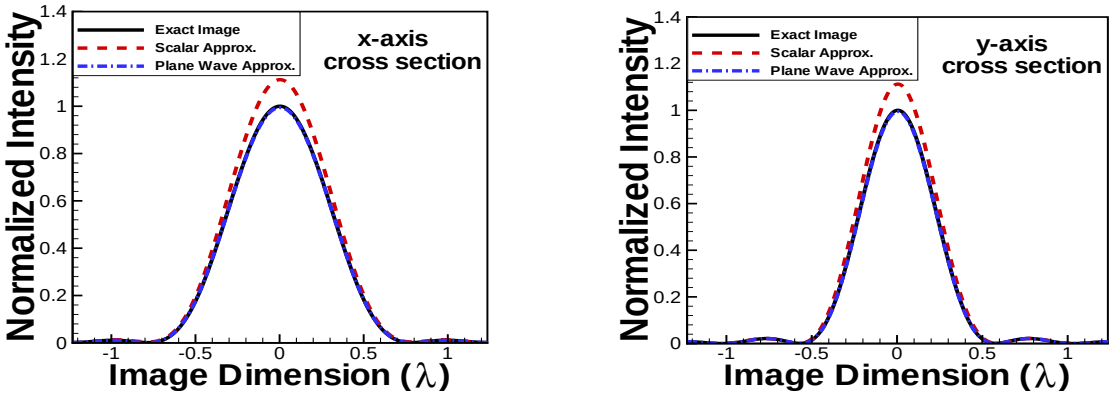
electromagnetic simulations [27], the following condition within the lens NA:

$$\boldsymbol{F}\{\mathbf{H_o}(\mathbf{r})\} \cdot \hat{\mathbf{e}}_\mathbf{y} \simeq \sqrt{\frac{\varepsilon_o}{\mu_o}} \left[\hat{\mathbf{r}} \times \boldsymbol{F}\{\mathbf{E_o}(\mathbf{r})\}\right] \cdot \hat{\mathbf{e}}_\mathbf{y} \tag{2.34}$$

Applying this relation together with the vector identity $\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = (\mathbf{A} \cdot \mathbf{C})\mathbf{B} - (\mathbf{A} \cdot \mathbf{B})\mathbf{C}$ to the integrand of equation (2.15) or (2.21), it can expressed as:

$$\begin{aligned}
&\boldsymbol{F}\Big\{[\eta(\hat{\mathbf{e}}_z \times \mathbf{H_o}(\mathbf{r}')) - \eta\big[(\hat{\mathbf{e}}_z \times \mathbf{H_o}(\mathbf{r}')) \cdot \hat{\mathbf{r}}\big]\hat{\mathbf{r}} - (\hat{\mathbf{e}}_z \times \mathbf{E_o}(\mathbf{r}')) \times \hat{\mathbf{r}}]\Big\} \\
&\approx \boldsymbol{F}\big\{(\hat{\mathbf{e}}_z \cdot \mathbf{E_o})\hat{\mathbf{r}} - (\hat{\mathbf{e}}_z \cdot \hat{\mathbf{r}})\mathbf{E_o} - \Big[\big((\hat{\mathbf{e}}_z \cdot \mathbf{E_o})\hat{\mathbf{r}} - (\hat{\mathbf{e}}_z \cdot \hat{\mathbf{r}})\mathbf{E_o}\big) \cdot \hat{\mathbf{r}}\Big]\hat{\mathbf{r}} + (\hat{\mathbf{r}} \cdot \mathbf{E_o})\hat{\mathbf{e}}_z - (\hat{\mathbf{r}} \cdot \hat{\mathbf{e}}_z)\mathbf{E_o}\Big\} \\
&= \boldsymbol{F}\big\{-2(\hat{\mathbf{e}}_z \cdot \hat{\mathbf{r}})\mathbf{E_o}(\mathbf{r}) + \big[(\hat{\mathbf{e}}_z \cdot \hat{\mathbf{r}})(\mathbf{E_o} \cdot \hat{\mathbf{r}})\big]\hat{\mathbf{r}} + (\hat{\mathbf{r}} \cdot \mathbf{E_o})\hat{\mathbf{e}}_z\big\}
\end{aligned} \tag{2.35}$$

This approximation produces a very small error on the evaluation of the image intensity distribution. This can be observed for the 1.4λ side square opening in figure 2.8(a) and (b), where both the image obtained from the exact field expression (2.21) and from the plane wave approximation (2.35) are being compared, and a RMS error less than 1% is measured.



(a) (b)

Figure 2.8: Cross sectional views along the x- and y-axes of the image intensity distribution on the focal plane of a 1.4λ square chrome opening as generated by the exact field expression (2.21) (solid line), the same expression after neglecting the cross-polarized components and applying the *plane wave approximation* (2.34) (dash-dotted line), and the image generated by the scalar approximation (2.36) (dashed line). The plane wave approximation produces produces less than 1% RMS error, while the scalar approximation produces a RMS error equal to 11.54%. (a) Cross-sectional view of the image along the direction of polarization of the illumination (x-axis). (b) Cross-sectional view of the image along the cross-polarization direction (y-axis).

Expression (2.35) can be further simplified for systems with high reduction ratios ($4X$ or beyond) such that $r_z \approx 1$. For such angles of divergence, the polarization vector remains approximately parallel to the plane of the reticle and, therefore normal to the direction of propagation of the ray. That is,

$\hat{\mathbf{r}} \cdot \boldsymbol{F}\{\mathbf{E_o}(\mathbf{r})\} << \boldsymbol{F}\{\mathbf{E_o}(\mathbf{r})\}$. Under this condition, equation (2.35) reduces to:

$$\boldsymbol{F}\Big\{[\eta(\hat{\mathbf{e}}_z \times \mathbf{H_o}(\mathbf{r}')) - \eta\big[(\hat{\mathbf{e}}_z \times \mathbf{H_o}(\mathbf{r}')) \cdot \hat{\mathbf{r}}\big]\hat{\mathbf{r}} - (\hat{\mathbf{e}}_z \times \mathbf{E_o}(\mathbf{r}')) \times \hat{\mathbf{r}}]\Big\} \approx -2\, cos\theta\, \boldsymbol{F}\{\mathbf{E_o}(\mathbf{r})\} \tag{2.36}$$

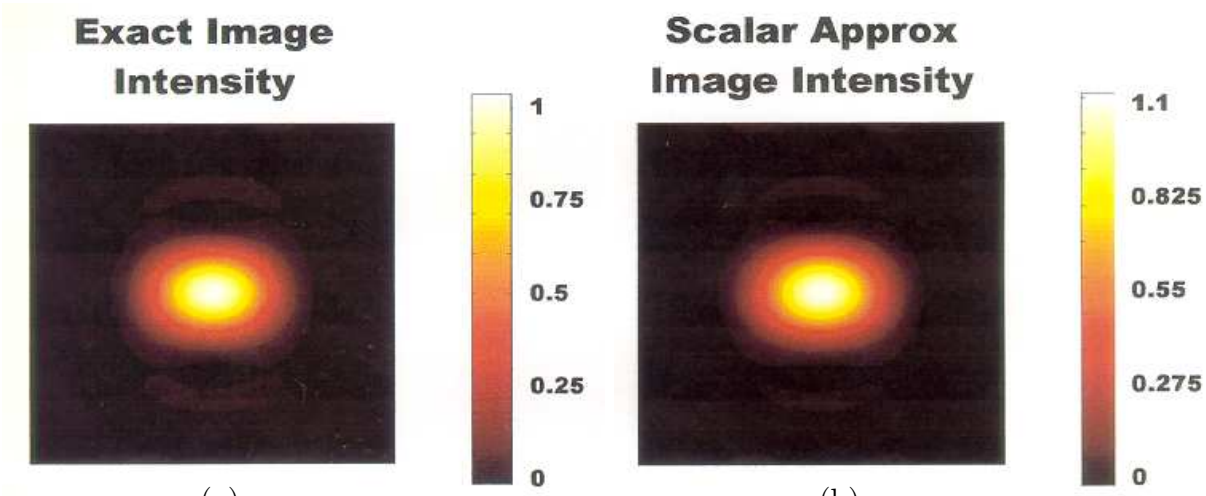
Condition (2.34) was first applied to all field components by Yeung in his extension of Hopkins theory [33], what may not be accurate for cross-polarized components based on the results on figure 2.7(b) and (c). This components are nonetheless ignored in the calculations and, therefore this inaccuracy is irrelevant. When applying the simplified result (2.36) to equation (2.15), it yields the simplest expression for the electric field on the entrance pupil:

$$\boldsymbol{E}_{EntrancePupil}(\mathbf{r}) = \hat{\mathbf{e}}_x \frac{j}{\lambda} \frac{e^{-jk_o r}}{r} \boldsymbol{F}\Big\{E_{o_x}; \frac{r_x}{\lambda}, \frac{r_y}{\lambda}\Big\} cos\theta \tag{2.37}$$

which can now be treated as a scalar quantity consistent with the scalar Huygens-Fresnel formulation [8, 36].

This last approximation, however, produces less accurate aerial images, as can be observed in both figures 2.8(a) and (b) and in figure 2.9 of the image intensity distribution on the focal plane for both the exact diffracted field expression (2.21) and its scalar approximation (2.37). The images correspond to the same 1.4λ square opening on the chrome layer illuminated by an x-polarized plane wave of 193nm wavelength and imaged through a $NA = 0.85$ system.

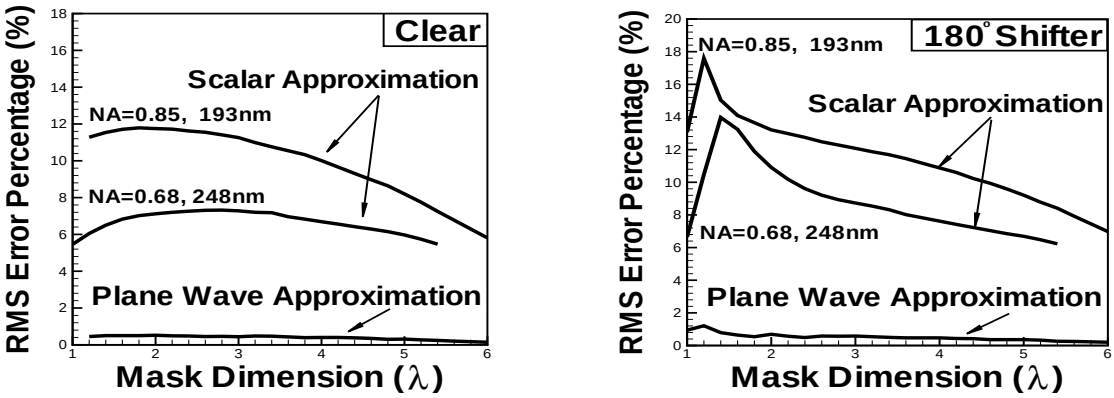
A plot of the total Root Mean Squared Error on the image intensity distribution, after being integrated over the wafer plane, due to the *plane wave* and the *scalar* approximations is shown in figure 2.10. For square openings of sizes ranging from λ to 6λ, the errors have been calculated for two combinations of illumination wavelength and numerical aperture: 248nm wavelength illumination with $NA = 0.68$ imaging and 193nm illumination with $NA = 0.85$ imaging. Ac-



(a) (b)

Figure 2.9: (a) Exact Image intensity distribution on the focal plane produced by a 1.4λ square mask aperture, where all field contributions are taken into account. (b) Image intensity distribution generated by the same aperture field after applying the approximation (2.36) and ignoring cross-polarized components, thus approximating the mask field with a single scalar component.

cording to these results, the error due to the application of equation (2.37) (scalar approximation) is significantly larger than the error due simply to the approximation (2.34) (plane wave approximation), which averages 1% for clear openings and 3% for 180^o-shifter openings, on both $NA = 0.68$ and $NA = 0.85$ imaging systems. A better description of clear and 180^o-shifter type of openings, as defined and used in this thesis, is provided in chapter 3. They refer to chrome apertures introducing either 0^o or 180^o phase shift to the light propagating through them, respectively. As expected, the RMS error due to the application of the scalar approximation increases with increasing numerical aperture as the electric field vector at the entrance pupil becomes less parallel to the plane of the mask and the projection along the optical axis, neglected in the scalar approximation, increases. This error is slightly higher for shifter openings than for clear ones, and



(a) (b)

Figure 2.10: a) RMS error of the image intensity on the focal plane due to the scalar approximation and due to plane wave approximation for clear openings b) RMS error due to the same approximations for 180^o shifter openings imaged through a $4X$ optical system of $NA = 0.85$ at $\lambda = 193nm$.

decreases for larger openings. A peak error value of around 12% is reached for clear openings of about 2λ with a $NA = 0.85$, and of about 7.5% at 2.8λ with $NA = 0.68$. For shifter openings, the peak error happens at about 1.2λ with values 18% and 14% for numerical apertures equal to 0.85 and 0.68, respectively.

This scalar approximation error, however, can be observed in figures 2.8(a) and (b) to be mostly due to a peak amplitude imbalance while the aerial image morphology remains unchanged. Thus, this 15% average RMS error on the image distribution can be tolerated with a simple amplitude scaling, and the fields on the image space can be evaluated with the following simpler expression:

$$\mathbf{E}_{image}(x', y', z') = -\frac{M}{\lambda^2} \iint\limits_{{s_x}^2+{s_y}^2 \leq NA^2} \sqrt{\frac{cos\theta}{cos\theta'}}\, \mathbf{T} \cdot \hat{\mathbf{e}}_x \boldsymbol{F}\Big\{E_{o_x}; \frac{Ms_x}{\lambda}, \frac{Ms_y}{\lambda}\Big\}\; e^{-jk\left[C+\Phi(s_x,s_y)+\hat{\mathbf{s}}\cdot\mathbf{r}'\right]}\, ds_x ds_y \tag{2.38}$$

Finally, while irradiance of the aerial image is evaluated as the magnitude of the time averaged Poynting vector, the absorbed energy within the photoresist and, therefore the final image distribution, is proportional to the squared electric field of the light propagating through it since it is a non-magnetic material [35]. Therefore only the electric field of the aerial image will be evaluated and analyzed in this thesis.

2.4 High Numerical Aperture Effects

The scalar approximation (2.37) of the fields at the entrance pupil introduced in section 2.3 is justified with large reduction ratios of the imaging lens. In systems with $4X$ and $5X$ reduction ratios, the angle described by the diffracted orders collected by the entrance pupil is small even with high numerical apertures (0.7-0.9), and the electric field polarization vector at the entrance pupil can be assumed approximately parallel to the plane of the reticle. As illustrated in figure 2.11, high numerical aperture lenses induce coupling between electromagnetic field components at the exit pupil even if we can assume horizontally polarized illumination at the entrance pupil. As the ray propagates through the lens, the horizontally polarized field component at the entrance pupil will couple onto the field component along the optical axis at the exit pupil.

Numerical apertures of the order of 0.9 are expected to be used in future technology nodes [19], or even beyond that value in the case of immersion lithography. The error calculations of section 2.3 indicated inaccuracies of up to 18% for $NA = 0.85$, mainly due to the approximation $r_z \approx 1$ in the object space. Through this approximation we are ignoring the projection of the electric field vector, initially horizontally polarized on the mask surface ($\hat{\mathbf{e}}_\mathbf{x}$), onto the optical axis direction ($\hat{\mathbf{e}}_\mathbf{z}$) and the cross-polarized direction ($\hat{\mathbf{e}}_\mathbf{y}$), which are becoming

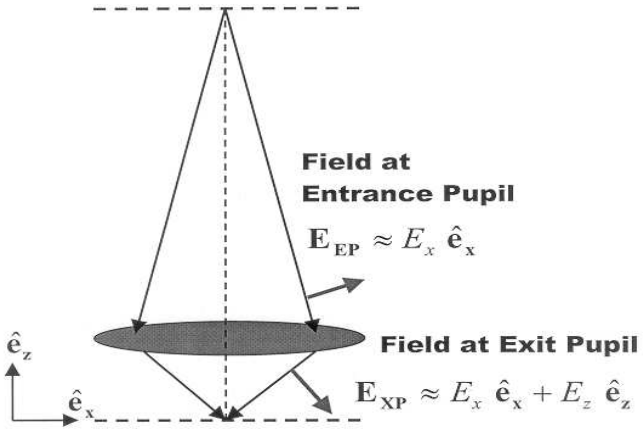


Figure 2.11: High numerical aperture lenses induce coupling between electromagnetic field components at the exit pupil. Assuming horizontally polarized illumination and demagnification such that the fields at the entrance pupil can still be assumed horizontally polarized, there will be coupling onto the field component along the optical axis as the ray propagates through the lenses.

significant for high NA. Denoting $\mathbf{E}^{EP}$ the electric field vector at the entrance pupil as given by equation (2.15), and $\mathbf{E}^{XP}$ the electric field at the exit pupil, according to equations (2.21) and (2.33) we have:

$$\mathbf{T} \cdot \begin{pmatrix} E_x^{EP} \\ E_z^{EP} \\ E_z^{EP} \end{pmatrix} = \begin{pmatrix} E_x^{XP} \\ E_z^{XP} \\ E_z^{XP} \end{pmatrix} = \begin{pmatrix} E_{xx}^{XP} + E_{yx}^{XP} + E_{zx}^{XP} \\ E_{xy}^{XP} + E_{yy}^{XP} + E_{zy}^{XP} \\ E_{xz}^{XP} + E_{yz}^{XP} + E_{zz}^{XP} \end{pmatrix} \tag{2.39}$$

Each of the exit field contributing components, E_{ij}^{XP}, are displayed in figure 2.12 for a 1.4λ square clear opening illuminated at normal incidence with horizontally polarized 193nm light. The field amplitudes on the first row of figure 2.12, that is, figures 2.12(a) to (c), represent the contribution from the main polarization component at the entrance pupil, E_x^{EP}, to E_x^{XP}, E_y^{XP} and E_z^{XP} on the exit pupil, respectively. Similarly, the fields on the second row represent the

contribution from E_y^{EP}, which has a negligible amplitude, and the fields from the third, the contribution from E_z^{EP} with a more significant amplitude.

The coherent sum of the fields plotted in figures 2.12(a), (d) and (g), provides E_x^{XP}, which carries the largest field amplitude as plotted in figure 2.13(a), while the fields on the second column, figures 2.12(b), (e) and (h) add up to form E_y^{XP}, which still has a negligible amplitude, even after some coupling from the polarized component. Figures 2.12(c), (f) and (i) combine to provide E_z^{XP} with a rather noticeable amplitude, plotted in figure 2.13(c), relative to the largest component, plotted in figure 2.13(a).

Two main effects are noticed as a consequence of high-NA imaging. First of all, coupling of the polarized field component E_x^{EP} onto E_z^{XP} gives place to a two-lobe image intensity distribution, figure 2.14(c), that adds incoherently to I_x^{XP}, figure 2.14(a), and is responsible for a broadening of the image along the polarization direction [36]. A second more subtle effect rises up from the coupling of the E_z^{EP} component on the entrance pupil, ignored by the scalar approximation (2.37), onto the polarization direction E_{zx}^{XP}. This field amplitude, although small in amplitude, adds coherently to E_x^{XP} and is responsible for the 15% average error due to the scalar approximation reported in section 2.3 and observed in figures 2.9 and 2.8.

2.5 Discussion

As the numerical aperture of the optical system reaches values equal and larger than 0.7, polarization effects, as well as coupling between the various electromagnetic components of the em field, are becoming increasingly noticeable. Simple scalar diffraction theory can no longer model satisfactorily this High-NA effects

and more rigorous vector diffraction theory is necessary.

In this chapter, a detailed description of the aerial image formation process for coherent normal illumination was provided through the application of rigorous electromagnetic field theory on both the object and image spaces. Under the assumption of nearly aberration-free lens quality, the meridional plane approximation is employed to link the electromagnetic fields at the entrance pupil to those at the exit pupil. Further approximations can be applied to simplify this rigorous theory. The so-called in this thesis *plane wave* approximation and the *scalar* approximation were analyzed, together with an estimate of the error introduced.

Finally, some of these High-NA effects are mitigated by the refractive effect of the photoresist layer, and the resist latent image generally exhibits improved quality relative to the aerial image [7, 39]. The focus of this thesis, however, is the analysis of the effect of the photomask near fields on the aerial image quality, and their modeling such that they are incorporated into the simulation of the imaging process. An accurate modeling of the mask fields will provide more accurate aerial image calculations and, consequently more accurate resist latent images regardless of the influence of the resist parameters. Analysis of the resist effects is above the scope of this thesis.

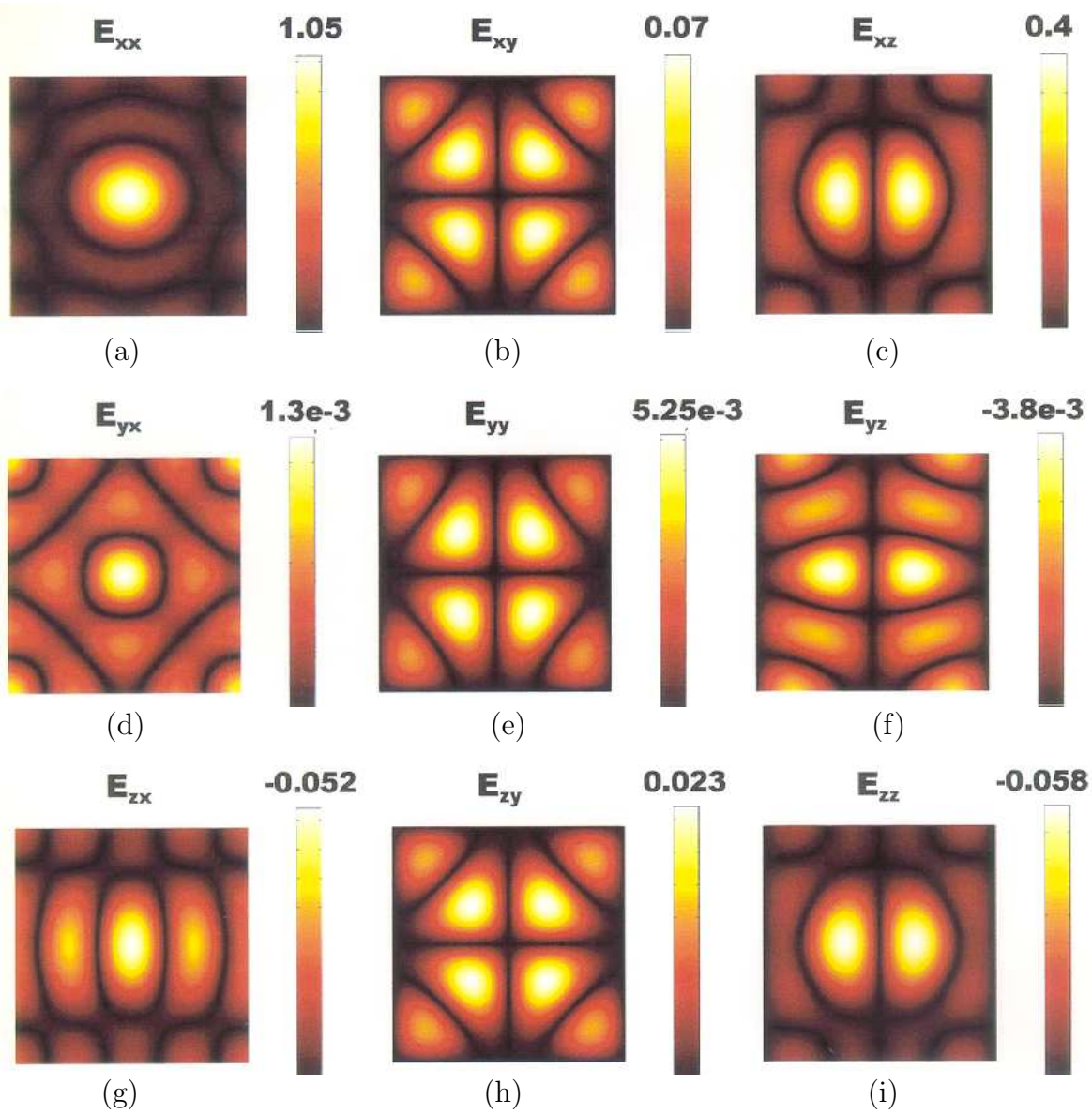
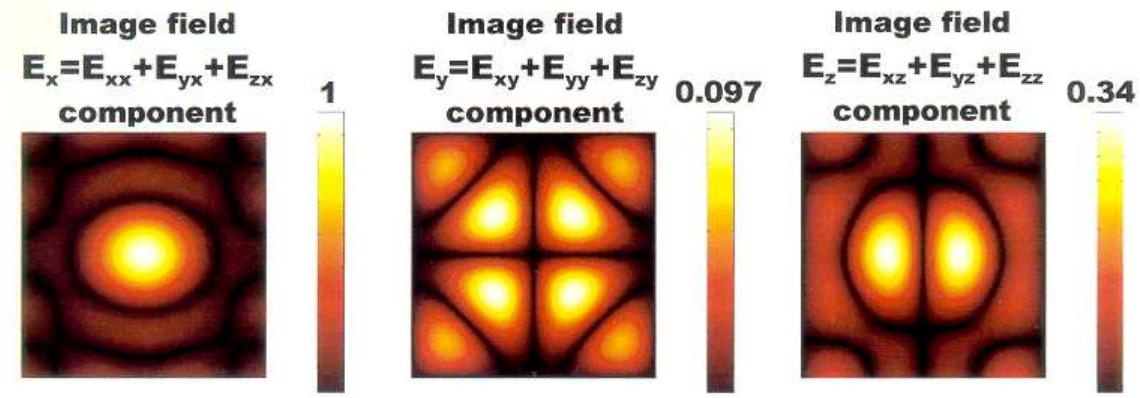
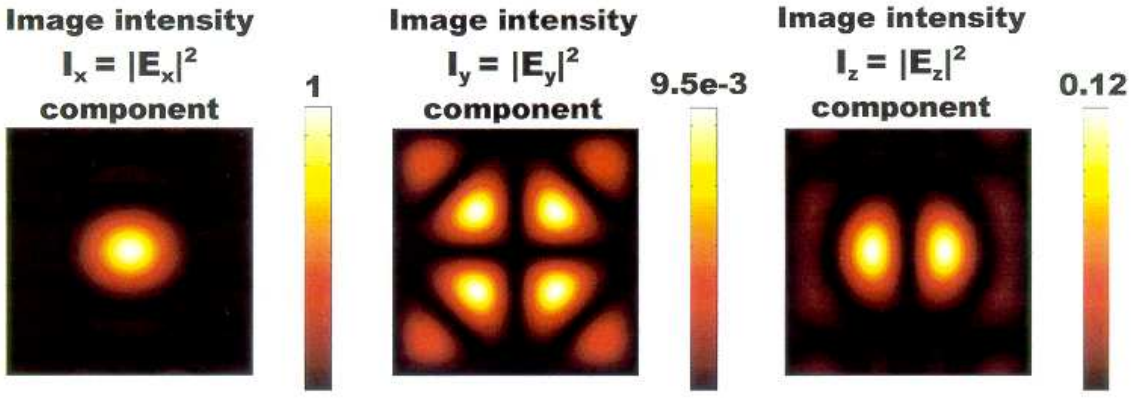


Figure 2.12: Electric field amplitude distribution over the focal plane of all contributions to the electric field of the image formed by a $NA = 0.85$, $4X$ imaging system at $\lambda = 193nm$. (superscripts XP have been omitted for clarity)



(a) (b) (c)

Figure 2.13: Electric field distribution over the focal plane of the three cartesian components of the image of a 1.4λ square opening. (superscripts XP omitted for clarity)



(a) (b) (c)

Figure 2.14: Intensity distribution over the focal plane of the three cartesian components of the image of a 1.4λ square opening.(superscripts XP omitted for clarity)

CHAPTER 3

Thick Mask Effects

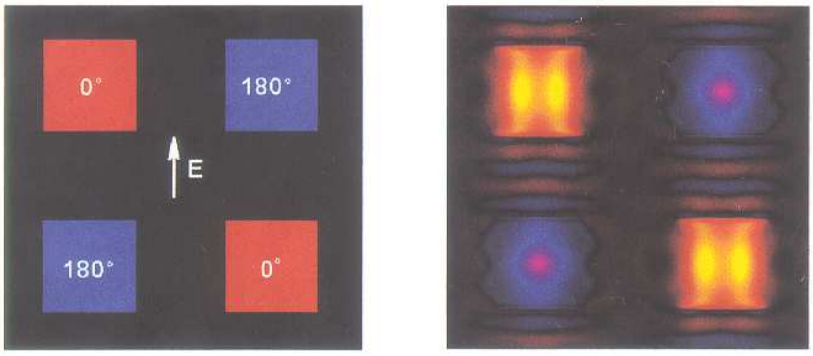
3.1 Kirchhoff Boundary Conditions

In the evaluation of the fields diffracted by the reticle it has been customary to apply the so-called *Kirchhoff's Boundary Conditions* to approximate the boundary value of the fields on the mask surface. This boundary conditions, which are introduced in its scalar form in appendix A, consist of replacing the fields on the aperture by those which would exist in the absence of the screen. Direct application of Kirchhoff's boundary conditions provides a *Thin Mask* approximation of the object field on the exit surface of the mask as illustrated in figure 3.1(a), obtained after multiplying the incident field times an ideal transmission function of the mask pattern. Figure 3.1(a) represents the Thin Mask model of a 2λ half-pitch array of alternating 180^o-shifter and clear square mask openings, with vertically polarized electric field. Figure 3.1(b), on the other hand, illustrates the actual object field obtained by rigorous electromagnetic simulation on the same mask pattern with the Finite-Difference Time-Domain software Tempest 6.0 [20].

This thin mask model ignores diffraction and polarization effects shown by the rigorously evaluated aperture field, which is being displayed in figure 3.1(b). Nevertheless it provides surprisingly accurate results when the diffraction from mask features much larger than the wavelength ($ka >> 1$, with k being the wavevector and a the aperture size) is evaluated at distances several wavelengths

from the aperture ($z > \frac{2a^2}{\lambda}$).

The utilization of 193nm wavelength lithography with a $0.85NA$ to print 65nm wafer features translates into k_1 factors approaching values around 0.3 and mask features of the order of the wavelength for $4X$ magnification. In addition, Alternating Phase-Shifting Masks (Alt. PSM) employ etching profiles with abrupt discontinuities and trench depths also in the order of the wavelength for 180^o phase-shifting openings. As a consequence of wavelength sized and high aspect ratio mask features, mask topography effects are becoming an increasing source of simulation errors [19, 21] in lithography, which are particularly critical for Alternating Phase-Shifting masks [10], and demand rigorous resource-consuming 3D electromagnetic field simulations in the subwavelength regime.



(a) (b)

Figure 3.1: (a) Kirchhoff scalar approximation (thin mask model) of the field on the mask plane of a 2λ half-pitch array of alternating 180^o-shifter and clear square mask openings, with vertically polarized electric field. (b) Sketch of the actual object field obtained by rigorous electromagnetic FDTD Tempest simulation on the same mask pattern.

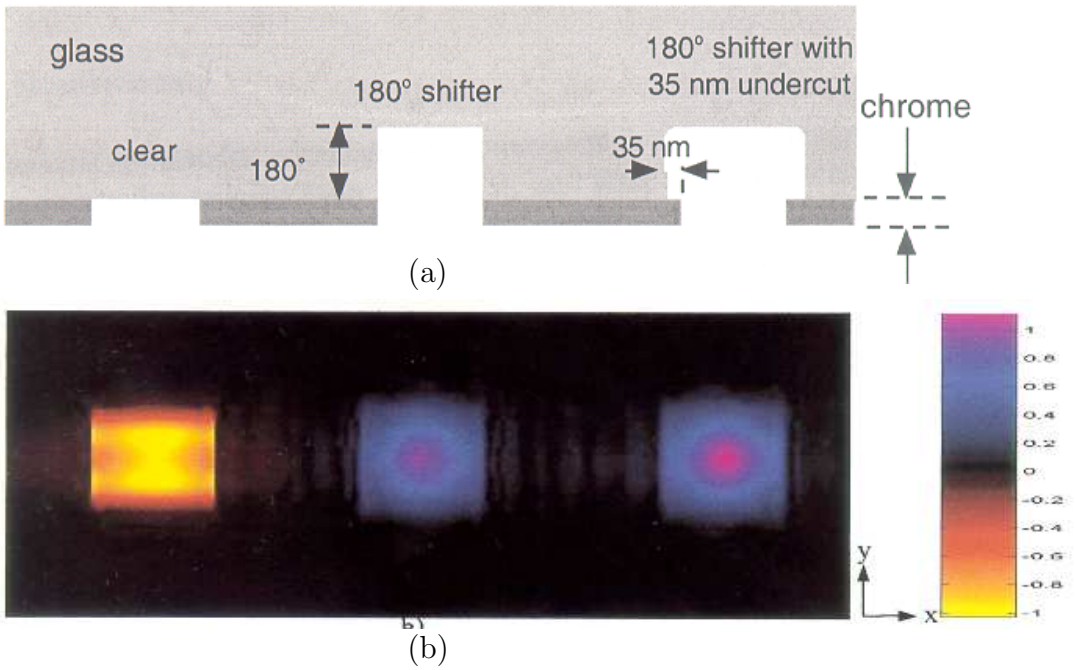
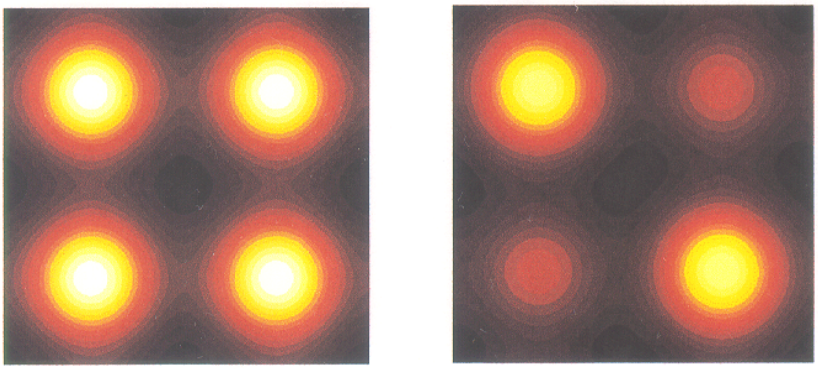


Figure 3.2: (a) Sketch of the three types of openings on a transmission lithography mask analyzed in this thesis. (b) Aperture field generated by each opening.

3.1.1 Thick Mask Effects

Alternating Phase-Shifting masks (Alt. PSM) have mask etching profiles with abrupt discontinuities as sketched in figure 3.2(a) for clear and 180^o phase-shifting openings with and without 35nm undercut. The rigorously evaluated electromagnetic fields on thick chrome mask apertures displayed in figure 3.2(b) exhibit effects that vary for different opening types as a consequence of the vector nature of light. These effects include polarization dependence of the boundary conditions with the edge orientation, diffraction at the edges and, as sketched for a typical case in figure 3.3(b), transmission imbalance between etched and unetched openings [55, 56, 57]. They also introduce deviations on the phase of the electromagnetic wave propagating through the 180^o-phase shifting openings relative to the unetched or clear openings. Phase errors become particularly critical for

Alternating Phase-Shifting Masks since they are based on the modulation of both amplitude and phase of the em fields propagating through them. An example of the amplitude imbalance between clear and shifter openings is displayed in figure 3.3(b), corresponding to the aerial image intensity at the focal plane of the rigorously evaluated object field of figure 3.1(b). As previously indicated, the mask features in this example had side dimensions equal to 2λ and, therefore their Kirchhoff's approximation introduces high errors, as high as 85% root mean squared error relative to the actual image in this particular example.



(a) (b)

Figure 3.3: (a) Aerial image field generated by the Thin Mask approximation of the field on the mask plane of a 2λ half-pitch array of alternating 180^o-shifter and clear square mask openings. (b) Sketch of the actual image field generated by the rigorously evaluated fields on the same mask. For both images, unpolarized, $\sigma = 0.6$ partially coherent illumination at 193nm was used with a $NA = 0.85$ imaging system.

Rigorous 3D electromagnetic field simulations of the fields on the mask surface are necessary but extremely resource and time-consuming, hence impractical even for small portions of the mask. More accurate physical models are in high demand, such that retain much of the Thin Mask simplicity, thus allowing fast and reliable simulations.

3.2 The Boundary Diffraction Wave

3.2.1 Historical Antecedents

As mentioned in chapter 2, Huygens was who first proposed a wave theory of light in 1678. His elementary-wave principle couldn't explain, however, how the various elementary waves summed up and, in consequence, couldn't explain the diffraction phenomena. More than one century later, in 1802, Thomas Young proposed a physical model for diffraction as the result of the interference of two waves [58]. Young believed that the incident light undergoes a kind of reflection at the boundary of the diffracting body and he considered diffraction to arise from the interference between the direct light beam and the light propagated from each point on the boundary [59]. This interpretation of diffraction in terms of edge or boundary waves was suggested prior to Fresnel's theory, published in 1818, which in fact combined Huygens' envelope construction with Young's interference principle to describe the effects of diffraction. However, because Young expressed his ideas in a rather qualitative manner, they did not gain general recognition and Huygens-Fresnel's theory of diffraction soon dominated the field [58].

In 1888, a few years after Kirchhoff, in 1882, had provided the mathematical foundations to the *Huygens-Fresnel* Principle, Young's interpretation of diffraction was given mathematical validity by Maggi, one of Kirchhoff's students in Berlin [58]. Maggi's formulation was independently derived by the polish scientist Rubinowicz (one of Sommerfeld's assistant in Munich) in 1917 in a slightly different form which has become more commonly accepted [60]. This so-called Maggi-Rubinowicz representation of Kirchhoff's formula divides Kirchhoff's surface integral, which is defined on the screen aperture, into two contributions: an incident wave propagated through the aperture in accordance with the laws of

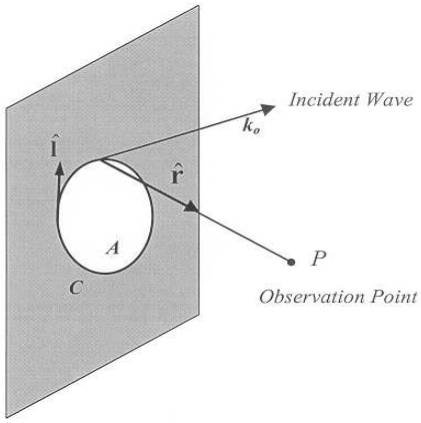


Figure 3.4: Vector notation for Rubinowicz's formulation

geometrical optics (Geometrical Wave), and a diffracted wave emanating from every point of the boundary of the aperture (the Boundary Diffraction Wave). Rubinowicz extracted the edge wave contribution in the Kirchhoff integral for the diffraction of waves by an aperture on an opaque screen as follows:

$$U(\mathbf{r}) = U^{GO}(\mathbf{r}) + U^{diff}(\mathbf{r}) \tag{3.1}$$

The first is the geometrical optics field

$$U^{GO}(\mathbf{r}) = \begin{cases} U^{o}(\mathbf{r}) & \text{if } \mathbf{r} \text{ is in the illuminated region} \\ 0 & \text{if } \mathbf{r} \text{ is in the shadow region} \end{cases} \tag{3.2}$$

and the second is the true diffraction field by the edges

$$U^{diff} = -\oint_C U_o(\mathbf{r}') \frac{[\hat{\mathbf{r}} \times \hat{\mathbf{k}}_\mathbf{o}] \cdot \hat{\mathbf{l}}}{1 - [\hat{\mathbf{r}} \cdot \hat{\mathbf{k}}_\mathbf{o}]} \frac{e^{-jk_o|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} \, dl \tag{3.3}$$

where the integral is taken along the edge C of the perimeter of the aperture A in the direction indicated by figure 3.4. The unit vector $\hat{\mathbf{l}}$ is the tangent, at each point $\mathbf{r}'$, to the contour C, while $\hat{\mathbf{r}}$ is directed from the integration point to the observation point. The unit vector $\hat{\mathbf{k}}_\mathbf{o}$ represents the propagation direction of

the incident wave impinging on the edge. Equation (3.3) can be interpreted as the integral of elementary edge waves originating at each point of the boundary, therefore providing a theoretical proof of the existence of diffracted rays associated with the presence of discontinuities in the obstacle. This representation in terms of single integrals has computational advantage over the double integral diffraction formulation based in Huygens-Fresnel principle, and was subsequently generalized for incident waves that are not plane or spherical by Miyamoto and Wolf [59, 61] in 1962. In the same year, Marchand and Wolf also extracted the edge wave contribution from the Rayleigh-Sommerfeld formulas [62]. All these representations, though, are based on Kirchhoff's high frequency approximation of the field behind the screen and, therefore do not provide the rigorous solution.

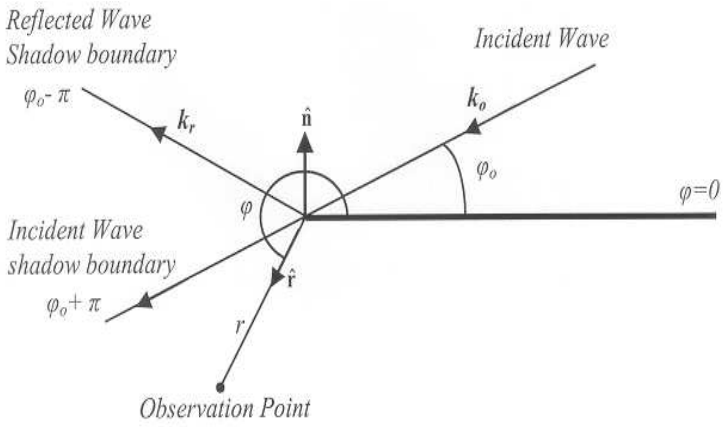


Figure 3.5: Geometry of Sommerfeld's half-plane diffraction problem.

Later, in 1896, A. Sommerfeld obtained a rigorous theoretical solution of the diffraction of a plane electromagnetic wave from a perfectly conducting half-plane. Ever since, several authors have rederived Sommerfeld's solution, for a half-plane as well as for a wedge, through different mathematical methods [63, 64, 65, 66]. Sommerfeld's solution is given in terms of Fresnel integrals [67] and its asymptotic

expansion becomes:

$$U^{tot}(r) = U^{inc}(r)\ \mu((\pi+\varphi_0)-\varphi) \mp U^{ref}(r)\ \mu((\pi-\varphi_0)-\varphi) + D_{s,h}(\varphi,\varphi_o)\frac{e^{-j(kr+\frac{\pi}{4})}}{\sqrt{2\pi kr}} + 0\left((kr)^{-3/2}\right) \tag{3.4}$$

where $\mu(x)$ is the unit step function, i.e., $\mu(x) = 1$ for $x > 0$ and $\mu(x) = 0$ for $x < 0$, and

$$D_{s,h}(\varphi,\varphi_o) = -\frac{1}{2}\Big\{sec\frac{\varphi-\varphi_o}{2} \mp sec\frac{\varphi+\varphi_o}{2}\Big\} \tag{3.5}$$

The upper (lower) sign refers to E- or soft polarization (H- or hard polarization), defined according to:

E-polarization: $U^{inc} = E_z$ such that $U^{total} = 0$ on half plane

H-polarization: $U^{inc} = H_z$ such that $\frac{\partial U^{total}}{\partial n} = 0$ on half plane

The angles φ and φ_o, sketched in figure 3.5 for the half-plane diffraction problem, correspond to the direction of the observation point and incident plane wave, respectively, which are measured from the illuminated metal face. The unit vector $\hat{\mathbf{n}}$ is the normal to the half-plane and the z direction is coincident with the plate edge. The subscript s refers to the case of the boundary condition $U = 0$ (Dirichlet's problem), also known as soft boundary conditions and corresponding to the E-polarization in Sommerfeld's problem, and the subscript h refers to the case of the boundary condition $\frac{\partial U}{\partial n} = 0$ (Neumann's problem), known as hard boundary conditions and corresponding to the H-polarization of Sommerfeld's problem.

In equation (3.4), the first and second terms describe the incident and reflected fields according to geometrical optics. The third term is the leading term of the high-frequency asymptotic expansion of the diffracted field, which has the form of a cylindrical wave emanating from the edge of the half-plane and propagating through both lit and shadow regions, hence rigorously confirming Young's ideas.

In 1912, experimental proof of the existence of these edge diffracted rays was provided by the Russian scientist Kalashnikov [68], who first recorded them on a photographic plate.

The next important step in the development of the diffracted ray concept was taken by Rubinowicz, who performed the first ray asymptotic expansion of the diffraction integral (3.3) at high frequencies by the method of stationary phase [69]. Rubinowicz showed that every stationary point on the edge created an entire cone of diffracted rays which satisfied Fermat's principle [70], $\hat{\mathbf{r}}_\mathbf{s} \cdot \mathbf{l}_\mathbf{s} = \hat{\mathbf{k}}_\mathbf{os} \cdot \mathbf{l}_\mathbf{s}$. In particular, for incident plane waves it is expressed as the following sum of the contribution arising from each stationary point l_s on the contour C [71]:

$$U^{diff}(r_s) \simeq -\frac{1}{2} \sum_s U_o(l_s) \, \frac{[\hat{\mathbf{r}}_\mathbf{s} \times \hat{\mathbf{k}}_\mathbf{os}] \cdot \hat{\mathbf{l}}_\mathbf{s}}{1 - [\hat{\mathbf{r}}_\mathbf{s} \cdot \hat{\mathbf{k}}_\mathbf{os}]} \, \frac{e^{i(kr_s + \frac{\pi}{4})}}{\sqrt{2\pi k r_s}} \tag{3.6}$$

where the vector $\hat{\mathbf{r}}_\mathbf{s}$ represents the unit vector pointing towards the observation point from the stationary point l_s, and the vectors $\hat{\mathbf{k}}_\mathbf{os}$ and $\mathbf{l}_\mathbf{s}$ represent the propagation direction of the incident field and the tangent to the contour at the same point l_s, respectively. If the boundary has discontinuities (corners), these will give rise to the next term in the asymptotic expansion of U^{diff} [61].

Both Sommerfeld's and Rubinowicz's results established the basis of the modern ray asymptotic theories as they were incorporated by Keller in his Geometrical Theory of Diffraction (GTD), enunciated in 1953 [72]. Geometrical Theory of Diffraction is an extension of Geometrical Optics which accounts for diffraction. Keller's hypothesis shares Young's idea that at optical frequencies diffraction is an edge effect. In other words, at high frequencies diffraction is a local phenomenon, determined only by the local properties of the incident field and the diffracting object. As a result, one can reduce the diffraction problem by complex objects to the superposition of simpler canonical solutions, such as that of Sommerfeld's half-plane. When the incident wave is propagating in a direction normal to the

screen, the diffracted wave is cylindrical originating from the edge. The straight lines perpendicular to the cylindrical wavefronts represent the diffracted rays leaving the edge in all directions. Keller used Fermat's principle to determine that when the incident rays are oblique to the edge, the diffracted wave front is conical as sketched in figure 3.6 extracted from Keller's work [72]. Straight lines orthogonal to these cones (also coming from the edge) represent the diffracted rays, which make equal angle with the edge as the incident ray. These diffracted rays form the so-called Keller Cone [72], also called Rubinowicz-Keller diffraction cone or just diffraction cone.

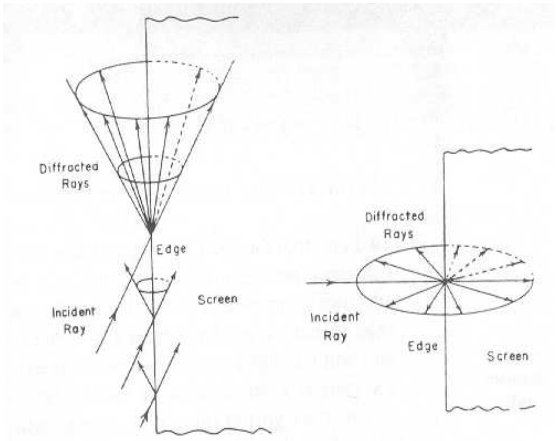


Figure 3.6: When the incident ray hits the edge of a thin screen obliquely, the diffracted cone of rays form a angle with the edge equal to the angle between the incident ray and the same edge. When the incident ray hits the edge normally the rays are diffracted in all directions on the same plane of incidence.

3.2.2 Physical Optics Approximation

On metallic objects of finite size the Physical Optics (PO) approximation assumes that the current induced by the incident field on the object illuminated surface

equals the current that would be induced on an infinite planar metallic plate:

$$\mathbf{J}^{PO} = 2\ \hat{\mathbf{n}} \times \mathbf{H}^{\mathbf{inc}} \tag{3.7}$$

The total scattered field is found through the vector potential $\mathbf{A}$ calculated by integration of the current:

$$\mathbf{A} = \int \mathbf{J}^{PO} \frac{e^{ikr}}{4\pi r} ds \tag{3.8}$$

The Physical Optics approximation is similar to Kirchhoff's boundary conditions in that the field is assumed unperturbed by the boundaries of the scattering object. Kirchhoff's approach, however, attempts to determine the field behind the opaque screen with an aperture, while the Physical Optics approximation refers to the equivalent currents induced on a reflecting surface being determined through the reflection coefficients on infinite surfaces. Some authors defend that Physical Optics approximation should be considered as an extension of Kirchhoff's approach [58]. In fact, by means of Babinet's principle [8], Rubinowicz's expansion of Kirchhoff's integral can be seen to derive from the Physical Optics approximation of the fields on Sommerfeld's half-plane.

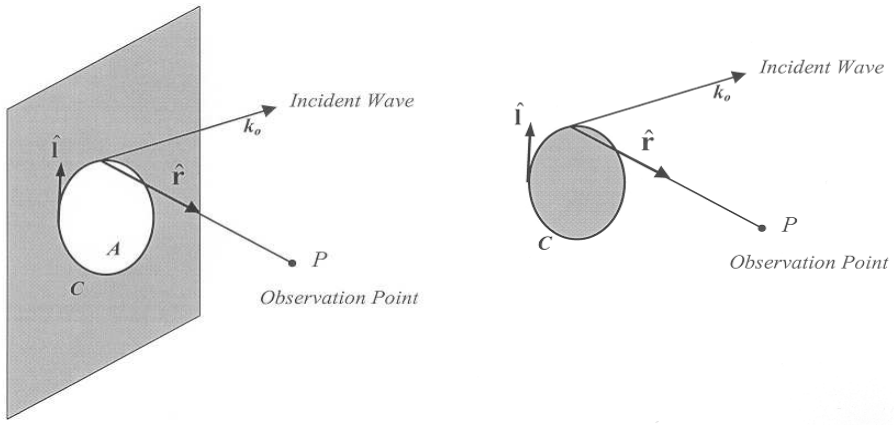


Figure 3.7: Graphical application of Babinet's principle to the aperture problem.

Kirchhoff's approximation suffers from some mathematical inconsistencies and, in consequence does not satisfy the boundary conditions on the screen surface. However, as discussed in appendix A, this problem is eliminated in the two Rayleigh-Sommerfeld integrals, which do satisfy the boundary conditions ($U = 0$ or $\frac{\partial U}{\partial n} = 0$). Marchand and Wolf extracted the boundary diffraction wave contribution from Rayleigh-Sommerfeld integrals [62]:

$$U^{diff}(r) = -\oint_C U_o \left\{ \frac{[\hat{\mathbf{r}} \times \hat{\mathbf{k}}_\mathbf{o}] \cdot \hat{\mathbf{l}}}{1 - [\hat{\mathbf{r}} \cdot \hat{\mathbf{k}}_\mathbf{o}]} \mp \frac{[\hat{\mathbf{r}}^* \times \hat{\mathbf{k}}_\mathbf{o}] \cdot \hat{\mathbf{l}}}{1 - [\hat{\mathbf{r}}^* \cdot \hat{\mathbf{k}}_\mathbf{o}]} \right\} \frac{e^{ik_o r}}{4\pi r}\, dl \tag{3.9}$$

where $\hat{\mathbf{r}}^*$ is the unit vector pointing from the image point P^* of the observation point P to a point on the diffracting edge. The negative (positive) sign of the expression between brackets in (3.9) refers to the first (second) Rayleigh-Sommerfeld formula (see appendix A for details).

Considering plane wave incidence, asymptotic expansion of the Rayleigh-Sommerfeld versions of Rubinowicz boundary wave also consists of the sum of the contributions from a finite number of stationary points on the edge [71]:

$$U^{diff}(r_s) \simeq -\frac{1}{2} \sum_s U_o(l_s) \left\{ \frac{[\hat{\mathbf{r}}_\mathbf{s} \times \hat{\mathbf{k}}_\mathbf{os}] \cdot \hat{\mathbf{l}}_\mathbf{s}}{1 - [\hat{\mathbf{r}}_\mathbf{s} \cdot \hat{\mathbf{k}}_\mathbf{os}]} \mp \frac{[\hat{\mathbf{r}}^*_\mathbf{s} \times \hat{\mathbf{k}}_\mathbf{os}] \cdot \hat{\mathbf{l}}_\mathbf{s}}{1 - [\hat{\mathbf{r}}^*_\mathbf{s} \cdot \hat{\mathbf{k}}_\mathbf{os}]} \right\} \frac{e^{i(kr_s + \frac{\pi}{4})}}{\sqrt{2\pi k r_s)}} \tag{3.10}$$

Application of Babinet's principle to Kirchhoff aperture problem of figure C.2(a) yields the scalar solution for diffracted rays from the aperture complementary plate C.2(b), which differs from equation (3.10) only by a sign [71]. Furthermore, considering the diffraction of a plane wave normally incident on a perfectly reflecting wedge, with the geometry of figure 3.8, only one ray reaches any observation point, and only one term of the summation (3.10) remains. In such situation we can further substitute, based on the geometry of figure 3.8 and noting that $\varphi* = 2\pi - \varphi$, the following relations:

$$\frac{[\hat{\mathbf{r}}_\mathbf{s} \times \hat{\mathbf{k}}_\mathbf{os}] \cdot \hat{\mathbf{l}}_\mathbf{s}}{1 - [\hat{\mathbf{r}}_\mathbf{s} \cdot \hat{\mathbf{k}}_\mathbf{os}]} = tan\frac{\varphi - \varphi_o}{2} \qquad \frac{[\hat{\mathbf{r}}^*_\mathbf{s} \times \hat{\mathbf{k}}_\mathbf{os}] \cdot \hat{\mathbf{l}}_\mathbf{s}}{1 - [\hat{\mathbf{r}}^*_\mathbf{s} \cdot \hat{\mathbf{k}}_\mathbf{os}]} = tan\frac{\varphi + \varphi_o}{2}$$

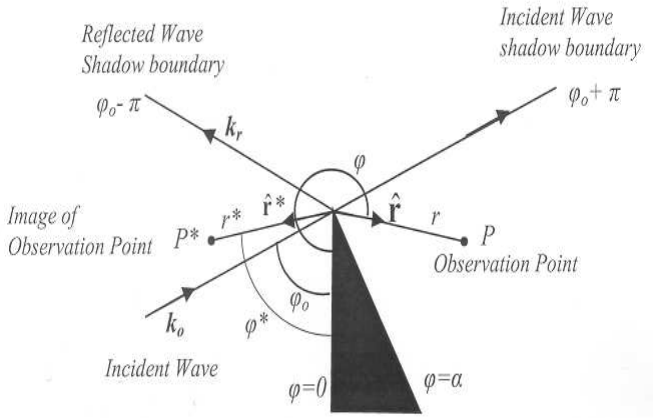


Figure 3.8: Geometry of the wedge diffraction problem (for $\alpha = 2\pi$ the wedge becomes a half-plane).

Asymptotic evaluation of the Physical Optics approximation applied to the fields diffracted by Sommerfeld's half-plane was carried out by Ufimtsev in his Physical Theory of Diffraction monograph [73]. He consider a plane wave normally incident on the perfectly reflecting wedge of figure 3.8. Ufimtsev's results showed that the total fields diffracted by the half-plane had a similar expression to (3.4), with the first two terms corresponding to the incident and reflected waves, each obeying geometrical optics laws, and an edge diffraction term in the form of cylindrical waves originating at the boundary. If the incident wave forms an angle φ_o with the plane of the wedge, then the lines $\varphi = \pi + \varphi_o$ and $\varphi = \pi - \varphi_o$ are the boundaries of the incident and reflected plane waves, respectively. The diffraction coefficient $D(\varphi, \varphi_o)$ multiplying the term representing the boundary diffraction waves, however, differs from that of the exact solution obtained by Sommerfeld (3.5). In particular, Ufimtsev found:

$$D_s^{PO}(\varphi, \varphi_o) = \frac{sin\varphi_o}{cos\varphi + cos\varphi_o} \qquad D_h^{PO}(\varphi, \varphi_o) = -\frac{sin\varphi}{cos\varphi + cos\varphi_o} \tag{3.11}$$

The subscript s refers to the case of the boundary condition $U = 0$ (Dirichlet's problem), also known as hard boundary conditions and corresponding to the E-polarization in Sommerfeld's problem, and the subscript h refers to the case of

the boundary condition $\frac{\partial U}{\partial n} = 0$ (Neumann's problem), known as soft boundary conditions and corresponding to the H-polarization of Sommerfeld's problem. It is easy to show that:

$$-\frac{1}{2}\{tan\frac{\varphi-\varphi_o}{2} - tan\frac{\varphi+\varphi_o}{2}\} = \frac{sin\varphi_o}{cos\varphi + cos\varphi_o} = D_s^{PO}(\varphi,\varphi_o)$$

$$-\frac{1}{2}\{tan\frac{\varphi-\varphi_o}{2} + tan\frac{\varphi+\varphi_o}{2}\} = -\frac{sin\varphi}{cos\varphi + cos\varphi_o} = D_h^{PO}(\varphi,\varphi_o)$$

hence confirming the equivalence between Kirchhoff's boundary conditions applied to Rayleigh-Sommerfeld first and second integrals, and the Physical Optics approximation for soft and hard boundary conditions, respectively. Note though, that the diffraction coefficients obtained through the approximation (3.11) differ from those of the exact solution obtained by Sommerfeld (3.5) and, therefore provide inaccurate results for the diffraction due to the edges. Regardless of its approximated nature, Rubinowicz's results can be considered as the foundation of the theory of edge diffracted waves [71].

On the geometrical optics boundaries, both Rubinowicz's, equation (3.11), and Sommerfeld's, equation (3.5), diffraction coefficients become singular and the asymptotic expansion losses its physical meaning. The field in the vicinity of these boundaries has a rather more complicated structure than the cylindrical wave and must be evaluated by means of the Fresnel integral.

3.3 Physical Theory of Diffraction

Asymptotic evaluation of the Physical Optics approximation (3.8) applied to Sommerfeld's half plane was observed in section 3.2.2 to yield a solution in terms of the incident and reflected geometrical optics waves, plus an edge diffraction term similar to Sommerfeld's exact solution, but which differed from it by the

expression for the diffraction coefficient (3.11). Ufimtsev observed this similarity between the exact and approximate solutions and reformulated Sommerfeld's results to develop his Physical Theory of Diffraction (PTD) for the analysis of high frequency scattering problems [73]. In PTD, the field scattered by an object is also considered as radiated by the current induced on the object surface by the incident field. The main concept of the theory is the separation of the total current ($\mathbf{J}^t$) into two components, a *uniform* current ($\mathbf{J}^{P0}$) associated to the Physical Optics approximation and a *nonuniform* or *fringe* current ($\mathbf{J}^{FR}$) attributed to diffraction due to deviations of the surface from an infinite plane.

$$\mathbf{J}^t = \mathbf{J}^{P0} + \mathbf{J}^{FR} \tag{3.12}$$

The essence of Ufimtsev's PTD is to improve the Physical Optics evaluation of the scattered field by perfectly conducting objects at high-frequency by adding a correction term to the field (or equivalently, the current) distribution on the wedge surface, the "nonuniform" component, due to the presence of the edge discontinuity.

At high frequencies and smooth (in terms of λ) curved edges, diffraction can be assumed a local phenomena, and Sommerfeld integral solutions for the canonical wedge problem can be used to determine the expression for these nonuniform currents in the vicinity of each point of the edge. For local plane wave incidence, the field radiated by $\mathbf{J}^t$ was found by Sommerfeld in the form of integrals in complex plane. Ufimtsev [73] found the field radiated by $\mathbf{J}^{P0}$ also as integrals in the same complex plane and subtracted them from the Sommerfeld integrals to deduce the field radiated by $\mathbf{J}^{FR}$.

3.3.1 Elementary Edge Waves

This nonuniform current, expressed in terms of the same kind of integrals in the complex plane, concentrates mainly near the edge, but its decreasing tail extends to infinite distance. Hence for each differential length of the edge, the field radiated by this current must be integrated over an infinitely narrow and infinitely long strip on both wedge surfaces with common boundary C as indicated in figure 3.9(a). Ufimtsev called this differential element of field radiation "Elementary Edge Wave" (EEW), dE^{FR}, and observed that, in the limit in which $kr \longrightarrow \infty$, only elementary integration strips oriented along the Rubinowicz-Keller diffraction direction yielded solutions with "true physical meaning", in the form of spherical waves originating at the edge [74] multiplied by trigonometrical directional factors. Linear superposition of these elementary spherical waves emanating from each infinitely small element of the edge provides the total field scattered by the object curved edge for all angles of observation. Hence the surface integral is reduced asymptotically to a line integral over the edge.

$$\mathbf{E}^t = \mathbf{E}^{PO} + \mathbf{E}^{FR} \qquad\qquad \mathbf{H}^t = \mathbf{H}^{PO} + \mathbf{H}^{FR} \tag{3.13}$$

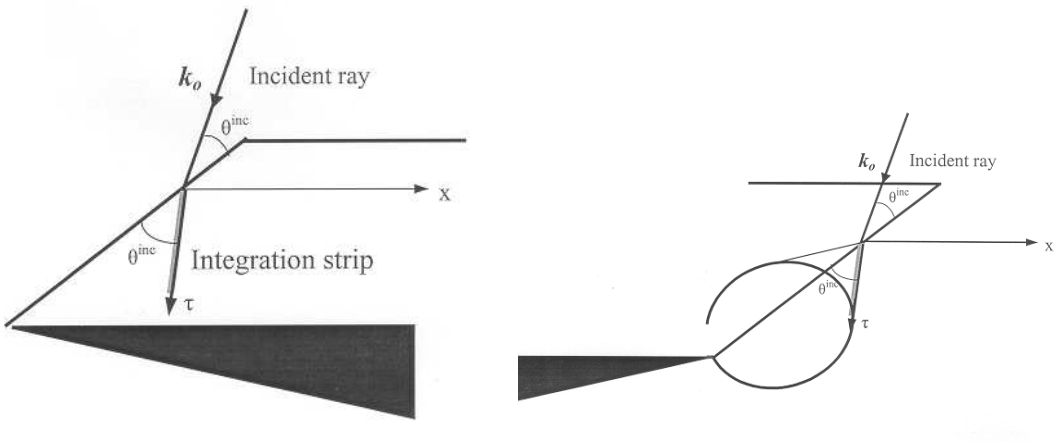
with

$$\mathbf{E}^{FR}(\mathbf{r}) = \int_{C'} 2\mathbf{E}^{EW}(\mathbf{r}, \mathbf{r}')G(\mathbf{r}, \mathbf{r}')\, dl' \tag{3.14}$$

and

$$\mathbf{E}^{EW}(\mathbf{r}, \mathbf{r}') = (\mathbf{E}_{inc} \cdot \hat{\mathbf{l}}')\mathbf{D}_{\mathbf{s}}^{EW}(\mathbf{r}, \mathbf{r}') + \eta\, (\mathbf{H}_{inc} \cdot \hat{\mathbf{l}}')\mathbf{D}_{\mathbf{h}}^{EW}(\mathbf{r}, \mathbf{r}') \tag{3.15}$$

In equation (3.14), $\mathbf{E}^{EW}$ is the elementary wave radiated by each differential arc-length dl' of C', and the unit vector $\hat{\mathbf{l}}'$ in equation (3.15) is tangent to the edge at every point. $G(\mathbf{r}, \mathbf{r}')$ is the usual free-space Green's function defined by



(a) (b)

Figure 3.9: (a) Geometry of the wedge tangential to each point of the object edge. The integration strips on the wedge surfaces along the diffraction cone are responsible for the Elementary Edge Waves diffracted by the object boundary. (b) Through Babinet's principle the integration strips are taken on the wedge complementary aperture.

equation (2.5) and $E_{l'}^{inc}$ and $H_{l'}^{inc}$ represent the incident field component parallel to the boundary. Asymptotic expressions for the functions $\mathbf{D_s^{EW}}$ and $\mathbf{D_h^{EW}}$ valid at every point in space were provided by Ufimtsev in terms of the incident and observation angles [74] and are collected in appendix C. When the observation point is located on the diffraction cone, the ray asymptotic of the scattered field reduce to the familiar results from Sommerfeld's solution in terms of cylindrical or conical wavefronts diverging from the edge:

$$E_z^{FR}(r) = E_{0z}\ D_s^{FR}(\varphi, \varphi_0)\ \frac{e^{i(kr+\frac{\pi}{4})}}{\sqrt{2\pi kr}} \tag{3.16}$$

$$H_z^{FR}(r) = H_{0z}\ D_h^{FR}(\varphi, \varphi_0)\ \frac{e^{i(kr+\frac{\pi}{4})}}{\sqrt{2\pi kr}} \tag{3.17}$$

where, from equations (3.5) and (3.11),

$$\begin{aligned} D_s^{FR}(\varphi,\varphi_o) &= D_s(\varphi,\varphi_o) - D_s^{PO}(\varphi,\varphi_o) \\ D_h^{FR}(\varphi,\varphi_o) &= D_h(\varphi,\varphi_o) - D_h^{PO}(\varphi,\varphi_o) \end{aligned} \tag{3.18}$$

An alternative Method of Equivalent Currents (MEC) was developed recently [75, 76], which in essence is a different interpretation of the PTD line integrals in which the integrands are identified as fictitious equivalent currents flowing along the edge rather than elementary edge waves. Specifically:

$$\mathbf{E}^{FR}(\mathbf{r}) = \int_{C'} \{ -j\omega\mu \mathbf{J}^{FR}(l') G(\mathbf{r},\mathbf{r}') + \frac{1}{j\omega\epsilon} (\mathbf{J}^{FR}(l')\cdot\nabla')\nabla' G(\mathbf{r},\mathbf{r}') + \mathbf{M}^{FR}(l')\times\nabla' G(\mathbf{r},\mathbf{r}')\} dl' \tag{3.19}$$

Equation (3.19) results from the asymptotic evaluation of the surface integrals over two infinitely narrow strips along the direction of Keller diffraction cone, axis τ in figure 3.9, on both surfaces of the wedge with common boundary C:

$$\mathbf{E}^{FR}(\mathbf{r}) = \int_{C'} \int_0^{\infty} \{ -j\omega\mu \mathbf{J}^{FR}(\mathbf{r}') G(\mathbf{r},\mathbf{r}') + \frac{1}{j\omega\epsilon} (\mathbf{J}^{FR}(\mathbf{r}')\cdot\nabla')\nabla' G(\mathbf{r},\mathbf{r}') + \mathbf{M}^{FR}(\mathbf{r}')\times\nabla' G(\mathbf{r},\mathbf{r}')\}\ sin\theta^{inc} d\tau'\ dl' \tag{3.20}$$

Close form expressions for $\mathbf{J}^{FR}$ and $\mathbf{M}^{FR}$ can be obtained through the asymptotic surface-to-edge integral reduction of the inner integrals over the coordinate τ along the integration strip, given that the main contribution arises at the endpoint $\tau = 0$. This was accomplished independently by several authors who provided expressions for the equivalent electric and magnetic line integrals in terms of the incident field and simple trigonometric functions [75, 76, 77]. A comparative study of these three PTD-MEC formulations can be found in [78].

3.3.2 PTD on Rectangular Aperture

The Physical Optics approximation on conducting surfaces is invariably linked to Kirchhoff integral on an aperture through Babinet's principle [8, 51], also

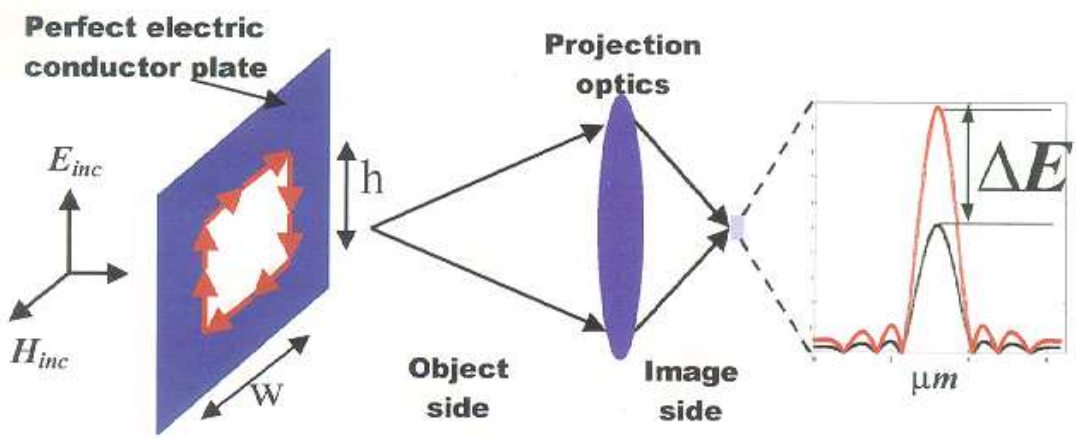


Figure 3.10: Application of PTD to rectangular aperture on perfect electric conducting plane and imaging through a $4X$ $NA = 0.85$ optical system. Relative error on the field amplitude due to edge diffraction is measured at the peak of the image.

referred to as the principle of duality. According to the equivalence theorem [52] over the free-space surface, the role of the true surface current is overtaken by the equivalent electric and magnetic currents: $\mathbf{J}_{eq} = \hat{\mathbf{n}} \times \mathbf{H}$ and $\mathbf{M}_{eq} = \mathbf{E} \times \hat{\mathbf{n}}$, where $\hat{\mathbf{n}}$ is the exterior normal to the surface. It follows that a correction in the spirit of PTD can be applied to Kirchhoff formula by considering the presence of some fringe equivalent currents on the aperture and integrating the effect over infinitely narrow strips lying on the aperture surface as in figure 3.9(b) [79]. In the asymptotic limit, $k \longrightarrow \infty$, equation (3.14) can be applied to calculate the aperture fringe fields by proper transformations according to Babinet's principle. In appendix C this is done for the rectangular aperture of width w and height h on the perfect conducting plate of figure 3.10. The resultant far field was then simulated to pass through an optical projection systems of $4X$ reduction factor as described in chapter 2, such that, even with high NA, the angle formed by the diffraction orders collected at the entrance pupil can be assumed small. Under

this circumstances, it is observed that the correction term of the field attributed to the fringe fields obeys, at the center of the image plane (x=y=0), the following relation:

$$\mathbf{E}_{image} = \mathbf{E}^{PO}_{image} + \mathbf{E}^{FR}_{image} = \mathbf{E}^{K}_{image} + \mathbf{E}^{FR}_{image} = \mathbf{E}^{K}_{image}\big(1 + \frac{\Delta\mathbf{E}}{\mathbf{E}}\big)$$

with

$$\frac{\Delta\mathbf{E}}{\mathbf{E}} = \frac{\mathbf{E}^{FR}_{image}}{\mathbf{E}^{K}_{image}} \approx -\frac{j}{2k}\frac{\big(2h+2w\big)}{h\ w} = -\frac{j}{2k}\frac{4}{\frac{2hw}{w+h}} = -j\frac{4\Delta d}{d_{effective}} \tag{3.21}$$

$$or \qquad = -j\frac{(2h+2w)\Delta d}{hw} = -j\frac{Boundary\ Layer\ Area}{Total\ Area} \tag{3.22}$$

Thus the relative error on the electric field amplitude due to the fringe fields, which are ignored in the Kirchhoff approximation, is inversely proportional to the effective opening size measured as the harmonic mean of the opening width (w) and height (h). The harmonic mean is defined as twice the product of the sides divided by its sum[1]. In other words, this inverse dependence of the relative error on the opening size can be expressed as the ratio of the area covered by a strip of width Δd placed along the edges of the openings to the total area of the opening as outlined in equation (3.22). This last interpretation of equation (3.22) suggests a transformation of the line integral (3.14) or (3.19) along the object boundary to a surface integral similar to (3.20) over a "Boundary Layer" of width Δd along the same edge, but where the integrand is the same as in the Physical Optics or Kirchhoff approximation, that is, a transformation in the form:

$$\mathbf{E}^{FR}(\mathbf{r}) = \beta \int_{C'}\int_{0}^{\Delta d} \big\{ -j\omega\mu\mathbf{J}^{K}(\mathbf{r}')G(\mathbf{r},\mathbf{r}') + \frac{1}{j\omega\epsilon}(\mathbf{J}^{K}(\mathbf{r}')\cdot\nabla')\nabla' G(\mathbf{r},\mathbf{r}') + \mathbf{M}^{K}(\mathbf{r}')\times\nabla' G(\mathbf{r},\mathbf{r}')\big\}\, dx'dl' \tag{3.23}$$

for normal incidence such that $\theta_{inc} = \frac{\pi}{2}$ and $d\tau = dx$. Moreover, the equivalent currents in the integrand are given now by $\mathbf{J}^{K} = \hat{\mathbf{n}} \times \mathbf{H}^{inc}$ and $\mathbf{M}^{K} = \mathbf{E}^{inc} \times \hat{\mathbf{n}}$,

[1]Harmonic mean is defined as: $\dfrac{2}{\left(\frac{1}{h}+\frac{1}{w}\right)}$

and the whole integral is multiplied by a coefficient β which, from equation (3.22), is given by $\beta = -j$ in this example.

Equation (3.23) is written in a form that resembles Franz's diffraction formula (2.11), utilized in chapter 2 to evaluate the field at the entrance pupil. It can also be interpreted as a correction width, Δd, applied on all four sides of the thin mask model to account for the error due to the thin mask approximation.

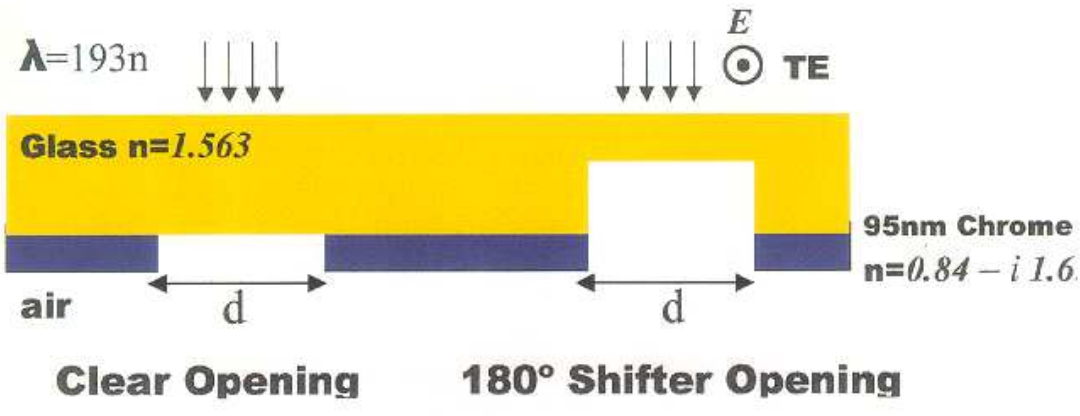
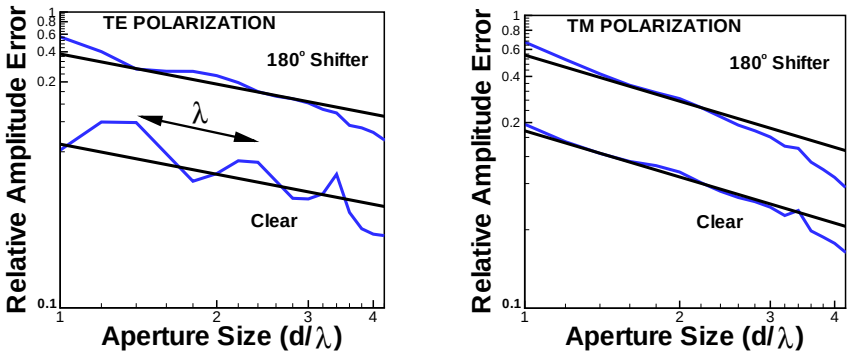


Figure 3.11: Evaluation of the relative error in amplitude due to the Kirchhoff approximation on apertures on Thick Chrome Masks.

The chrome layer where the photomask is patterned, however, has a finite index of refraction at the wavelengths of lithography operation and cannot be regarded as a perfect conductor. In addition it has a finite thickness and is covered by a thick layer of glass which is etched to create phase-shifting features on the mask. Rigorous evaluation of the field on the mask surface, for both clear and 180^o-shifter 2D openings, of the 95nm thick chrome layer of figure 3.11 was performed by the Finite-Different Time-Domain software Tempest 6.0 [20]. The relative error in amplitude was measured at the peak of the image field and, as can be observed in figure 3.12, the same reciprocal relation with the size of the opening hold for apertures on thick masks. Figure 3.12(a) exhibits resonant



(a) (b)

Figure 3.12: The inverse law on the opening dependence of the relative image field amplitude due to edge diffraction effects holds for Thick Masks of finite chrome thickness and index of diffraction. Resonant effects are reduced and almost eliminated for 180^o-shifter openings.

effects due to multiple diffraction interference between edges, which are more significative for clear openings with TE-polarized illumination. TE polarization in figure 3.12 refers to the electric field being parallel to the opening edges and normal to the plane of the page, and TM polarization on the same figure refers to the electric field lying on the plane of the page normal to the opening edges.

3.4 Boundary Layer Approximation

The conclusions of section 3.3.2 provide the basis for the development of a simple "Boundary Layer" model to account for thick mask effects in optical lithography. This model consists of a sophisticated version of the Kirchhoff or Thin Mask approximation, simply adding a layer of fixed width at every edge. The boundary layer parameters are based on the comparison of the electric field on the wafer

produced by various thick phase-shifting masks against that produced by their corresponding ideal thin mask models. Systematic simulations of rectangular openings with side lengths ranging from 1λ to 6λ, aspect ratios ranging from 3:1 to 1:3, and three different etching profiles per opening size: 0^o, 180^o shifter and 180^o shifter with 35nm undercut, provided the means to extract those parameters [80, 81]. Rigorous electromagnetic results for the fields on the masks were obtained from 3D FDTD Tempest [27] simulations performed with coherent, on-axis and off-axis illumination, and with a grid of 40 nodes per wavelength (more than twice as fine as recommended) to ensure accuracy. Figure 3.13 shows a sample of the simulated mask apertures.

Full vectorial formulation of the imaging process, developed in chapter 2, was applied to the evaluation of the aerial field distribution produced by both rigorous em solutions of the mask field and those due to the thin mask approximation. Two sets of illumination parameters were evaluated: a $4X$ projection system with NA equal to 0.68, operating at $248nm$ illumination wavelength corresponding to case I of Table 3.1; and a similar $4X$ system with $NA = 0.85$, operating at $193nm$ and corresponding to case II in Table 3.1.

The indices of refraction used for glass and chrome are 1.5 and $2.5 - i2$, respectively, at 248nm, and 1.563 and $0.84 - i1.65$ at 193nm. Finally, a chrome layer thickness of 95nm was employed at 193nm wavelength such that it provided the same absorption of incident light as the 80nm chrome layer used at 248nm wavelength. Section 4.4 analyzes the variations of the mask model parameters for different chrome thicknesses.

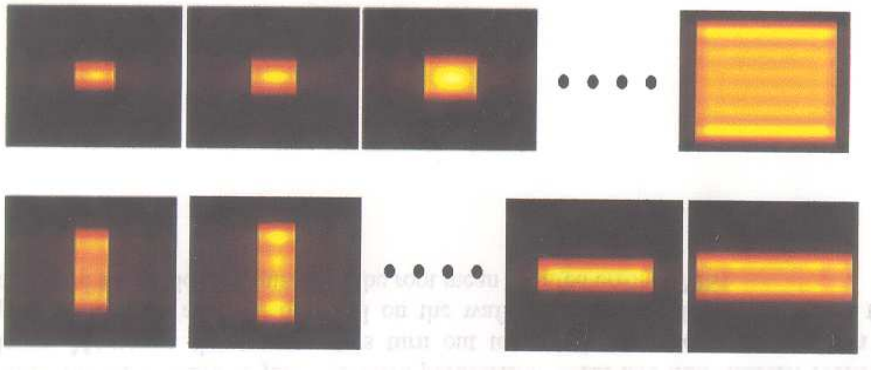


Figure 3.13: Example of the mask patterns simulated with the Finite-Difference Time–Domain program Tempest [20] for clear openings with side lengths ranging from λ to 6λ and aspect ratios from 1:3 to 3:1.

3.4.1 Relative Error in Amplitude

The relative error in aerial field amplitude produced by the thin mask approximation was measured by the deviations of its real component from the rigorously evaluated electromagnetic field. As indicated by figure 3.14, the error measurements were performed over the main polarization component, as provided by the simplified expression for the fields at the entrance pupil (2.37), at its peak value.

In agreement with the results of section 3.3.2, the deficit on the real part of the field main polarization component was observed to follow a reciprocal dependence on the size of the opening, for both square and rectangular openings, as long as it was measured as the harmonic mean of its width (w) and its height (h). This "inverse law" on the mean size of the opening is plotted in figure 3.15(a) for the case of 248nm illumination with a $NA = 0.68$ optical system. Weighting the width and height unequally or combining them by a formula other than the harmonic mean produced a poorer fit to the inverse law.

As outlined in equation (3.24), this inverse dependence of the real component

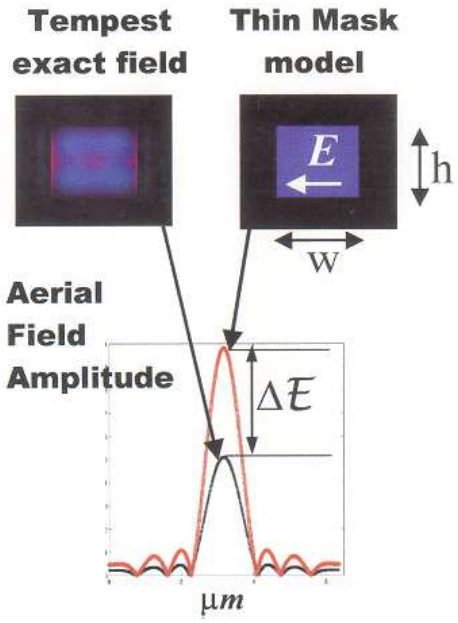


Figure 3.14: Error measurement on aerial image field amplitude due to the Thin Mask approximation relative to the rigorously evaluated mask field.

of the error on the opening size indicates that it can be assumed proportional to a correction width, Δd, placed along the edges of the thin mask model of the opening, independently of its size and regardless of polarization. This yields the real part of the Boundary Layer model as displayed in figure 3.17(a), where an opaque layer of zero transmission is placed on all four sides of the opening to account for amplitude errors.

$$Amplitude\ \ deficit = Re\left\{\frac{\Delta \mathbf{E}}{\mathbf{E}}\right\} = -\frac{4\Delta d}{d} = -\frac{4\Delta d}{\frac{2wh}{w+h}} \tag{3.24}$$

The main conclusion to be derived from the observed "inverse law" on the opening mean size is that the relative error in amplitude due to Kirchhoff approximation, and the consequently boundary layer correction width, can be inter-

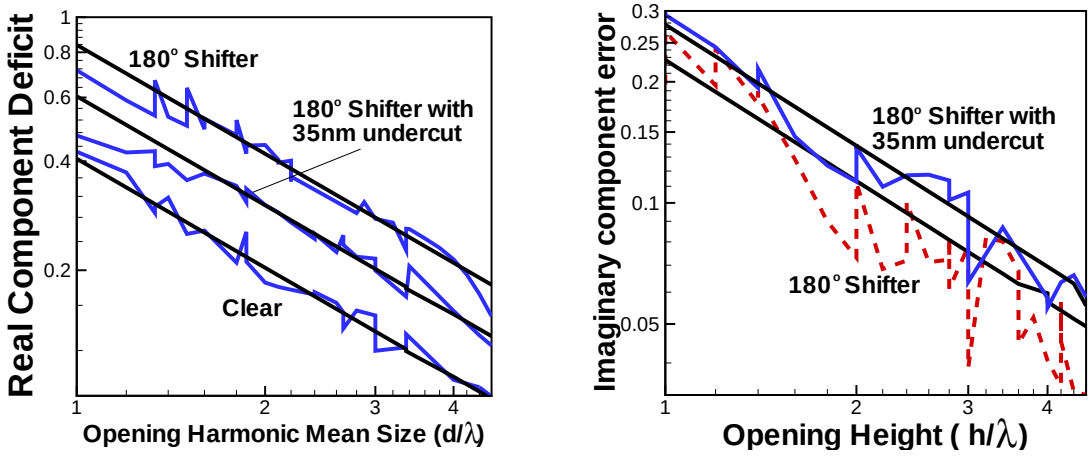


Figure 3.15: (a) Log-log plot of the relative error in the real component of the electric field on the wafer produced by the thin mask approximation, as compared to the rigorously evaluated EM field, versus the harmonic mean of the opening height and width. (b) Log-log plot of the relative error in the imaginary component as a function of opening height (opening size in the direction perpendicular to polarization).

preted, to a good approximation, as an intrinsic edge property. In other words, thick mask effects can be interpreted as an edge property analogous to the edge diffraction waves initially suggested by Young [58].

3.4.2 Relative Error in Phase

Relative errors of phase produce non-zero values of the imaginary part of the aerial field relative to the ideal 0^o or 180^o phase of the thin mask field. The imaginary component of the electric field error will then account for the phase deviations of the model. Following the same procedure as in section 3.4.1, it was observed that only when plotted versus the opening height h (dimension normal to polarization) did the imaginary component of the error follow the same "inverse law" with the opening size. This conclusion is evident from the results displayed in figure 3.16

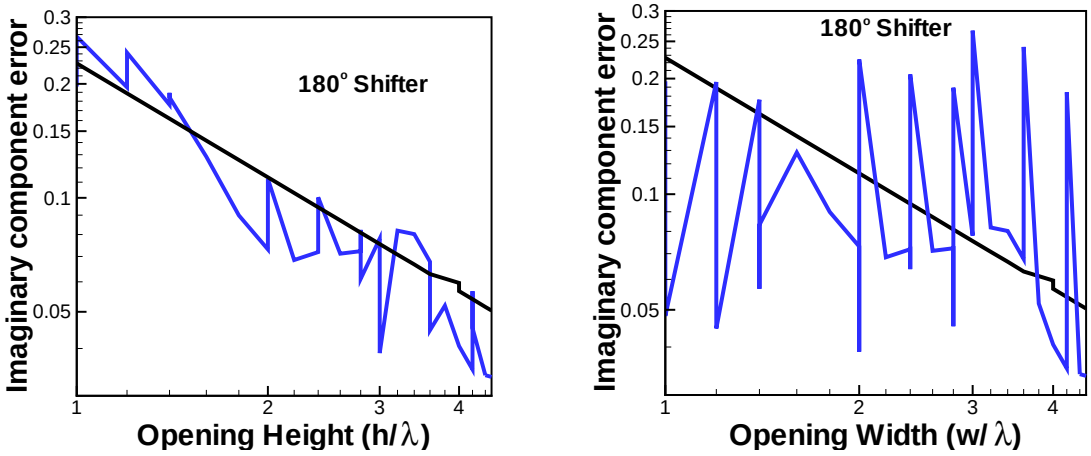


Figure 3.16: (a) Log-log plot of the relative error in the imaginary electric field component versus the opening dimension normal to polarization (h). (b) Log-log plot of the relative error in the imaginary electric field component versus the opening dimension parallel to polarization (w).

of the same imaginary component errors versus either the dimension normal or parallel to polarization. Only the former produces a good fit to the inverse law. Any other functional dependence of the error such as the one of figure 3.16(b) or any average of the opening sides produced the same bad fit to the inverse law. Figure 3.15(b) is a plot of the imaginary error observed on both 180^o-phase shifter and shifter with 35nm undercut (clear openings produced negligible phase errors) for the 248nm illumination with $NA = 0.68$ case.

This result, formulated in equation (3.25), yields the conclusion that the phase errors are compensated by a purely imaginary transmission coefficient, $-j\beta$, through the boundary layers that lie parallel to the electric field polarization when rectangular features are considered. The imaginary part of the final boundary layer model, displayed in figure 3.17(b), compensates for phase devia-

tions of the Thin Mask model.

$$Relative\,Imaginary\,Error = Im\{\frac{\Delta \mathbf{E}}{\mathbf{E}}\} = \beta\frac{2\Delta d}{h}. \tag{3.25}$$

On arbitrary geometries, the Boundary Layer has an imaginary transmission coefficient proportional to the absolute value of the cosine angle between the electric field polarization and the edge direction, reaching its maximum value when the electric field is tangent to the boundary, and diminishing to zero when the field is normal to it. This simple geometric rule (cosine rule) accounts for the orientation dependence of the boundary conditions at the metal edges. Tangential components of the electric field must vanish on metal surfaces, setting up conditions for standing-wave effects on the aperture field that contribute to the phase errors. Normal electric field components, on the other hand, exhibit a discontinuity due to the accumulation of charges on the chrome surface, but its contribution to the final projected image is mostly filtered out by the optical lens [82].

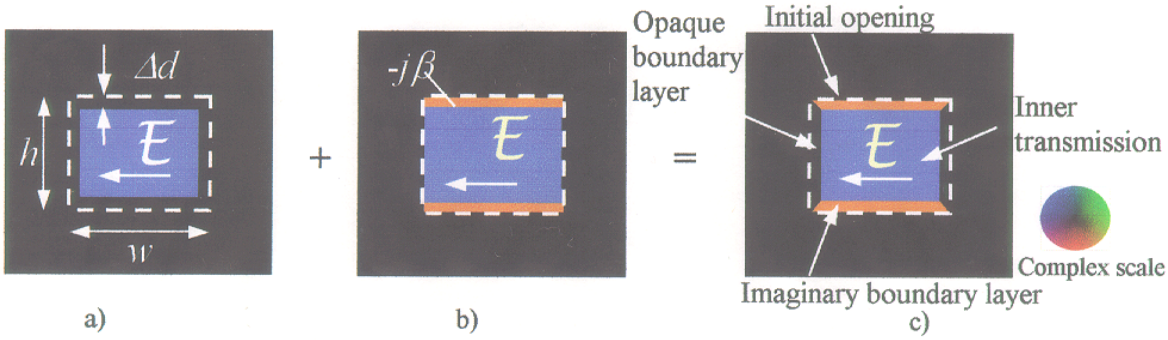


Figure 3.17: (a) Real component of the Boundary Layer model. (b) Imaginary component of the Boundary Layer model. (c) Final Boundary Layer model as the superposition of both real and imaginary parts.

Final superposition of both the real and imaginary parts of the model provides the final boundary layer model, sketched in figure 3.17(c). The key result of our simulations is that the thick mask effects can be interpreted, to a good

approximation, as an intrinsic edge property, and modeled with just two fixed parameters: width and transmission coefficient of a locally-applied boundary layer. As sketched in figure 3.18, the width of the boundary layer controls the variation in peak amplitude, while the imaginary transmission coefficient corrects phase deviations of the thick mask. These errors differ for different types of mask edge cross-sections, giving place to different boundary layer parameters to be applied locally in the vicinity of each type of mask edge profile. Thus each mask edge specification (clear, shifter and shifter with undercut) has its own set of boundary layer parameters.

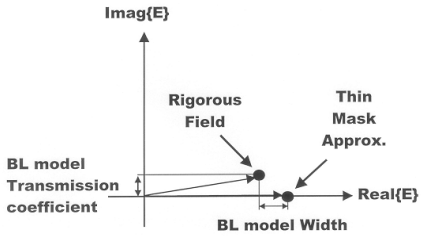


Figure 3.18: Simple graphical notion of the effect of the boundary layer on the complex electric field plane. The width of the boundary layer controls the variation in peak amplitude, while the imaginary transmission coefficient corrects phase deviations of the thick mask.

3.5 Boundary Layer Parameters

The boundary layer parameters of width and transmission coefficient are derived, for each etching profile, from the slope of the best fitting straight line to the data points, in a least squares sense, and further optimized to minimize the errors. Data was collected to determine the boundary layer parameters of the three types of etching profiles commonly used in Alt. PSM and illustrated in figure 3.2(a): clear and 180^o phase-shifting opening with and without 35nm undercut, for the

Table 3.1:

Width and transmission coefficients of the boundary layer model for different type of openings in typical Alt. PSM geometry (cases I and II refer to $248nm$ and $193nm$ wavelength imaging systems, respectively)

Opening type	Boundary Width (nm)		Tangential Boundary Transmission		Normal Boundary Trans.	Interior Trans.	Minimun opening (nm)	
	I	II	I	II			I	II
Clear	24.8	14.5	0.0i	0.8i	0	1	248	200
Shifter	55.8	53.0	-0.52i	-0.30i	0	-1	300	250
Shifter with 35nm undercut	37.2	33.7	-0.66i	-0.635i	0	-1	350	200

two imaging configurations indicated in the beginning of section 3.4, corresponding to cases I and II of table 3.1. Only the etched mask profiles of case I yielded nonzero imaginary coefficients since phase errors for unetched openings were too small to show systematic variations. Note that a slightly different set of parameters could have been obtained by allowing the boundary layer to take complex values rather than being purely opaque or with an imaginary transmission coefficient. Requiring purely imaginary boundary layer transmission leads to the simplest type of model yet providing satisfactory results in terms of root mean square errors of the field on the wafer.

3.6 Discussion

The Thin Mask approximation, commonly employed in optical lithography simulations, is becoming increasingly inaccurate as the mask dimensions shrink with

every technology node and the mask topography exhibits higher aspect ratios. This approximation arises from the application of Kirchhoff's boundary conditions on the mask surface, which ignore the effect of the chrome edges on the electromagnetic field propagating through the mask aperture.

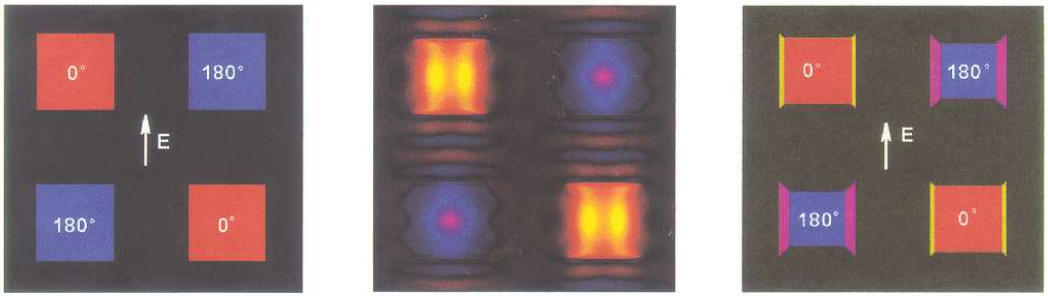
In this chapter we observed the relation, through Babinet's principle, between the Physical Optics approximation and Kirchhoff's boundary conditions. The Physical Theory of Diffraction is a technique developed to correct for edge effects in the application of the Physical Optics approximation on high-frequency scattering by perfect conducting surfaces. It considers the effect of the edges to arise from some "fringe" line currents flowing along the object border. Lithographic reticles, however, cannot be regarded as perfect conductors given the finite index of refraction of the chrome layer at the operation wavelength, but a similar "fringe" edge-field concept can be explored to account for the electromagnetic effects.

For rectangular apertures on both perfect conducting plates or more realistic chrome layers of finite thickness and refractive index, we observed how the relative errors of the field real and imaginary components on the wafer follow an inverse law on the opening mean size and height, respectively, what allowed us to reduce the model to a simple boundary layer of fixed width and transmission coefficient. The proposed model, therefore, consists of a sophisticated version of Kirchhoff approximation, simply adding a boundary layer to every edge.

CHAPTER 4

Boundary Layer Model Accuracy

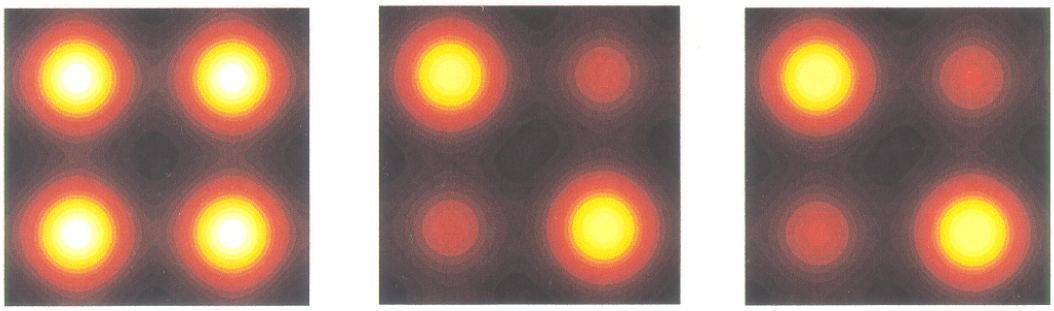
In chapter 3 a new model of the photomask near-fields has been derived, which incorporates electromagnetic effects arising from the vector nature of light and the finite thickness and refractive index of the chrome layer. This model, which we call the *Boundary Layer Model* (BL Model), retains much of Kirchhoff's simplicity but adds a fixed-width strip of imaginary transmission at the boundary of the mask features. Figure 4.1(c) illustrates the corresponding Boundary Layer model of the 2λ half-pitch array of alternating 180^o-shifter and clear square mask openings of figure 3.1, reproduced in figures 4.1(a) and (b), with vertically polarized electric field.



(a) (b) (c)

Figure 4.1: (a) Kirchhoff scalar approximation (thin mask model) of the field on the mask plane of a 2λ half-pitch array of alternating 180^o-shifter and clear square mask openings, with vertically polarized electric field. (b) Sketch of the actual object field obtained by rigorous electromagnetic FDTD Tempest simulation on the same mask. (c) Boundary Layer model for the same mask features and illumination conditions.

A remarkable reduction of the Root Mean Squared error is achieved by the application of this simple boundary layer to the customary Thin Mask approximation. The later produced an error of 85.15% at 193nm for a system of $\sigma = 0.6$ partially coherent illumination and $NA = 0.85$, while only 3.48% RMS error was observed for the former. This improvement is evident from figure 4.2(a)-(c), which is a plot of the aerial images generated by either model as compared to the rigorously evaluated field on the mask.

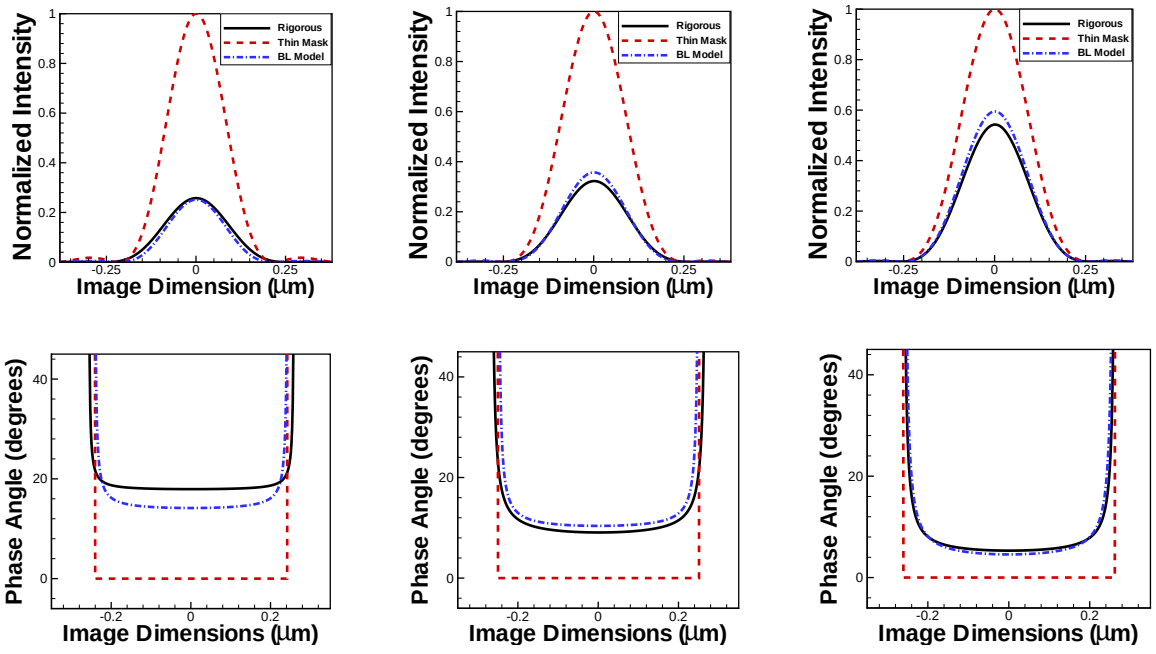


(a) (b) (c)

Figure 4.2: (a) Aerial image intensity distribution at the focal plane obtained from the thin mask approximation of a 2λ half-pitch array of alternating 180^o-shifter and clear square mask openings, and producing a rms error of 85.15%. Unpolarized (TE and TM components were added incoherently), $\sigma = 0.6$ partially coherent illumination at 193nm and $NA = 0.85$ was used. (b) Aerial intensity distribution for the rigorously evaluated object field, and (c) aerial intensity distribution obtained with the Boundary Layer model under identical illumination conditions and producing a rms error of 3.48%.

4.1 Boundary Layer Model with Coherent Illumination

Equally satisfactory are the results in figure 4.3(a)-(c), for the aerial images generated by 180^o phase-shifting isolated square openings of sizes equal to 1.2λ, 1.8λ and 3λ, with normal, coherent illumination. The rigorously simulated fields



(a) (b) (c)

Figure 4.3: Comparison of both total intensity and phase of the wafer images produced by the exact field at the aperture (solid) and both its corresponding Kirchhoff approximation (dash-dot) and Boundary Layer model (dashed), for three different sizes of 180^o-phase-shift square opening: (a) 1.2 λ, (b) 1.8 λ and (c) 3 λ, respectively. Total intensity was calculated by incoherently superposing the contributions of all field components, while the phase corresponds to just the main polarization component.

are contrasted to their corresponding Thin Mask and Boundary Layer models. Simulations were performed at 248nm wavelength with a $4X$ reduction optical system of Numerical Aperture equal to 0.68. The total intensity was calculated by incoherently superposing the contributions of all field components, while the phase plots correspond to just the component along the polarization direction. They reveal a close agreement between the images generated by both the exact field (solid line) and its Boundary Layer model (dash-dotted line) in both ampli-

tude and phase plots, what yields to a remarkable reduction of the error caused by the conventional Kirchhoff approximation (dashed line), even for the smallest opening sizes.

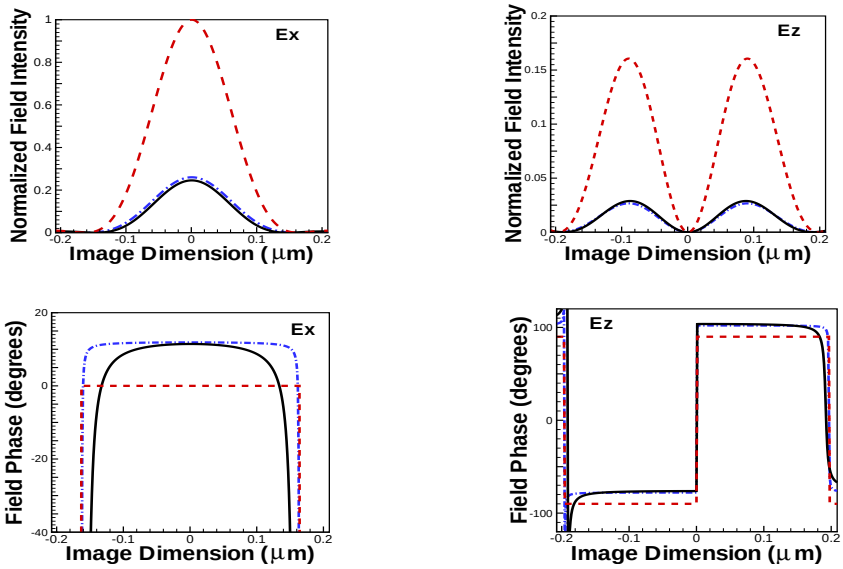
4.1.1 Root Mean Squared Error

The Root Mean Squared (RMS) error of the image intensity integrated over the wafer plane is used to measure the model accuracy. However, given that the photoresist is a non-magnetic material, the squared electric field and not the time averaged Poynting vector, will determine the image intensity distribution. The RMS error was therefore evaluated according to equation (4.1), where the model field refers to both the Thin Mask model and the Boundary Layer Model.

$$Root\ Mean\ Squared\ Error = RMS = \sqrt{\frac{\iint ||\mathbf{E}_{image}|^2 - |\mathbf{E}_{model}|^2|^2 ds}{\iint |\mathbf{E}_{image}|^2 ds}} \quad (4.1)$$

4.1.2 High NA Effects

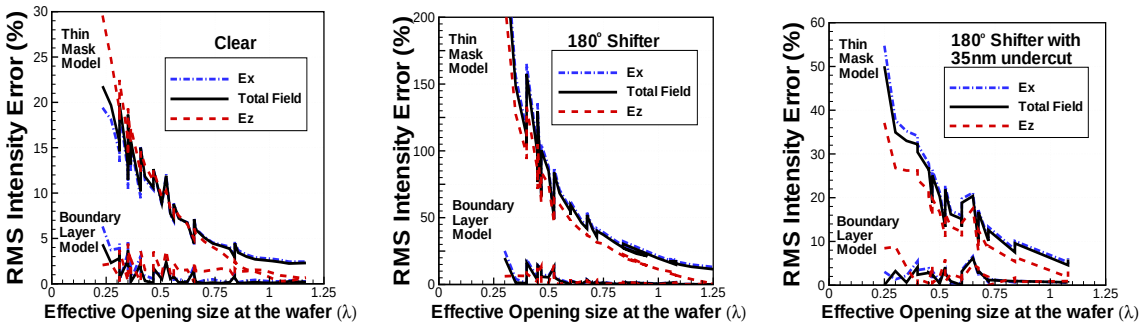
The boundary layer model was optimized for the field component along the direction of polarization of the incident light, E_x, for both amplitude and phase, as illustrated in figure 4.4(a). This approximation, however, provided similar error reductions for all cross-field components arising at the exit pupil of a high NA lens. Figure 4.4(b) illustrates an example of accurate fit of the boundary layer approximation to the rigorous field coupled along the optical axis, E_z. The root mean squared error at the plane of the wafer, measured according to equation (4.1) for both the conventional thin mask model and the boundary layer model, as compared to the actual image, exhibit a remarkable reduction of the error even for mask features close to the wavelength. RMS error calculations for case II of table 3.1 are being displayed for the total electric field as well as for the



(a) (b)

Figure 4.4: Comparison between the aerial field components produced by rigorously evaluated EM Tempest field solutions of the object field (solid) and both the corresponding "thin mask" approximation (dashed) and our Boundary Layer model (dash–dotted), of a 1.6λ 180^o-phase-shift square mask opening. (a) Intensity and phase of the field component along the polarization direction (x-axis) and (b) intensity and phase of the field coupled onto the component along the optical axis (z-axis).

two most significative individual components, E_x and E_z, in figures 4.5(a)-(c). It is observed that the error reduction due to the BL approximation is larger for smaller sizes, while lowering the overall error bound in all sizes. Figures 4.6(a) and (b), on the other hand, are plots of the phase deviations from 180^o of the phase-shifting openings with and without 35nm undercut, at 248nm and 193nm wavelength illumination, respectively. Error reductions of up to a factor of 7 are achieved for the smallest sizes at 248nm (case I of table 3.1), where higher phase deviations were observed. Although the boundary layer is able of compensating phase deviations of the fields, some residual phase error remains, which can be



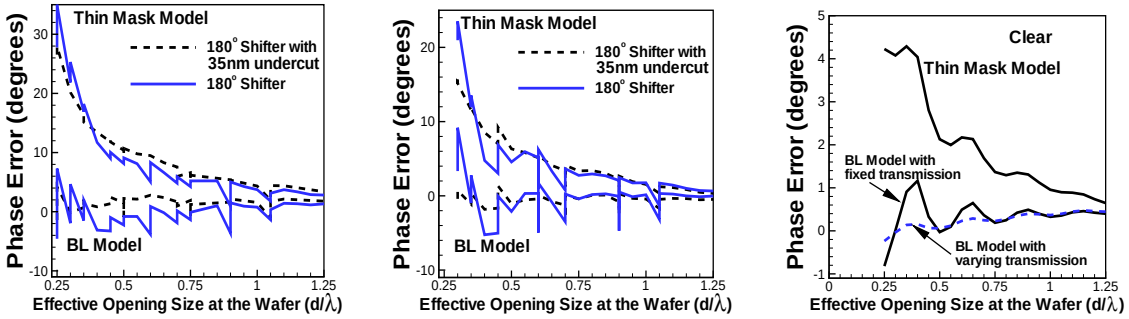
(a) (b) (c)

Figure 4.5: Root mean squared error on the projected image intensity integrated over the wafer plane for the total field and both the E_x and E_z components with coherent normal $193nm$ illumination and $0.85NA$ imaging system.(a) Clear openings, (b)Shifter openings and (c) Shifter with 35nm undercut.

attributed to resonance. To illustrate the effect of the resonance, the phase error for clear, only square openings is plotted in figure 4.6(c) for the imaging conditions of case II of table 3.1. The residual phase error due to these resonances can be further reduced by allowing the imaginary transmission coefficient of the Boundary Layer to vary with the opening size rather than maintaining it constant. This simple modification improves the phase compensation capability of the model as illustrated also in figure 4.6(c).

4.2 Partially Coherent Imaging Formulation

Illumination in industrial applications consists of an unpolarized source of partially coherent light that reduces some of the effects associated with coherent illumination. Contrary to the simple source point generating coherent illumination, partially coherent light results from the spatially extended source aperture



(a) (b) (c)

Figure 4.6: Phase error in degrees for the main polarization component E_x of the projected field at the wafer plane with (a) coherent normal $248nm$ illumination and $0.68NA$ imaging system (Case I of table 3.1), (b) and (c) coherent normal $193nm$ illumination and $0.85NA$ imaging system (Case II of table 3.1). Only square openings are plotted in (c).

about the optical axis. Under the assumption of Köhler illumination configuration described below, two equivalent methods, both based on the spatial discretization of the source into a number of spatially incoherent point sources, are usually utilized in lithography to model imaging with partially coherent illumination. The thermal nature of the light source guaranties that the illumination produced by two such source points is mutually incoherent. In the *Source Integration or Abbe's Method* [16, 17], the coherent images generated by each source point are incoherently added together to produce the final partial coherent image. In the equivalent *Transfer Cross Coefficient or Hopkins Method* [8, 18], the integration over the source is carried out first and the result provides directly the aerial image intensity distribution generated by the partially coherent light.

4.2.1 Illumination Configuration

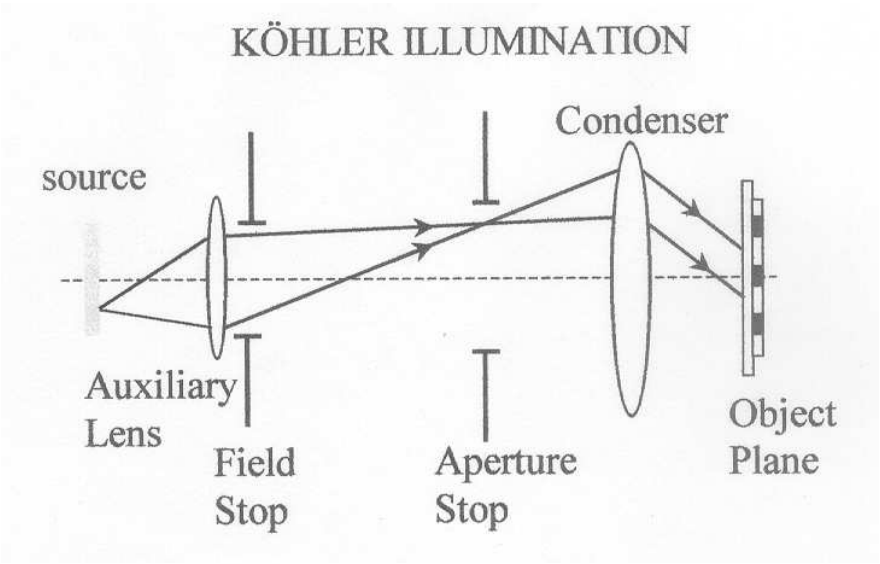


Figure 4.7: With Köhler configuration each source point can be treated as generating a coherent, linearly polarized plane wave of spatial frequency determined by the position of the source point relative to the optical axis.

Köhler illumination configuration [8] predominates in lithography because it provides uniform illumination from a source that in general is non-uniform, provided well-corrected condenser lenses are employed [9]. In Köhler illumination, the source can be considered to be located at the focal plane of the condenser and the object at the condenser exit pupil. As sketched in figure 4.7, if a telecentric condenser is assumed, each of these point sources originates a coherent, linearly polarized plane wave emerging from the lens with an angle determined by the source point location relative to the optical axis [8]. Alternatively, the source cross section can be imaged on the entrance pupil of the projection lens, in which case each source point produces spherical waves converging toward one point on the entrance pupil plane. Given that the distance between the object plane and the entrance pupil in real lithography steppers is at least several wave-

lengths, the curvature of the spherical wavefronts that illuminate the mask can be assumed negligible over several resolution units, allowing their approximation as plane wavefronts [9].

4.2.2 Abbe's Formulation

This procedure, also known as Source Integration method, breaks down the partially coherent imaging system into a set of coherent systems, which, based on equation (2.38), behave as linear systems in field amplitude. The total aerial image is then obtained by the incoherent superposition of all the contributions within the condenser Numerical Aperture (NA_c):

$$I_{image}(x',y',z') = \iint\limits_{p_x^2+p_y^2 \leq NA_c^2} I_{source}(p_x,p_y)\ \mathbf{E_{image}}(x',y';p_x,p_y)\cdot\mathbf{E^*_{image}}(x',y';p_x,p_y)\ dp_x dp_y \quad (4.2)$$

Each point on the source generates an incident plane wave at the object with direction cosines $(p_x, p_y, \sqrt{1-p_x^2-p_y^2})$ and time-averaged intensity $I_{source}(p_x, p_y)$. The aerial image intensity distribution at the wafer plane for each incident plane wave is obtained following the procedure described in chapter 2, that is, after some reasonable approximations, applying equation (2.38) to the electromagnetic field on the object surface produced by each incident plane wave.

The set of plane waves incident on the mask correspond to points in k-space. Figure 1.3 of section 1.2.4 showed the diagrams of some advanced illumination schemes which can all be simulated following the same Abbe's procedure. All results of partially coherent illumination presented in this thesis correspond to circular illumination schemes.

Finally, separate results for both orthogonal (TE and TM) polarizations need to be incoherently superposed in order to emulate unpolarized light.

4.2.3 Hopkins' Method

The second formulation is due to Hopkins [8, 18, 83], who observed that the integration over the source can be carried out before summing up the diffraction angles accepted by the lens. Vector extension of the Hopkins's method was performed by Yeung [33] and partly reproduced here for completeness. Yeung's formulation also incorporated thin-film interference effects, but these are above the scope of this thesis and are ignored in the following paragraphs.

In order to be able to apply the Hopkins formalism to equation (4.2), it is necessary to separate the expression for the mask field into the product of two terms, one depending only on the reticle coordinates (x, y) and one depending only on the source point coordinates (p_x, p_y). The effects due to the reticle and due to the angle of incidence can not, in principle, be separated in rigorous solutions of Maxwell's equation on the entire mask topography. Typical printing situations encountered in optical lithography involve reduction factors or $4X$ or $5X$ and partial coherence factors of σ between 0.3 and 0.8 and, therefore incident angles no larger than 10^o with respect to the normal direction. Under such circumstances, several authors have verified independently, based on rigorous field evaluations, that the diffracted harmonics at the exit pupil of the reticle remain approximately constant [82, 84, 85, 86]. Furthermore, Pistor [85] analyzed the limits of this approximation for mask features of small dimensions and high aspect rations and showed that it holds for typical phase shifting masks in printing lithography. It breaks down, however, in inspection lithography situations where $1X$ reduction factors and, in consequence, incident angles of 50^o are used [84]. This result makes possible the following approximation of the object fields for most imaging systems used in optical lithography:

$$\mathbf{E_o}(x, y; p_x, p_y) \approx \mathbf{E_o}(x, y; p_x = 0, p_y = 0)e^{-ik(p_x x + p_y y)} \tag{4.3}$$

$$\mathbf{H_o}(x, y; p_x, p_y) \approx \mathbf{H_o}(x, y; p_x = 0, p_y = 0)e^{-ik(p_x x + p_y y)} \tag{4.4}$$

Equation (2.38) then transforms into:

$$\begin{aligned}&\mathbf{E}_{image}(x',y') = \\ &-\frac{M}{\lambda^2}\iint_{-\infty}^{\infty}\mathbf{K}(\frac{s_x}{\lambda},\frac{s_y}{\lambda};z')\ \cdot \boldsymbol{F}\Big\{\mathbf{E_O};\frac{Ms_x-p_x}{\lambda},\frac{Ms_y-p_y}{\lambda}\Big\}\ e^{-jk(s_xx'+s_yy')}\ ds_xds_y\end{aligned} \tag{4.5}$$

where $\mathbf{E_O}$ can be interpreted as the Fourier Transform integrand of either equation (2.38) or the rather more complex and accurate integrand in equation (2.15), which provide the fields at the entrance pupil of the imaging system. Moreover, according to equation (2.38),

$$\mathbf{K}(\frac{s_x}{\lambda},\frac{s_y}{\lambda};z') = \sqrt{\frac{cos\theta}{cos\theta'}}\ \mathbf{T}(s_x,s_y)e^{-jk\left[C+\Phi(s_x,s_y)-s_z\triangle_z\right]}\ CIRC\Big(\frac{\sqrt{s_x^2+s_y^2}}{NA_i}\Big) \tag{4.6}$$

with $\mathbf{T}$ representing the polarization tensor (2.33) derived in section 2.2, and $\triangle_z$ represents the defocus distance in the direction along the optical axis, $\hat{\mathbf{e}}_\mathbf{z}$, measured from the focal point at $z'=0$. In addition, the filtering effects of the projection system entrance pupil have been incorporated into the function $\mathbf{K}$ by means of the circular step function, $CIRC$.

Under the assumption of isoplanatic or space-invariant optical systems and after the application of Hopkins approximation, equation (4.2) can be transformed, with the help of the new set of integration variables $(f,g) = (\frac{Ms_x-p_x}{M\lambda},\frac{Ms_y-p_y}{M\lambda})$, $(f',g') = (\frac{Ms'_x-p_x}{M\lambda},\frac{Ms'_y-p_y}{M\lambda})$ and $(f'',g'') = (\frac{p_x}{M\lambda},\frac{p_y}{M\lambda})$, into [33]:

$$\begin{aligned}I_{image}(x',y',z') =&\iiiint\sum_{ij}TCC_{ij}(f,g;f',g';z') \\ &\boldsymbol{F}\Big\{M^2E_{O_i};Mf,Mg\Big\}\ \boldsymbol{F}^*\Big\{M^2E_{O_j};Mf',Mg'\Big\}e^{-j2\pi[(f-f')x'+(g-g')y']}\ dfdgdf'dg'\end{aligned} \tag{4.7}$$

where the indexes i,j,k refer to the cartesian components x,y,x of the fields.

Hence the aerial image intensity can be expressed as a weighted sum of the interference between every pair of waves at the entrance pupil with frequencies (f,g) and (f',g'), respectively [2]. The weighting factors in this relation are the

so-called Transmission Cross-Coefficients (TCC), which are independent of the object pattern and, according to Yeung's vector extension of Hopkins theory, are given in the following matrix form:

$$TCC_{ij}(f,g;f',g';z') = \iint I_{source}(f'',g'') \sum_{k} K_{ki}(f''+f,g''+g;z')K^{*}_{kj}(f''+f',g''+g';z')\, df''dg'' \tag{4.8}$$

This formulation is computationally more advantageous than Abbe's method because, given an optical system with fixed illumination, numerical aperture, defocus, and other aberrations, its transmission cross-coefficients (TCCs) need to be calculated only once. Both illumination and imaging processes are modeled by the TTCs. Thereafter they can be reused for repeated aerial image simulations of different mask patterns printed by the same optical system.

Several implementations of Hopkins' theory of partially coherent imaging are due to the members of the Advanced Lithography Group in the Electrical Engineering and Computer Science Department of the University of California, Berkeley [87, 88]. Specifically, SPLAT is a program written in FORTRAN 77 for two-dimensional projection-printing. Yeung's vector extension of the Hopkins's method was recently implemented into the new version 6.0 of SPLAT [89]. This simulator allows the possibility of input as object field the rigorously evaluated Thick Mask diffraction orders or those generated by the Thin Mask approximation.

Existing simulation tools with extended physical models and convenient user interfaces are not always commercially available. Some of the software most commonly used by the lithography community include PROLITH [38, 90] and SOLID-C [91] among others.

4.3 Boundary Layer Model with Partially Coherent Illumination

Two equivalent formulations of the imaging process with partial coherent illumination were briefly described in section 4.2. These are Abbe's and Hopkins methods. As an intermediate step, it was shown that Abbe's method yields the image field amplitude and phase distribution for each illumination angle prior to the source integration, while Hopkins yields directly the final intensity distribution. Information about the field amplitude and phase is needed to determine our boundary layer model and, therefore Abbe's method was chosen over Hopkins', despite the computational superiority of the later. Once the boundary layer parameters have been deduced, however, they can be used in combination with either imaging theory.

4.3.1 Simulation Practical Aspects

Hopkins Approximation

As explained in section 4.2, the so-called Hopkins approximation, assumes that, for small angles of incidence relative to the normal, the photomask diffraction orders are equal to those obtained at normal incidence, except for the corresponding shift in frequency. This approximation is necessary to derive Hopkins formalism, and can as well be used to greatly reduce the computational cost of evaluating the object field for each possible incident plane wave in Abbe's procedure. A single rigorous mask field evaluation with normal incidence provides reasonably accurate results for the diffraction orders due to plane waves propagating within the condenser numerical aperture. Hence partially coherent light can be simulated through Abbe's formalism by means of equations (4.3) and

(4.4), that is, by appropriately shifting in k-space the diffracted orders obtained at normal incidence. Then the image intensity results due to all possible incident plane waves are incoherently added to form the final aerial image.

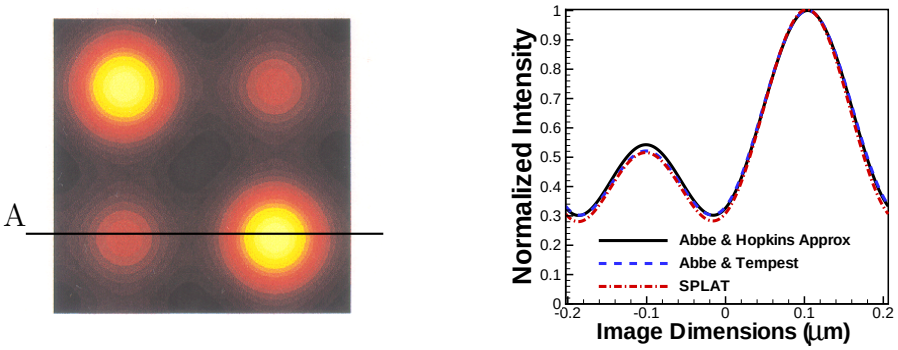
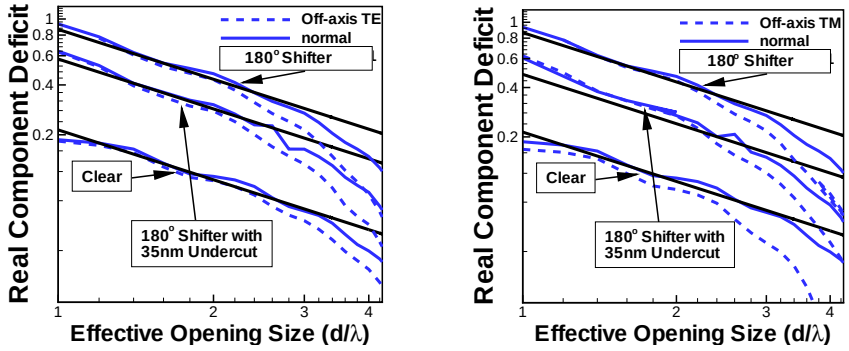


Figure 4.8: Aerial image intensity distribution cross section along plane A of a 2λ half-pitch array of alternating 180^o-shifter and clear square mask openings illuminated with 193nm partially coherent light of $\sigma = 0.55$. Partially coherent light was simulated through Abbe's formalism by incoherently adding image results due to plane waves at different angles of incidence. Results for mask fields rigorously evaluated with Tempest at different incident angles are compared to results obtained with the application of Hopkins approximation. Results are also compared against aerial images evaluated with the software SPLAT [89] based on Hopkins formulas.

In figure 4.8, aerial image results for mask fields rigorously evaluated with Tempest [20] at different angles of incidence are compared to results obtained through the application of Hopkins approximation. The mask pattern of figure 4.1(b) was illuminated with 193nm partially coherent light of $\sigma = 0.55$. The cross-sectional view along plane A on figure 4.8 shows a near overlap between both images. Results are also compared against aerial images evaluated with the software SPLAT [89] based on Hopkins formulas. The relative agreement justifies the application of Hopkins approximation with $4X$ imaging systems and NA up

to 0.85, even with phase-shifting mask, whenever the partial coherence factor does not exceed values equal to 0.8-0.9.



(a) (b)

Figure 4.9: Relative deficit of the aerial field real component follows a similar reciprocity relation with the opening effective size for normal or off-axis illumination. The angle of incidence corresponds to a value of $\sigma = 0.55$ for a $4X$ imaging system of $NA_i = 0.85$.(a) TE polarization. (b) TM polarization.

Utilization of the Boundary Layer Model to the simulation of partially coherent light requires the determination of the model parameters for plane waves with off-axis incidence. It is expected, however, that Hopkins approximation holds for the boundary layer width and transmission coefficient parameters, that is, that these parameters remain approximately constant for incident angles up to $8^o - 10^o$. This is shown in figures 4.9(a) and (b) for TE and TM polarizations, respectively, at normal incidence and with a 8.5^o angle relative to the normal. The same inverse relation of the relative deficit on the real field component with the opening effective size is observed for opening sizes up to about 3λ. Larger opening sizes exhibit smaller errors due to the Kirchhoff approximation with off-axis illumination. This result is a consequence of choosing the peak amplitude as a measure of the error and does not necessarily imply an improvement of the

thin mask accuracy at off-axis angles. Nevertheless, the use of the boundary layer parameters of table 3.1 for off-axis plane waves within the range $\sigma = 0.3 - 0.6$ for a $4X$ imaging systems is justified by the results of figures 4.9(a) and (b).

Source Discretization

Abbe's model of partial coherent illumination as described in section 4.2 requires the discretization of a continuous source into a finite number of source integration points, and a separate TEMPEST simulation needs to be performed for each grid point.

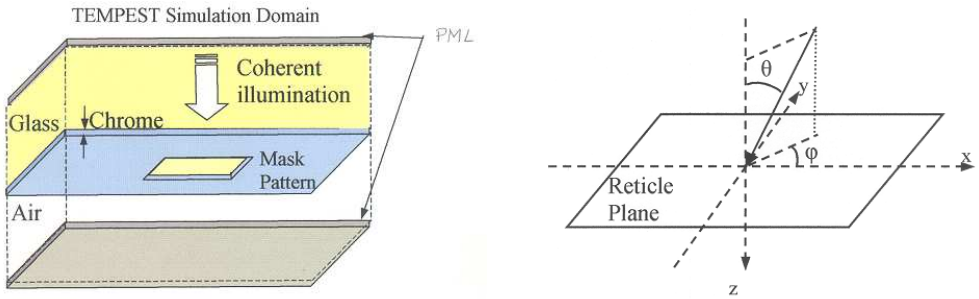


Figure 4.10: (a) Tempest simulation domain. (b) Plane wave incident angles.

Tempest simulations are performed in the domain as illustrated in figure 4.10 of dimensions $10\lambda \times 10\lambda \times 4\lambda$. Light propagates in the z-direction and Perfectly Matched Layers are placed on either end to emulate an infinite extension of the domain in the z-axis. Periodic boundary conditions are applied to the lateral directions, however, which restrict the allowed values of propagating wave vectors to a discrete set inside the Tempest domain. Given a simulation domain of dimensions $L_x \times L_y$ on the XY plane, any wave vector $\mathbf{k} = (k_x, k_y, k_z) = (kp_x, kp_y, k\sqrt{1 - p_x^2 - p_y^2})$, with $k = \frac{2\pi}{\lambda}$, can only take the

following values:

$$k_x = p\frac{2\pi}{L_x} \quad with \quad p = 0, \pm 1, \pm 2..$$
$$k_y = q\frac{2\pi}{L_y} \quad with \quad q = 0, \pm 1, \pm 2..$$

Assuming that $a = ML_x$ and $b = ML_y$ are the dimensions of the repeating cell at the wafer plane, and substituting $k = \frac{2\pi}{\lambda}$ in the previous relation, we have:

$$p_x = p\frac{M\lambda}{a} \quad with \quad p = 0, \pm 1, \pm 2..$$
$$p_y = q\frac{M\lambda}{b} \quad with \quad q = 0, \pm 1, \pm 2..$$

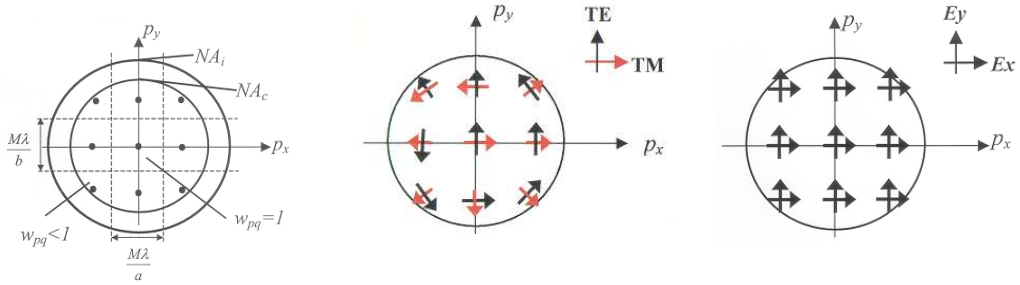
where (p,q) represent the integer indexes of each excitation plane wave. The angle θ formed by the incident wave vector with the optical axis z is given by: $\theta = arcsin(\frac{k_z}{k})$ with $k_z = \sqrt{k^2 - k_x^2 - k_y^2}$. The azimuthal angle φ formed by the projection of the incident wave vector on the reticle plane and the x-axis is defined as $\varphi = arctan(\frac{k_y}{k_x})$. As a result, only a discrete set of angles of incidence serve to represent a continuous source of light, and the integral in equation (4.2) transforms into a weighted sum of mutually incoherent images:

$$I_{image}(x', y', z') = \frac{1}{W} \sum_{p,q \subset \ \sigma} w_{pq} \ \mathbf{E_{image},}_{pq}(x', y') \cdot \mathbf{E^*_{image},}_{pq} \tag{4.9}$$

The intensity of one point source is made proportional to the discretization area w_{pq} as illustrated in figure 4.11(a), and the intensity is normalized to the total area W.

Tempest Polarization

Polarized waves are deterministic, the field quantities are definite functions of time and position and, therefore the field components add coherently. If the field quantities vary randomly, the field components are not correlated and add



(a) (b) (c)

Figure 4.11: (a) Source discretization diagram with periodic Tempest simulation domain. (b) Source polarization states TE and TM. (c) Approximated source polarization states E_x and E_y.

incoherently. Unpolarized light can thus be modeled by the incoherent superposition of two mutually orthogonal and normal to the propagation direction, linearly polarized plane waves [8]. In the situation encountered in lithography, where a plane wave impinges on a planar interface between two media, these two orthogonal polarizations, referred to as TE and TM, have the following polarization of the electric field:

$$\hat{\mathbf{e}}_{TE} = -sin\varphi\ \hat{\mathbf{e}}_x + cos\varphi\ \hat{\mathbf{e}}_y \tag{4.10}$$

$$\hat{\mathbf{e}}_{TM} = sin\theta cos\varphi\ \hat{\mathbf{e}}_x + sin\theta cos\varphi\ \hat{\mathbf{e}}_y - cos\theta\ \hat{\mathbf{e}}_z \tag{4.11}$$

based on the incident angles sketched in figure 4.10(b).

Polarization properties are assumed to be unaffected by the field propagation throughout the imaging system. Diffraction by the photomask, however, is greatly dependent on the polarization of the incident light and introduces coupling between orthogonal polarization as was illustrated in chapter 2. Thus two separate Tempest simulations are needed per source grid point, with incident TE and TM field polarizations, in the orientation indicated in figure 4.11(b). As Adam [92] pointed out, the application of Hopkins approximation means that

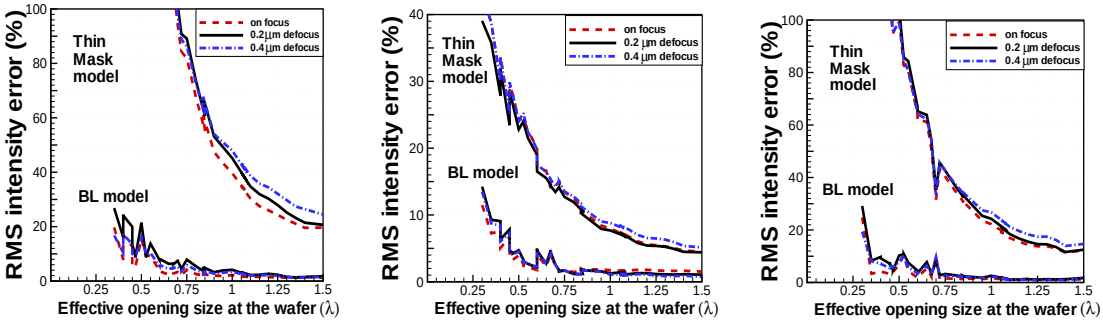
the two polarizations simulated at normal incidence, $\hat{\mathbf{e}}_{TE} \equiv \hat{\mathbf{e}}_y$ and $\hat{\mathbf{e}}_{TM} \equiv \hat{\mathbf{e}}_x$, are being shifted in k-space and used to deduce the scattering by off-axis point sources. This procedure is illustrated graphically in figure 4.11(c), where it becomes obvious that certain amount of inaccuracy is introduced as a result.

Variations of the incident angle θ affect uniquely to the TM polarization, but for values of $\theta \leq 10^o$, the z-component of the electric field varies less than 2% of the total field amplitude. Thus the effect of ignoring this component in our simulations is negligible. The amplitude variation of both field components with the azimuthal angle φ, however, is more significative. In fact, a source point of equal values of p_x and p_y has field components of equal strength along x- and y-axis with either TE or TM polarizations. When normal incident E_x and E_y polarizations are used in place of TE and TM, some error is inevitable. Nevertheless, for angles of incidence within the range of application of the Hopkins approximation, this error is assumed very small.

4.3.2 Isolated Features

Partial coherent illumination of the same isolated features that were analyzed in section 4.1 is studied in this section. The root mean squared (RMS) error in the intensity distribution was calculated for the aerial images generated by the modeled mask fields relative to the rigorous em fields. Figure 4.12(a) to (c) is a plot of the RMS error for the clear, 180^o-shifter and 180^o-shifter with 35nm undercut openings. Unpolarized (both TE and TM components were added incoherently), partially coherent illumination was used at $193nm$, $NA = 0.85$ and σ=0.6. A $4X$ reduction ratio was assumed in the simulations, such that the angle described by the waves collected at the entrance pupil is small and the same boundary layer parameters are applied to on-axis and off-axis illumination

components. Phase errors manifest themselves more critically when the image is out of focus [2] and, therefore need to be appropriately compensated by the Boundary Layer model of the object field. In order to evaluate phase effects on the model, RMS errors are evaluated at the focal plane and at two out-of-focus planes in figure 4.12. No significative variation of the error is observed for defocus amounts up to $0.4\mu m$, proving the accuracy of the Boundary Layer model within typical values of Depth-of-Focus.



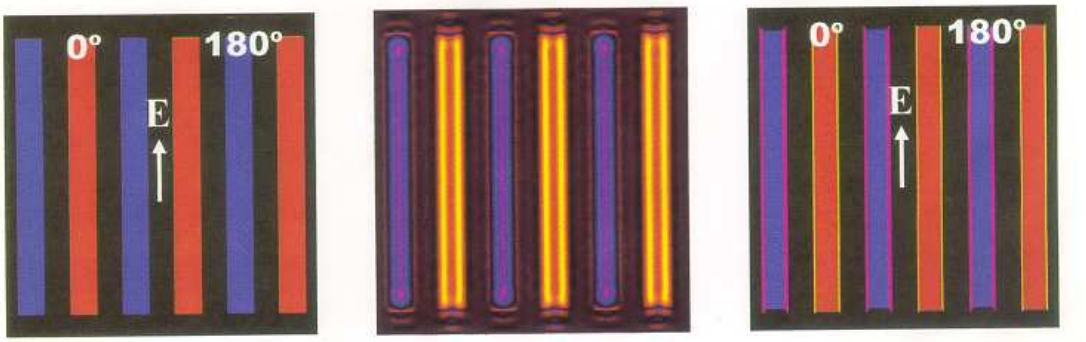
(a) (b) (c)

Figure 4.12: Root mean squared error in the intensity distribution, integrated over the focal plane and at two out-of-focus positions, of the approximated images relative to the rigorous em fields for unpolarized, partially coherent illumination at $193nm$, $NA = 0.85$, $\sigma = 0.6$. Three model parameters are analyzed: (a) 180^o-phase-shift openings, (b) clear openings and (c) 180^o-phase-shift with $35nm$ undercut openings.

4.3.3 Dense Patterns

Isolated features were simulated in Tempest by defining relative large domains. Figure 4.2 showed an example of small square features closely arranged in a periodic pattern. The accuracy of the Boundary Layer model was judged remarkable in that case, what suggested that the model capabilities extended as well to dense

mask patterns. In order to provide a more exhaustive analysis of the modelling accuracy, our boundary layer model has been applied to periodic mask features such as those of figure 4.13. Figure 4.13(a) represents the Kirchhoff scalar approximation (thin mask model) of a 78nm (as measured at the wafer plane) half-pitch array of alternating 180^o-shifter and clear lines. Unpolarized, $\sigma = 0.5$ partially coherent illumination at 193nm was used with a $NA = 0.85$ imaging system. Figure 4.13(b) shows a sketch of the actual object field obtained by rigorous electromagnetic FDTD Tempest simulation on the mask plane of the same mask pattern and vertically polarized electric field. Finally, figure 4.13(c) shows the Boundary Layer model for the same mask features and illumination conditions.



(a) (b) (c)

Figure 4.13: (a) Kirchhoff scalar approximation (thin mask model) of the field on the mask plane of a 78nm (as measured at the wafer plane) half-pitch array of alternating 180^o-shifter and clear line openings, with vertically polarized electric field at 193nm. (b) Sketch of the actual object field obtained by rigorous electromagnetic FDTD Tempest simulation on the same mask. (c) Boundary Layer model for the same mask features and illumination conditions.

Figures 4.14(a)-(c), on the other hand, correspond to the aerial image intensity at the focal plane for the rigorously evaluated object field 4.13(b) and its

two modeling attempts. Unpolarized, $\sigma = 0.5$ partially coherent illumination at $193nm$ was used with a $NA = 0.85$ imaging system. Figures 4.14(b), produced by the rigorously evaluated object field, and 4.14(c), corresponding to the Boundary Layer model, show a close agreement with only a 4.63% rms error due to the approximation. However, Figure 4.14(a), obtained from the Thin Mask approximation, produced a much higher RMS error of 50.97% as compared to the actual image.

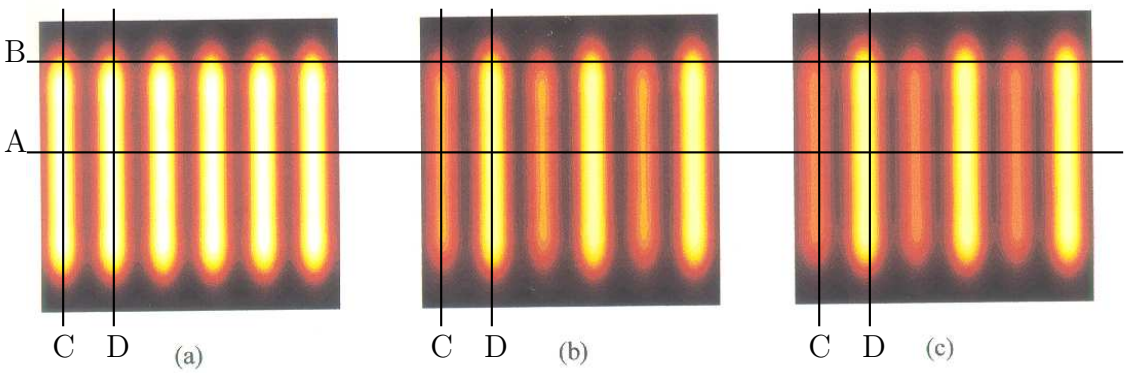
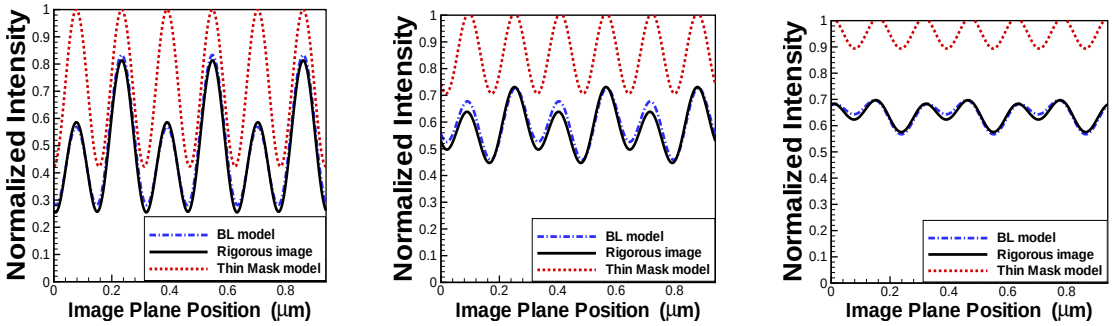


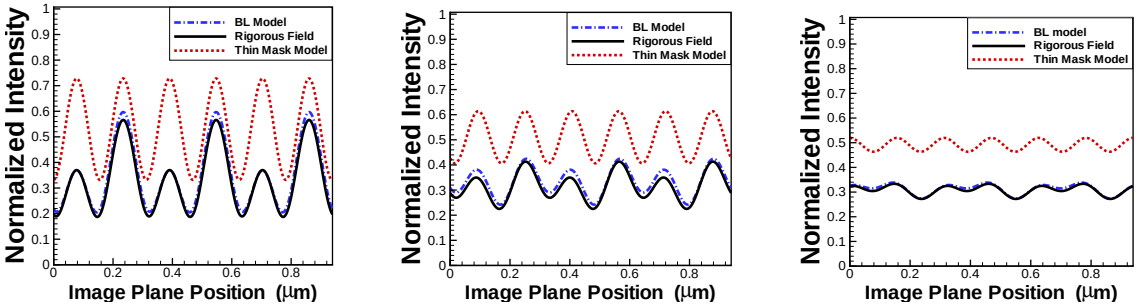
Figure 4.14: Aerial image intensity results at the focal plane of a 78nm (as measured at the wafer plane) half-pitch array of alternating 180^o-shifter and clear mask lines with an unpolarized, $\sigma = 0.5$ partially coherent illumination at 193nm and $NA = 0.85$. (a) Aerial image produced by the thin mask approximation with an rms error of 50.97%. (b) Aerial image produced by the rigorously evaluated object field, and (c) aerial image produced by the Boundary Layer with an rms error of 4.63%.

The cross-sectional planes marked in figure 4.14 by the lines A, B, C and D are plotted in figures 4.15 to 4.18, respectively. These views are plotted for the three aerial images at the focal plane as well as at $0.2\mu m$ and 0.4 μm defocus. All of the plots reveal a close overlap between the rigorous solutions and those obtained through the boundary layer model, even at points close to the end of



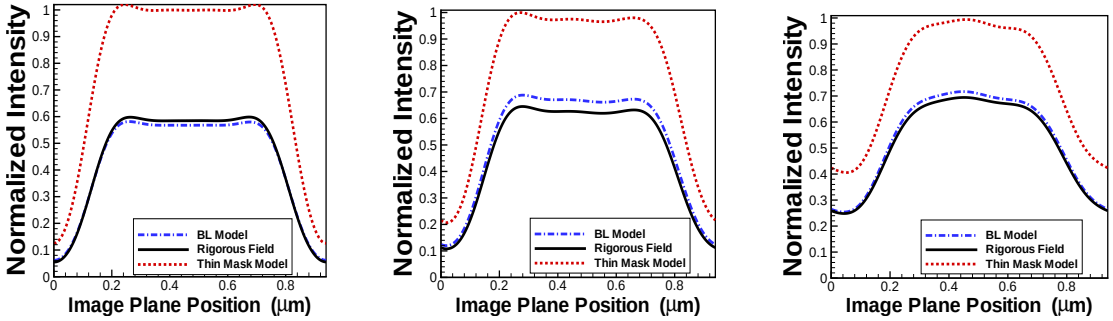
(a) (b) (c)

Figure 4.15: Cross-sectional view along plane marked by A of the aerial image intensity distribution of a 78nm (as measured at the wafer plane) half-pitch array of alternating 180^o-shifter and clear lines. Unpolarized, $\sigma = 0.5$ partially coherent illumination at $193nm$ and $NA = 0.85$ was used. (a) Image at the focal plane, (b) effect of defocus at $0.2\mu m$ out-of-focus and (c) at $0.4\mu m$ out-of-focus.



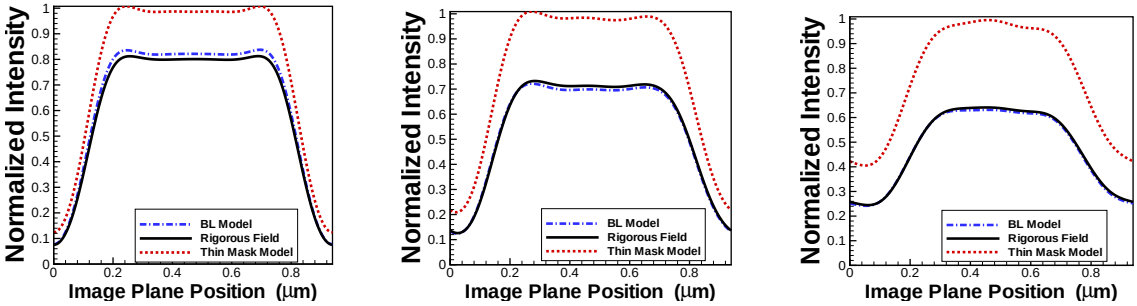
(a) (b) (c)

Figure 4.16: Cross-sectional view along plane B towards the end of the lines. Same parameters. (a) Image at the focal plane, (b) effect of defocus at $0.2\mu m$ out-of-focus and (c) at $0.4\mu m$ out-of-focus.



(a) (b) (c)

Figure 4.17: Cross-sectional view along plane C, corresponding to the 180^o-shifter opening. Same parameters. (a) Image at the focal plane, (b) effect of defocus at $0.2\mu m$ out-of-focus and (c) at $0.4\mu m$ out-of-focus.



(a) (b) (c)

Figure 4.18: Cross-sectional view along D, corresponding to the clear opening. Same parameters. (a) Image at the focal plane, (b) effect of defocus at $0.2\mu m$ out-of-focus and (c) at $0.4\mu m$ out-of-focus.

the lines and with high amounts of defocus.

This improved modeling accuracy appears to be independent of the pitch value according to plot in figure 4.19 of the root mean squared error measured

for different pitches at the wafer plane.

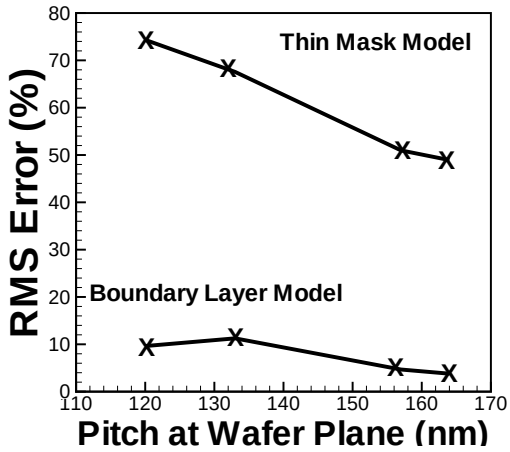


Figure 4.19: Root Mean Squared Error integrated over the wafer plane of periodic lines with different pitch dimensions, modelled by either the Thin Mask or the Boundary Layer Models.

As a final example of our model accuracy on square corners, the mask pattern of figure 4.20 is analyzed. As before, figure 4.20(a) depicts the Kirchhoff scalar approximation of the field on the mask plane of a 79nm (as measured at the wafer plane) half-pitch array of alternating 180^o-shifter and clear lines, which form a 90^o corner. Figure 4.20(b), on the other hand, illustrates the actual object field with vertically polarized electric field at 193nm, obtained by rigorous electromagnetic FDTD Tempest simulation on the same mask. And finally, figure 4.20(c) represents the Boundary Layer model for the same mask features and illumination conditions.

Aerial image intensity results at the focal plane for the mask pattern of figures 4.21(a)-(c) were calculated for unpolarized light and an imaging system of $\sigma = 0.4$ partially coherent illumination at 193nm and $NA = 0.85$. The enhanced modeling accuracy due to the boundary layer can be appreciated directly from

(a) (b) (c)

Figure 4.20: (a) Kirchhoff scalar approximation (thin mask model) of the field on the mask plane of a 79nm (as measured at the wafer plane) half-pitch array of alternating 180^o-shifter and clear corner openings, with vertically polarized electric field at 193nm. (b) Sketch of the actual object field obtained by rigorous electromagnetic FDTD Tempest simulation on the same mask. (c) Boundary Layer model for the same mask features and illumination conditions.

the plots. Quantitatively, the aerial image produced by the thin mask approximation exhibited a RMS error of 56.78%, while that produced by the Boundary Layer showed an error of 3.28%.

4.4 Sensitivity of the BL model Parameters with the Chrome Thickness

In section 3.4 it was explained that a chrome layer thickness of 95nm was employed at 193nm wavelength such that it provided the same absorption of incident light as the 80nm chrome layer used at 248nm wavelength. This section focuses of the variation of the boundary layer parameters at 193nm lithography when the chrome thickness changes from 95nm to 80nm. It can be seen in figure 4.22

(a) (b) (c)

Figure 4.21: Aerial image intensity results at the focal plane of a 79nm (as measured at the wafer plane) half-pitch array of alternating 180^o-shifter and clear square corners with an unpolarized, $\sigma = 0.4$ partially coherent illumination at 193nm and $NA = 0.85$. (a) Aerial image produced by the thin mask approximation with an rms error of 56.78%. (b) Aerial image produced by the rigorously evaluated object field, and (c) aerial image produced by the Boundary Layer with an rms error of 3.28%.

that the same "Inverse Law" of the field real component deficit with respect to the effective opening size can be identified for chrome layers of 80nm thickness. Moreover, the same exact coefficients would be extracted for the 180^o-shifter and 180^o-shifter with 35nm undercut, for both 95nm and 80nm cases. The higher transmission through the chrome, however, has a larger impact on the error measured for clear openings. The curve for clear openings deviates from the straight inverse line, and provides different values for the model parameters. This result is consistent with the fact that the boundary layer model is founded on the Physical Theory of Diffraction, which was developed for perfect conducting surfaces. Errors due to the thin mask approximation are expected to fit the "inverse law" better for more opaque chrome layers.

Nevertheless, although the boundary layer coefficients do not vary abruptly

with changes in the mask transversal profile, it is suggested that a new set of parameters is calculated each time the mask profile is modified.

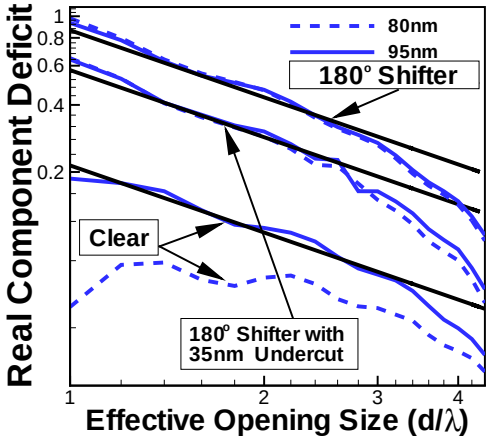
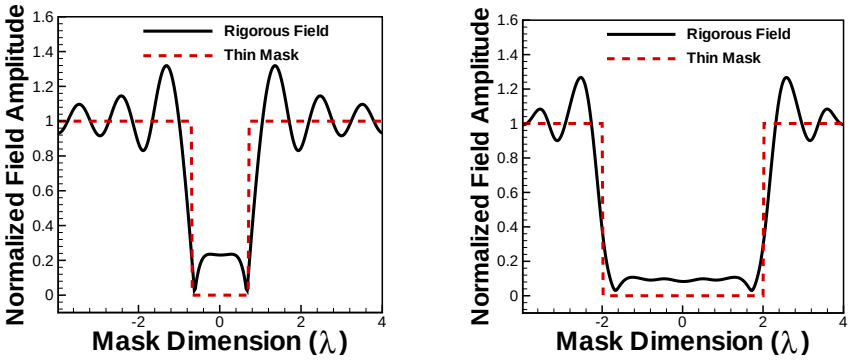


Figure 4.22: Relative deficit of the real component of the electric field as measured at the peak of the aerial image on the focal plane.

4.5 Opaque Mask Features for Negative Resist

The type of mask patterns studied in most of this thesis are those utilized for printing on positive photoresist. Positive resists become soluble in developer solution upon exposure of light, while negative resists lose their solubility on those areas exposed to light. This means that in order to produce contact holes on positive resists, the mask feature consists of an aperture on the chrome layer. On the other hand, in order to produce a contact hole on a negative resist, one needs a mask consisting of an opaque chrome feature surrounded by glass. Due to diffraction at the chrome edges and the small dimensions of the mask feature needed for the creation of contact holes with negative resist, unwanted light can reach the center of the contact, leading to closed holes. Imaging over positive

resists also suffers from reduced field transmission through the chrome aperture, but its effect over the aerial image is less critical and, therefore preferred in practice to print contact holes [7].



(a) (b)

Figure 4.23: Mask Fields produced by a square opaque features of different sizes: (a) 1.4λ, (b) 3λ. The chrome shapes block the transmission of light.

Diffraction effects, as well as the nonzero transmission through the chrome layer, suggest the possibility that the boundary layer model might not be able to model the electromagnetic effects on opaque chrome features of small dimensions. Figures 4.23(a) and (b) are two examples of mask fields due to two opaque mask square features of sizes 1.4λ and 3λ, respectively. They illustrate the nonzero field amplitude inside the area covered by chrome, where it was expected that the light would be blocked.

A similar evaluation of the peak amplitude error of the aerial image due to Kirchhoff approximation was carried out for opaque square features surrounded by glass. No phase-shift was introduced. The relative error on both real and imaginary field components of the aerial image are plotted in figure 4.24. Errors on the imaginary component fit the reciprocal law on the opening size more closely than the errors on the real component, which deviates slightly for sizes

larger than 3λ. Nonetheless, these results suggest the possibility of defining a new model for opaque features with dimensions of the order of the wavelength, consisting of a boundary layer to be applied at the edges on the clear areas of the mask.

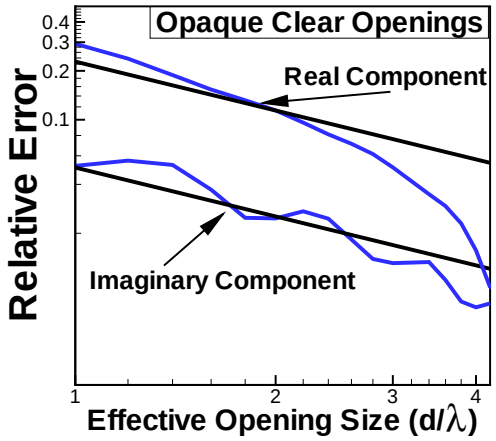


Figure 4.24: Relative error on both real and imaginary field components of the aerial image.

4.6 Discussion

Next generation Alternating Phase-Shifting masks have etching profiles with abrupt discontinuities, mask features of the order of the wavelength and high aspect ratios. Rigorously evaluated fields on thick chrome mask apertures exhibit diffraction and polarization effects (Thick Mask effects) that can not be accurately modeled by the conventional Thin Mask approximation.

The key result of our simulations is that the thick mask effects can be interpreted, to a good approximation, as an intrinsic edge property, and modeled with just two fixed parameters: width and transmission coefficient of a locally-determined

boundary layer. Thus our proposed model consists of a sophisticated version of Kirchhoff approximation, simply adding a boundary layer to every edge.

We proved in the previous paragraphs that the BL model accurately accounts for thick mask effects of the fields on the mask, incorporating effects of electromagnetic coupling due to the high numerical aperture ≥ 0.7, and accurately compensates for phase errors even at planes out of focus. This greatly improves the accuracy of aerial image computation in photolithography simulations at a reasonable computational cost.

APPENDIX A

Scalar Diffraction Theory

Deviation from rectilinear propagation of light, the accepted light propagation theory in the 17^{th} century, was termed *diffraction* by Sommerfeld [70] but first observed and reported by Grimaldi in 1665. It was Huygens, however, who first proposed a wave theory in 1678. Huygens postulated that "each element of a wave-front may be regarded as the center of a secondary disturbance which gives rise to spherical, diverging wavelets, and the envelope of these secondary waves forms the wavefront at later times" [8]. Huygens seemed to have been unaware of Grimaldi's observations and the wave theory was not used to explain the phenomena of diffraction until the year 1818, when Fresnel supplemented Huygens' envelope construction with Young's interference principle to describe these effects. Years later, in 1882, the *Huygens-Fresnel Principle* was given mathematical foundations by Kirchhoff, in the form of an integral theorem expressing the solution of the homogeneous wave equation at any arbitrary point in the field, in terms of the values of the solution and its first normal derivatives on an arbitrary closed surface enclosing the observation point. This theorem constitutes the basis of the *Scalar Diffraction Theory*.

A.1 Kirchhoff Diffraction by a Planar Screen

Scalar diffraction theory is based on the scalar Green's theorem (A.1) on a closed surface S enclosing a volume V, applied to two complex functions, U and G, continuous and with continuous derivatives inside and on S.

$$\iiint_V (U\nabla^2 G - G\nabla^2 U)dv = \iint_S \Big(U\frac{\partial G}{\partial n} - G\frac{\partial U}{\partial n}\Big)ds, \tag{A.1}$$

where $\frac{\partial}{\partial n}$ indicates the partial derivative in the outward normal direction at each point on S.

Scalar diffraction theory treats light as a scalar phenomenon, ignoring the coupling between the various electromagnetic components through Maxwell equations at boundaries, lenses or other optical elements. Thus if we assume a monochromatic wave, $U(\mathbf{r},t) = U(\mathbf{r})e^{i\omega t}$, with U representing any of the six electromagnetic field components, then $U(\mathbf{r})$ must obey the time-independent wave equation: $(\nabla^2 + k^2)U(\mathbf{r}) = 0$. Finally, choosing the free-space green's function, solution of the equation $(\nabla^2 + k^2)G(\mathbf{r},\mathbf{r}') = -\delta(\mathbf{r}-\mathbf{r}')$, as the auxiliary function $G(\mathbf{r},\mathbf{r}') = \frac{e^{-jk|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|}$, equation (A.1) becomes:

$$U(\mathbf{r}) = \iint_{S'} \Big(U(\mathbf{r}')\frac{\partial G(\mathbf{r},\mathbf{r}')}{\partial n'} - G(\mathbf{r},\mathbf{r}')\frac{\partial U(\mathbf{r}')}{\partial n'}\Big)ds', \tag{A.2}$$

providing the value of the field U at an observation point $\mathbf{r}$ in terms of its values on the boundary surface S. Note that the integration is performed on the primed variables and the unit vector $\hat{\mathbf{n}}$ has been exchanged by $\hat{\mathbf{n}}' = -\hat{\mathbf{n}}$, the inward normal vector as sketched in figure A.1.

For the diffraction problem by a plane screen like the one sketched on figure A.1, the integration surface is segmented into three disjoint parts. A planar surface S_1 laying directly on the opaque region of the screen exit surface, the surface A covering the area of the diffracting aperture on the screen, and a large half sphere S_2 containing the observation point $\mathbf{r}$. The entire integral over S_2 must

vanish as $R = |\mathbf{r}|$ becomes arbitrarily large, provided the disturbance satisfies the Sommerfeld radiation condition [70]:

$$\lim_{r\to\infty} \; R\big(\frac{\partial U}{\partial n} - jkU\big) = 0 \tag{A.3}$$

which is true if $U(\mathbf{r})$ vanishes at least as fast as a spherical wave. Since the illuminating light $U_s(\mathbf{r})$ invariably consists of a linear combination of spherical waves, this requirement is always fulfilled.

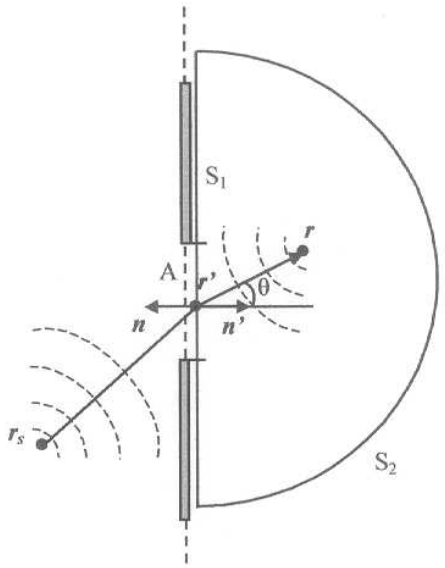


Figure A.1: Schematic diagram for the application of Kirchhoff diffraction theory to a plane screen.

A.2 Kirchhoff Boundary Conditions

Application of Kirchhoff formula requires knowledge of both U and $\frac{\partial U}{\partial n}$ on the surface S, what can become impracticable for many complicate shaped scatterers. Approximate estimates of the boundary values of the field are of common practice and often accurate enough. The most common approximation is the so-called

Physical-Optics approximation, which was applied by Kirchhoff to the diffraction by apertures in plane screens and is usually referred to as Kirchhoff's Boundary Conditions. Kirchhoff found the solution to the diffraction problem by specifying the following boundary conditions [41]:

1. Across the aperture A, the field distribution $U(\mathbf{r}')$ and its normal derivative $\frac{\partial U(\mathbf{r}')}{\partial n'}$ are exactly as they would be in the absence of the screen:

$$U(\mathbf{r}') = U_s(\mathbf{r}') \qquad \frac{\partial U}{\partial n'} = \frac{\partial U_s}{\partial n'} \qquad \text{for } \mathbf{r}' \in \text{A} \tag{A.4}$$

2. Over the portion of S_1 that lies in the geometrical shadow of the screen, the field distribution $U(\mathbf{r}')$ and its normal derivative $\frac{\partial U(\mathbf{r}')}{n'}$ are identically zero:

$$U(\mathbf{r}') = 0 \qquad \frac{\partial U}{\partial n'} = 0 \qquad \text{for } \mathbf{r}' \in S_1 \tag{A.5}$$

A.3 Rayleigh-Sommerfeld Diffraction Formulae

Kirchhoff's theory suffers from two inconsistencies. From the mathematical viewpoint, the boundary values are over-specified, which would not affect the result as long as they were the correct ones, but can produce errors when the boundary conditions (A.4) and (A.5) are applied. Kirchhoff's boundary conditions force both the potential function U and its normal derivative $\frac{\partial U}{\partial n'}$ to vanish on the surface S_1. Sommerfeld [70] used Riemann's theory of functions to argue that any solution of the wave equation vanishes in the whole space if both the solution and its normal derivative are zero on any finite boundary surface element. Then, according to this theorem, U and $\frac{\partial U}{\partial n'}$ must vanish over both S and A and, therefore equation (A.2) combined with Kirchhoff's boundary condition (A.5) imply that the field must be zero in all space. Moreover, Kirchhoff's integral does not reproduce the enforced boundary conditions as the observation

point approaches the aperture A. This inconsistencies in Kirchhoff theory were removed by Sommerfeld, who eliminated the need of imposing boundary values on both the disturbance and its derivative simultaneously. In the first Rayleigh-Sommerfeld solutions, $G = 0$ on the entire surface $A \cup S_1$, thus we only need to apply boundary conditions to U and not to its derivative. Similarly, by imposing $\frac{\partial G}{\partial n'} = 0$ on the entire screen $S_1 \cup A$, the second Raylegh-Sommerfeld solution is obtained, which only needs to apply boundary conditions over $\frac{\partial U}{\partial n'}$. The first Raylegh-Sommerfeld form of Kirchhoff's diffraction formula is as follows:

$$U_I(\mathbf{r}) = 2 \iint_{A'} U(\mathbf{r}') \frac{\partial G(\mathbf{r},\mathbf{r'})}{\partial n'} \, ds', \tag{A.6}$$

while the second reads:

$$U_{II}(\mathbf{r}) = -2 \iint_{A'} G(\mathbf{r}, \mathbf{r}') \frac{\partial U(\mathbf{r}')}{\partial n'} \, ds', \tag{A.7}$$

And the initial Kirchhoff's formula is recovered by averaging (A.6) and (A.7):

$$U_K = \frac{1}{2}(U_I + U_{II}), \tag{A.8}$$

All three formulae yield the same result for the diffracted fields, provided the boundary values are correctly specified. Any of the two Rayleigh-Sommerfeld solutions has the advantage that it requires knowledge of only one of the two field quantities, but they are applicable only to planar surfaces. The reason is that only over planar surfaces we can find, by means of image theory, Green's functions G or $\frac{\partial G}{\partial n}$ with the desired property of being zero on the entire surface $S_1 \cup A$. When the boundary values are incorrectly specified, as when Kirchhoff boundary conditions are employed, then they usually yield different results [42] at close distances from the aperture. Attempts to resolve or interpret Kirchhoff formulation in a consistent manner have been carried out by several authors [42, 93]. However, the most important point is to determine which formulation provides the most

accurate results in the far field region when small dimensions of the aperture are considered.

Wolf and Marchand [94] showed that when the linear dimensions of the aperture are large compared to the wavelength ($ka >> 1$) and the field are evaluated in the Fraunhofer or far field region of the aperture at not too large diffraction angles, then the predictions on the basis of the inconsistent Kirchhoff theory and on the basis of either one of the two consistent Rayleigh-Sommerfeld theories differ from each other by inappreciable amounts. Physically, the differences between formulations are due to the distinct expressions for the fringing effects at the edges, which can be neglected provided the dimensions of the aperture are large compared to the wavelength. The work by Totzeck [95] deals with the diffraction near fields of small objects with dimensions comparable to the wavelength. He compared all scalar Kirchhoff and Rayleigh-Sommerfeld diffraction theories with microwave measurements, and concluded that agreement between calculations and measurements was depended upon the size of the phase objects. His results showed that U_{II} yielded the closest results for scattering objects of dimensions much smaller than the wavelenght ($d << \lambda$), while U_I was preferable if the object thickness was approximately λ. The Kirchhoff diffraction integral, defined as the mean value of the Rayleigh-Sommerfeld diffraction integrals (A.8), yielded the best agreement for object thicknesses of approximately $\frac{\lambda}{2}$. Nevertheless, Totkzeck observed that the deviations between the three theories become negligible as the distance to the object increases only over a few wavelengths ($> 2\lambda$).

Based on the previous arguments, either one of the Rayleigh-Sommerfeld of Kirchhoff integrals will provide accurate results of the fields diffracted by a photomask at the entrance pupil of the imaging system, located several wavelengths apart. For simplicity, the first Rayleigh-Sommerfeld formula will be adopted.

Assuming the observation point located at a distance several wavelengths from the screen, that is, $|\mathbf{r}-\mathbf{r}'| >> \lambda$, then we can approximate $|\mathbf{r}-\mathbf{r}'|$ by its binomial expansion, $r-(\mathbf{r}\cdot\mathbf{r}')/r$, and replace:

$$\frac{\partial G(\mathbf{r},\mathbf{r}')}{\partial n'} = (jk+\frac{1}{r})\frac{e^{-jk|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|}(\hat{\mathbf{n}}'\cdot\hat{\mathbf{r}}) \approx jk\frac{e^{-jkr}}{4\pi r}\, e^{jk\hat{\mathbf{r}}\cdot\mathbf{r}'} cos\theta \tag{A.9}$$

such that the final scalar diffraction formula of the field amplitude on the entrance pupil simplifies to:

$$U_{Entrance\ Pupil}(\mathbf{r}) = \frac{j}{\lambda}\frac{e^{-jkr}}{r}\iint_{A'} cos\theta\, U(\mathbf{r}')\, e^{jk\hat{\mathbf{r}}\cdot\mathbf{r}'}\, ds' \tag{A.10}$$

An additional approximation can be imposed to this already approximate theory that will be applicable only to numerical apertures up to about 0.4 [31]. In the regime of small NA the maximum angle θ that the diffracted light ray $\hat{\mathbf{r}}$ forms with the optical axis $\hat{\mathbf{e}}_\mathbf{z} = \hat{\mathbf{n}}'$ is also very small (of the order of $sin^{-1}(\frac{NA}{4})$ for an imaging system of $\frac{1}{4}$ demagnification), and the paraxial or small angle approximation can be made, that is, $cos\,\theta \approx 1$. For higher values of NA, about $0.6-0.7$, the nonparaxial scalar diffraction theory can still provide adequate predictions of the diffracted fields with decreasing accuracy as NA increases. For $NA > 0.7$, polarization effects and the inherent coupling between the vector components of the em field became noticeable and vector diffraction theory is necessary [31, 32].

Only image points of small linear dimensions around the optical axis, $\hat{\mathbf{e}}_z = \hat{\mathbf{n}}'$, as compared to the distance r, are of interest in this analysis, what means that, according to the method of stationary phase [8], only points about the optical axis will contribute significantly to the diffraction integral. Therefore the origin of the unit vectors $\hat{\mathbf{r}}$, pointing towards the observation point $\mathbf{r}$ on the entrance pupil, can be taken as the center of coordinates on the object plane, and it can

be expressed as:

$$\hat{\mathbf{r}} = \frac{\mathbf{r}}{r} = r_x\,\hat{\mathbf{e}}_x + r_y\,\hat{\mathbf{e}}_y + r_z\,\hat{\mathbf{e}}_z = sin\theta cos\varphi\,\hat{\mathbf{e}}_x + sin\theta sin\varphi\,\hat{\mathbf{e}}_y + cos\theta\,\hat{\mathbf{e}}_z \tag{A.11}$$

Equation (A.9) can then be interpreted as the Fourier transform of the field amplitude $U(\mathbf{r}')$ evaluated at the spatial frequencies $(\frac{r_x}{\lambda}, \frac{r_y}{\lambda})$.

A.4 Focusing of Scalar Waves

The phase transformation induced in the propagating wave by a lens composed of spherical surfaces has the property of mapping incident waves into spherical waves, which converge towards the focal point if the lens is a converging one. Until 1909, most theoretical treatments of wave focusing were based on Huygens-Fresnel principle which utilizes spherical-wave representations of the fields. In 1909 Debye reformulated the focusing problem using plane waves rather than spherical waves. Debye's original formulation can be derived in a simplified manner starting from the far field representation of the first Rayleigh-Sommerfeld integral (A.10). Each point of the exit pupil is assumed to lie on the spherical wavefront of a monochromatic wave converging towards the focal point located, for simplicity, on the optical axis. According to Kirchhoff Boundary Conditions applied on the exit pupil surface of linear dimensions assumed large compared to λ, the scalar field amplitude emerging from the exit pupil can be expressed as:

$$U_{ExitPupil}(\mathbf{r}') = A(x,y)e^{-jkC(x,y)}\frac{e^{jkR'}}{R'} \tag{A.12}$$

where A and C are, respectively, the amplitude and phase of the wave propagating through the lens as it reaches each point (x, y) of the exit pupil. Hence, based on figure A.2, the field at any image point $\mathbf{r}'$ will consist of the contribution from all spherical waves $\frac{e^{-jkL}}{L}$, originating at each point of the exit pupil, each of them having the strength given by equation (A.12):

$$U_{image}(x', y', z') = \frac{j}{\lambda} \iint_{ExitPupil} cos\theta' A(x, y) \frac{e^{-jk[C-R']}}{R'} \frac{e^{-jkL}}{L} da, \tag{A.13}$$

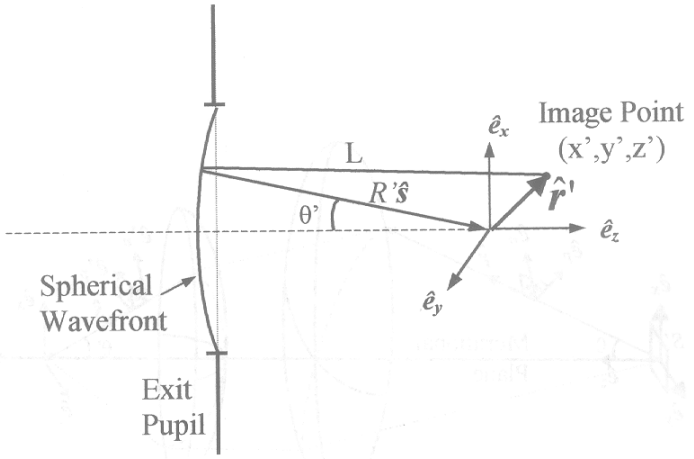


Figure A.2: Notation for Debye's integral formulation. Converging spherical wave to the gaussian image point on the optical axis from a circular aperture.

From figure A.2, the relation $L - R' = \hat{\mathbf{s}} \cdot \mathbf{r}'$ can be extracted, where

$$\hat{\mathbf{s}} = s_x\, \hat{\mathbf{e}}_x + s_y\, \hat{\mathbf{e}}_y + s_z\, \hat{\mathbf{e}}_z = -sin\theta' cos\varphi\, \hat{\mathbf{e}}_x - sin\theta' sin\varphi\, \hat{\mathbf{e}}_y + cos\theta'\, \hat{\mathbf{e}}_z \tag{A.14}$$

and $s_x = \frac{x}{MR'}$ and $s_y = \frac{y}{MR'}$, where M denotes the demagnification of the lens (usually $\frac{1}{4}$ or $\frac{1}{5}$). Hence using $da' = R'^2 d\Omega = R'^2 \frac{ds_x ds y}{s_z}$, with $R' \sim L$, we obtain the final expression for the fields in the neighborhood of the focal point as given by equation (A.15).

$$U_{image}(x', y', z') = \frac{j}{\lambda} \iint_{{s_x}^2 + {s_y}^2 \leq NA^2} \frac{A(s_x, s_y)}{s_z}\, e^{-jk\left[C + \hat{\mathbf{s}} \cdot \mathbf{r}'\right]}\, ds_x ds_y, \tag{A.15}$$

This formulation is based in geometrical optics and provides the optical image in the neighborhood of the lens focal plane in terms of a superposition of plane

waves propagating with all directions cosines, (s_x, s_y), that fill the exit pupil aperture [8].

Direct transformation of the first Rayleigh-Sommerfeld integral formula (A.6) into an equivalent expansion in terms of plane waves can be carried out through the asymptotic evaluation of the angular-spectrum of the exit pupil field (A.12) [42]. In Debye's approximation, only the contribution from the stationary points inside the pupil are retained, which represent plane waves with all possible propagation vectors (s_x, s_y) that lay inside the solid angle Ω subtended by the aperture at the focal point. This procedure was followed by Wolf in his vector generalization of Debye's representation of the electromagnetic fields on the image space [43], and justifies the use of geometrical optics to trace the polarization state of the field along each ray. Wolf's generalization of Debye's integral is described in chapter 2.

Assuming negligible losses due to reflection and absorption through the lens, the field amplitudes at the entrance and exit pupil must satisfy the energy conservation relation:

$$|U_{ExitPupil}(s_x, s_y)|^2 \ da' = |U_{EntrancePupil}(r_x, r_y)|^2 \ da \tag{A.16}$$

Based on the notation of figures 2.1 and A.2, $da = r^2 sin\theta d\theta d\varphi$ and $da' = R'^2 sin\theta' d\theta' d\varphi'$. Further, the angle θ between the incident ray and the entrance pupil, and the angle θ' between the outgoing ray and the exit pupil, must satisfy the Abbe's sine condition [8], $sin(\theta) = Msin(\theta')$ with M being the demagnification of the lens ($\frac{1}{4}$ or $\frac{1}{5}$), such that the spatial frequencies in the image space are related to those in the object space according to $s_x = -\frac{r_x}{M}$ and $s_y = -\frac{r_y}{M}$. Assuming an index of refraction equal to unity on both object and image spaces and noting that $d\varphi = d\varphi'$, these two conditions result in the following obliquity factor for the field magnitude at the exit pupil:

$$R'U_{ExitPupil}(s_x, s_y) = rU_{EntrancePupil}(r_x, r_y) \ \ M \ \sqrt{\frac{cos\theta'}{cos\theta}} \tag{A.17}$$

and the final expression for the image field is as follows:

$$U_{image}(x', y', z') =$$

$$-\frac{M}{\lambda^2} \iint\limits_{{s_x}^2+{s_y}^2 \leq NA^2} \sqrt{\frac{cos\theta}{cos\theta'}}\ \boldsymbol{F}\{U(\mathbf{r}'); \frac{Ms_x}{\lambda}, \frac{Ms_y}{\lambda}\} e^{-jk\left[C+\hat{\mathbf{s}}\cdot\mathbf{r}'\right]}\ ds_x ds_y \tag{A.18}$$

APPENDIX B

Vector Green's Theorem

Both Stratton-Chu's and Franz's formulas for the diffraction of electromagnetic fields can be derived from the Vector Green Theorem [50]:

$$\iiint_V \Big(\mathbf{Q}\cdot\nabla\times\nabla\times\mathbf{P}-\mathbf{P}\cdot\nabla\times\nabla\times\mathbf{Q}\Big)dv=\iint_S\Big(\mathbf{P}\times\nabla\times\mathbf{Q}-\mathbf{Q}\times\nabla\times\mathbf{P}\Big)\cdot d\mathbf{s} \quad \text{(B.1)}$$

where the gradient ∇ is operated on the source coordinates, denoted henceforth as primed variables. Choosing $\mathbf{Q}=\mathbf{E}(\mathbf{r})$ and $\mathbf{P}=\hat{\mathbf{a}}G(\mathbf{r},\mathbf{r}')$, with $\hat{\mathbf{a}}$ being an arbitrary unit vector and $G(\mathbf{r},\mathbf{r}')$ the free-space scalar Green's function (B.2), solution of the equation:

$$\big(\nabla^2+k^2\big)G(\mathbf{r},\mathbf{r}')=-\delta(\mathbf{r}-\mathbf{r}') \quad \rightarrow \quad G(\mathbf{r},\mathbf{r}')=\frac{e^{-jk|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|}, \quad \text{(B.2)}$$

Stratton-Chu formula is derived. By selecting $\mathbf{P}=\mathbf{H}$ and the modified free-space Green function $\mathbf{Q}=\nabla' G\times\hat{\mathbf{a}}$, Franz equation is deduced. The proof of both formulations follows below.

B.1 Stratton-Chu Formula

Starting with Stratton-Chu formulation, the first term of the volume integral on the left hand side of equation (B.1) takes the form:

- $\nabla'\times\nabla'\times\mathbf{P}=\nabla'\times\nabla'\times(\hat{\mathbf{a}}G)=\nabla'(\nabla'\cdot\hat{\mathbf{a}}G)-\nabla'^2(\hat{\mathbf{a}}G)=$

$$=\nabla'(\hat{\mathbf{a}}\cdot\nabla'G)+\hat{\mathbf{a}}\big(k^2G+\delta(\mathbf{r}-\mathbf{r}')\big)$$

- $\mathbf{E}\cdot\nabla'(\hat{\mathbf{a}}\cdot\nabla'G)=\nabla'\cdot\big(\mathbf{E}(\hat{\mathbf{a}}\cdot\nabla'G)\big)-(\hat{\mathbf{a}}\cdot\nabla'G)(\nabla'\cdot\mathbf{E})$

$$\Rightarrow \iiint_{V'}\mathbf{Q}\cdot\nabla'\times\nabla'\times\mathbf{P}\,dv'=\iiint_{V'}\big[\nabla'\cdot(\mathbf{E}(\hat{\mathbf{a}}\cdot\nabla'G))-(\hat{\mathbf{a}}\cdot\nabla'G)(\nabla'\cdot\mathbf{E})$$
$$+\hat{\mathbf{a}}\cdot k^2G\mathbf{E}+\hat{\mathbf{a}}\cdot\delta(\mathbf{r}-\mathbf{r}')\mathbf{E}\big]dv',$$

which turns, after applying the divergence theorem [50] to the first term in the volume integral, Maxwell equation $\nabla' \cdot \mathbf{E} = \frac{\rho}{\varepsilon}$ to the second term, and further rearranging the vectors, into:

$$\Rightarrow \; = \hat{\mathbf{a}} \cdot \iint\limits_{S'} (\hat{\mathbf{n}} \cdot \mathbf{E}) \nabla' G \, ds' + \hat{\mathbf{a}} \cdot \iiint\limits_{V'} \left[k^2 G \, \mathbf{E} - \tfrac{\rho}{\varepsilon} \nabla' G + \delta(\mathbf{r} - \mathbf{r}') \mathbf{E}(\mathbf{r}') \right] dv',$$

for $\hat{\mathbf{n}}$ an outgoing unit vector normal to the bounding surface S' of the integration volume V' such that $d\mathbf{s}' = \hat{\mathbf{n}} ds'$. The integral property of the delta function, $\iiint\limits_{V'} \delta(\mathbf{r} - \mathbf{r}') \mathbf{E}(\mathbf{r}') dv' = \mathbf{E}(\mathbf{r})$, can be used to simplify the last term in the previous volume integral, however, care must be exercised when the point $\mathbf{r}$ lies on the volume bounding surface S', since the volume integral contributes half the amount contributed when $\mathbf{r}$ is completely inside V'. As a result:

$$\iiint\limits_{V'} \delta(\mathbf{r} - \mathbf{r}') \mathbf{E}(\mathbf{r}') dv' = \begin{vmatrix} \mathbf{E}(\mathbf{r}) & if & \mathbf{r} \in V', \notin S' \\ \frac{1}{2}\mathbf{E}(\mathbf{r}) & if & \mathbf{r} \in S' \\ 0 & if & \mathbf{r} \notin V', S' \end{vmatrix} \tag{B.3}$$

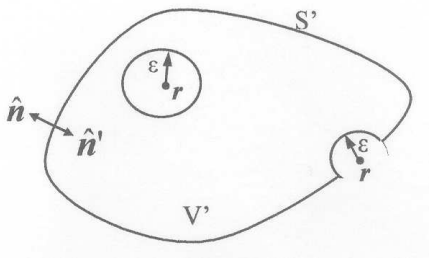


Figure B.1: Vector notation for the Stratton-Chu and Franz formulations.

For the second term inside the volume integral on the left hand side of equation (B.1), we use the vector wave equation $\nabla' \times \nabla' \times \mathbf{E} - k^2 \mathbf{E} = -j\omega\mu\mathbf{J} - \nabla' \times \mathbf{J_m}$, such that it results in:

- $\mathbf{P} \cdot \nabla' \times \nabla' \times \mathbf{Q} = \hat{\mathbf{a}} \cdot \nabla' \times \nabla' \times \mathbf{E} = \hat{\mathbf{a}} \cdot G(k^2 \mathbf{E} - j\omega\mu\mathbf{J} - \nabla' \times \mathbf{J_m})$

Using the vector identity $\nabla' \times (G\mathbf{J_m}) = G\nabla' \times \mathbf{J_m} - \mathbf{J_m} \times \nabla' G$, and applying Stoke's

theorem in $3D$ to the volume integral of $\nabla' \times (G\mathbf{J_m})$ we obtain:

$$\Rightarrow \iiint_{V'} \mathbf{P} \cdot \nabla' \times \nabla' \times \mathbf{Q}\, dv' = \hat{\mathbf{a}} \cdot \Big(\iiint_{V'} \big[k^2 G\mathbf{E} - j\omega\mu\mathbf{J}G - \mathbf{J_m} \times \nabla' G\big] dv' - \iint_{S'} \hat{\mathbf{n}} \times G\mathbf{J_m}\, ds' \Big).$$

Moving now to the right hand side of equation (B.1), we find:

- $\mathbf{P} \times \nabla' \times \mathbf{Q} \cdot \hat{\mathbf{n}} = \hat{\mathbf{a}}\, G \times \nabla' \times \mathbf{E} \cdot \hat{\mathbf{n}} = \hat{\mathbf{a}}\, G \times (-j\omega\mu\mathbf{H} - \mathbf{J_m}) \cdot \hat{\mathbf{n}}$

after using Maxwell equation $\nabla' \times \mathbf{E} = -j\omega\mu\mathbf{H} - \mathbf{J_m}$. And by means of the vector identity $\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = \mathbf{B} \cdot (\mathbf{C} \times \mathbf{A})$, it turns into:

$$\Rightarrow \iint_{S'} \mathbf{P} \cdot \nabla' \times \mathbf{Q} \cdot \hat{\mathbf{n}}\, ds' = \hat{\mathbf{a}} \cdot \iint_{S'} \hat{\mathbf{n}} \times (j\omega\mu G\mathbf{H} + G\mathbf{J_m})\, ds'.$$

Similarly with the last term on the right hand side of equation (B.1), it can be expressed as:

- $\mathbf{Q} \times \nabla' \times \mathbf{P} \cdot \hat{\mathbf{n}} = \big(\mathbf{E} \times \nabla' \times \hat{\mathbf{a}}\, G\big) \cdot \hat{\mathbf{n}} = \big(\mathbf{E} \times \nabla' G \times \hat{\mathbf{a}}\big) \cdot \hat{\mathbf{n}}$

with the aid of the vector identity $\nabla' \times \hat{\mathbf{a}}\, G = G\, \nabla' \times \hat{\mathbf{a}} - \hat{\mathbf{a}} \times \nabla' G = -\hat{\mathbf{a}} \times \nabla' G$. And by means of the same vector identity as the one used previously, $\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = \mathbf{B} \cdot (\mathbf{C} \times \mathbf{A})$, it becomes:

$$\Rightarrow \iint_{S'} \mathbf{Q} \cdot \nabla' \times \mathbf{P} \cdot \hat{\mathbf{n}}\, ds' = \hat{\mathbf{a}} \cdot \iint_{S'} (\hat{\mathbf{n}} \times \mathbf{E}) \times \nabla' G\, ds'.$$

Finally, adding all the terms, eliminating those that cancel out and rearranging the rest, the following relation is obtained which is satisfied for each arbitrary vector $\hat{\mathbf{a}}$ multiplying on both sides of the equality:

$$\begin{aligned} &\hat{\mathbf{a}} \cdot \iiint_{V'} \Big[j\omega\mu\mathbf{J}G + \mathbf{J_m} \times \nabla' G - \frac{\rho}{\varepsilon} \nabla' G\Big] dv' = \\ &\hat{\mathbf{a}} \cdot \mathbf{E}(\mathbf{r}) + \hat{\mathbf{a}} \cdot \iint_{S'} \Big(j\omega\mu G(\hat{\mathbf{n}} \times \mathbf{H}) - (\hat{\mathbf{n}} \times \mathbf{E}) \times \nabla' G - (\hat{\mathbf{n}} \cdot \mathbf{E}) \nabla' G\Big)\, ds' \end{aligned} \tag{B.4}$$

Due to the singularity of the free space scalar Green function (B.2), in the

previous result (B.4), the surface integral must be understood as the Cauchy Principal Value of the integral over the closed surface S' excluding an ϵ neighborhood of the singularity, and a factor of $\frac{1}{2}$ must be placed in front of the term $\mathbf{E}(\mathbf{r})$.

The final Stratton-Chu expression for the value of $\mathbf{E}(\mathbf{r})$ at any point in V' is given by:

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}_{inc}(\mathbf{r}) + T \iint_{S'} \Big(- j\omega\mu(\hat{\mathbf{n}}' \times \mathbf{H})G + (\hat{\mathbf{n}}' \times \mathbf{E}) \times \nabla' G + (\hat{\mathbf{n}}' \cdot \mathbf{E})\nabla' G \Big)\, ds'$$
$$with \;\; T = \begin{cases} 1 \text{ if } \mathbf{r} \notin S' \\ 2 \text{ if } \mathbf{r} \in S' \end{cases} \tag{B.5}$$

In equation (B.5) $\hat{\mathbf{n}}$ has been exchanged with $\hat{\mathbf{n}}' = -\hat{\mathbf{n}}$ pointing inwards towards the integration volume, and the volume integral over the sources have been identified as the incident field $\mathbf{E}_{inc}$. When applied to the evaluation of the fields scattered by the lithographic mask, the integration over S' surrounding the photomask reduces to the integration over the area of the photomask due to the radiation condition that ensures that the fields vanish as $r \to \infty$.

B.2 Franz Formula

In order to derive Franz formula of direct integration of the fields, we need to begin by substituting $\mathbf{P} = \mathbf{H}$ and $\mathbf{Q} = \nabla' \times G\hat{\mathbf{a}}$ in equation (B.1). The first term on the left hand side takes the form:

- $\mathbf{Q} \cdot \nabla' \times \nabla' \times \mathbf{P} = (\nabla' \times G\hat{\mathbf{a}}) \cdot \nabla' \times \nabla' \times \mathbf{H} =$
$$= (\nabla' \times G\hat{\mathbf{a}}) \cdot (k^2\mathbf{H} - j\omega\varepsilon\mathbf{J_m} + \nabla' \times \mathbf{J}) =$$
$$= (\nabla' \times G\hat{\mathbf{a}}) \cdot (k^2\mathbf{H} + \nabla' \times \mathbf{J}) + j\omega\varepsilon(\mathbf{J_m} \times \nabla' G) \cdot \hat{\mathbf{a}}$$

with the aid of the identities $\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = \mathbf{C} \cdot (\mathbf{A} \times \mathbf{B})$ and $\nabla' \times G\hat{\mathbf{a}} = \nabla' G \times \hat{\mathbf{a}}$.

For the second term of the left hand side we have:

- $\mathbf{P}\cdot\nabla'\times\nabla'\times\mathbf{Q} = \mathbf{H}\cdot\nabla'\times\nabla'\times\nabla'\times G\hat{\mathbf{a}} = \mathbf{H}\cdot\nabla'\times\Big[\nabla'(\hat{\mathbf{a}}\cdot\nabla' G)+\hat{\mathbf{a}}(k^2 G+\delta(\mathbf{r}-\mathbf{r}'))\Big] =$
$= \mathbf{H}\cdot\nabla'\times\big(\hat{\mathbf{a}}k^2 G + \hat{\mathbf{a}}\delta(\mathbf{r}-\mathbf{r}')\big)$

using the identity $\nabla'\cdot G\hat{\mathbf{a}} = G\nabla'\cdot\hat{\mathbf{a}} + \hat{\mathbf{a}}\cdot\nabla' G = \hat{\mathbf{a}}\cdot\nabla' G$, and given that the curl of the gradient of a scalar is zero.

The first term on the right hand side of the vector Green theorem (B.1) can be written as:

- $\mathbf{P}\times\nabla'\times\mathbf{Q} = \mathbf{H}\times\nabla'\times(\nabla'\times\hat{\mathbf{a}}G) = \mathbf{H}\times\Big[\nabla'(\hat{\mathbf{a}}\cdot\nabla' G) + \hat{\mathbf{a}}\big(k^2 G + \delta(\mathbf{r}-\mathbf{r}')\big)\Big]$

or, after applying $\nabla'(\hat{\mathbf{a}}\cdot\nabla' G) = (\hat{\mathbf{a}}\cdot\nabla')\nabla' G - \nabla' G\cdot\nabla'\hat{\mathbf{a}} = (\hat{\mathbf{a}}\cdot\nabla')\nabla' G$, as:

$\Rightarrow \iint\limits_{S'}\Big\{\mathbf{H}\times\big[(\hat{\mathbf{a}}\cdot\nabla')\nabla' G + k^2 G\hat{\mathbf{a}} + \delta(\mathbf{r}-\mathbf{r}')\hat{\mathbf{a}}\big]\Big\}\cdot\hat{\mathbf{n}}ds'$,

By means of the Divergence theorem [50] applied to the last term within the above surface integral and using the vector identity $\mathbf{A}\cdot(\mathbf{B}\times\mathbf{C}) = \mathbf{C}\cdot(\mathbf{A}\times\mathbf{B})$, on the remaining first two terms, this last integral is expressed as:

$\Rightarrow \iint\limits_{S'}\Big[(\hat{\mathbf{n}}\times\mathbf{H})\cdot(\hat{\mathbf{a}}\cdot\nabla')\nabla' G + (\hat{\mathbf{n}}\times\mathbf{H})\cdot k^2 G\hat{\mathbf{a}}\Big]ds' + \iiint\limits_{V'}\nabla'\cdot\big(\mathbf{H}\times\hat{\mathbf{a}}\delta(\mathbf{r}-\mathbf{r}')\big)dv'$,

It is convenient to expand the term within the last volume integral by means of another vector identity:

- $\iiint\limits_{V'}\nabla'\cdot\big(\mathbf{H}\times\hat{\mathbf{a}}\delta(\mathbf{r}-\mathbf{r}')\big)dv' = \iiint\limits_{V'}\big(\hat{\mathbf{a}}\delta(\mathbf{r}-\mathbf{r}')\cdot\nabla'\times\mathbf{H} - \mathbf{H}\cdot\nabla'\times\hat{\mathbf{a}}\delta(\mathbf{r}-\mathbf{r}')\big)dv'$

$= \iiint\limits_{V'}\big(\hat{\mathbf{a}}\delta(\mathbf{r}-\mathbf{r}')\cdot(j\omega\varepsilon\mathbf{E}+\mathbf{J}) - \mathbf{H}\cdot\nabla'\times\hat{\mathbf{a}}\delta(\mathbf{r}-\mathbf{r}')\big)dv'$

after using Maxwell equation $\nabla'\times\mathbf{H} = j\omega\varepsilon\mathbf{E}+\mathbf{J}$ for the magnetic field on the first term. This same Maxwell equation is used to expand the last term of the surface integral on the right hand side of equation (B.1) as follows:

- $\iint\limits_{S'} \big[\mathbf{Q} \times \nabla' \times \mathbf{P}\big] \cdot \hat{\mathbf{n}} ds' = \iint\limits_{S'} \big[(\nabla' \times G\hat{\mathbf{a}}) \times (j\omega\varepsilon\mathbf{E} + \mathbf{J})\big] \cdot \hat{\mathbf{n}} ds'$

$$= \iint\limits_{S'} \big[(\nabla' \times G\hat{\mathbf{a}}) \times j\omega\varepsilon\mathbf{E}\big] \cdot \hat{\mathbf{n}} ds' + \iiint\limits_{V'} \nabla' \cdot \big[(\nabla' \times G\hat{\mathbf{a}}) \times \mathbf{J}\big] dv'$$

Several vector identities can be applied to both the above surface and volume integrals in order to formulate them in a more suitable form. To the volume integral we apply $\nabla' \cdot (\mathbf{A} \times \mathbf{B}) = \mathbf{B} \cdot \nabla' \times \mathbf{A} - \mathbf{A} \cdot \nabla' \times \mathbf{B}$, and the the surface integral $\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = \mathbf{B} \cdot (\mathbf{A} \times \mathbf{C})$, such that:

- $\iiint\limits_{V'} \nabla' \cdot \big[(\nabla' \times G\hat{\mathbf{a}}) \times \mathbf{J}\big] dv' = \iiint\limits_{V'} \big[\mathbf{J} \cdot \nabla' \times (\nabla' \times G\hat{\mathbf{a}}) - (\nabla' \times G\hat{\mathbf{a}}) \cdot (\nabla' \times \mathbf{J})\big] dv'$

$= \iiint\limits_{V'} \big[\mathbf{J} \cdot \nabla'(\nabla' G \cdot \hat{\mathbf{a}}) + \mathbf{J} \cdot k^2 G\hat{\mathbf{a}} + \mathbf{J} \cdot \hat{\mathbf{a}}\delta(\mathbf{r} - \mathbf{r}') - (\nabla' \times G\hat{\mathbf{a}}) \cdot (\nabla' \times \mathbf{J})\big] dv'$

- $\iint\limits_{S'} \big[(\nabla' \times G\hat{\mathbf{a}}) \times (j\omega\varepsilon\mathbf{E} + \mathbf{J})\big] \cdot \hat{\mathbf{n}} ds' = \iint\limits_{S'} j\omega\varepsilon(\nabla' G \times \hat{\mathbf{a}}) \cdot (\hat{\mathbf{n}} \times \mathbf{E}) ds'$

$= \iint\limits_{S'} j\omega\varepsilon\hat{\mathbf{a}} \cdot \big((\hat{\mathbf{n}} \times \mathbf{E}) \times \nabla' G\big) ds'$

Adding together all terms on both sides of the equality (B.1), and eliminating those terms that cancel out, we get for every vector $\hat{\mathbf{a}}$ the following relation:

$$\begin{aligned} &\hat{\mathbf{a}} \cdot \iiint\limits_{V'} \big[j\omega\varepsilon\mathbf{J_m} \times \nabla' G + (\mathbf{J} \cdot \nabla')\nabla' G + k^2 \mathbf{J} G\big] dv' = \\ &= \hat{\mathbf{a}} \cdot j\omega\varepsilon\mathbf{E} + \hat{\mathbf{a}} \cdot \iint\limits_{S'} \big[((\hat{\mathbf{n}} \times \mathbf{H}) \cdot \nabla')\nabla' G + k^2(\hat{\mathbf{n}} \times \mathbf{H})G + j\omega\varepsilon(\hat{\mathbf{n}} \times \mathbf{E}) \times \nabla' G\big] ds' \end{aligned} \tag{B.6}$$

The final expression of Franz formula as it will be used for photolithographic calculations is obtained after identifying the volume integral over the sources as

generating the incident field, and after interchanging $\hat{\mathbf{n}}$ by $\hat{\mathbf{n}}' = -\hat{\mathbf{n}}$:

$$\begin{aligned}\mathbf{E}(\mathbf{r}) =& \mathbf{E}_{inc}(\mathbf{r}) + \\ &+ \; T \iint\limits_{S'} \Big(- j\omega\mu(\hat{\mathbf{n}}' \times \mathbf{H})G + \frac{1}{j\omega\varepsilon}\big((\hat{\mathbf{n}}' \times \mathbf{H}) \cdot \nabla'\big)\nabla' G + (\hat{\mathbf{n}}' \times \mathbf{E}) \times \nabla' G \Big)\, ds' \\ &with \;\; T = \begin{cases} 1 \text{ if } \mathbf{r} \notin S' \\ 2 \text{ if } \mathbf{r} \in S' \end{cases}\end{aligned} \tag{B.7}$$

A more compact expression of Franz formula, which also includes the electric and magnetic sources, reads as follows:

$$\begin{aligned}\mathbf{E}(\mathbf{r}) = \mathbf{E}_{inc}(\mathbf{r}) + \frac{1}{j\omega\varepsilon}\nabla \times \nabla \times \Big[\iiint\limits_{V'} \mathbf{J} G dv' + \iint\limits_{S'} \hat{\mathbf{n}}' \times \mathbf{H} G \, ds' \Big] \\ - \nabla \times \Big[\iiint\limits_{V'} \mathbf{J_m} G dv' - \iint\limits_{S'} \hat{\mathbf{n}}' \times \mathbf{E} G \, ds' \Big]\end{aligned} \tag{B.8}$$

Both expressions (B.7) and (B.8) can also be obtained by means of the free-space magnetic dyadic Green function, $\mathbf{Q} = \nabla' G \times \bar{\mathbf{I}}$ [47], and still more compactly formulated if we employ Dyadic notation [96].

From the previous derivations we observe that by utilizing the modified free-space Green function $\nabla \times G\hat{\mathbf{a}}$, the term $\hat{\mathbf{n}} \cdot \mathbf{E}(\mathbf{r})$, which was present in Stratton-Chu formula, is eliminated and only the field components tangent to the surface S' contribute to the integral. When evaluating the fields in the Fraunhofer region, the term $\hat{\mathbf{n}} \cdot \mathbf{E}(\mathbf{r})$ introduces component along the propagation direction $\hat{\mathbf{r}}$ which result in a non-TEM wave, contrary to the expected behavior of the fields in the far field due to a source of finite spatial extent. Franz formula, on the other hand, does predict TEM diffracted waves in the far field.

The relation between Franz and Stratton-Chu formulas was deduced in section 2.1.1, equation (2.10), which is repeated here for completeness:

$$\mathbf{E}^F(\mathbf{r}) = \mathbf{E}^{SC}(\mathbf{r}) + \frac{1}{j\omega\varepsilon} \oint_C \Big[(\hat{\mathbf{l}} \cdot \mathbf{H_o}) \nabla' \varphi \Big] dl \tag{B.9}$$

As was recently discussed by Tai [47], it is observed that the line integrals added by Stratton and Chu to account for the discontinuities of the fields on the surface S' are inherently contained in Franz's formulas. Tai concluded that Franz's formulation of the vectorial Huygens principle is superior, because it satisfies Maxwell equations for both continuous and discontinuous of fields. Both formulations are equivalent for continuous fields on S.

Relation (B.9) is obtained after expanding ∇ into components such that we can apply the vector identity $\nabla \cdot (\hat{\mathbf{a}}\varphi) = \varphi \nabla \cdot \hat{\mathbf{a}} + \hat{\mathbf{a}} \cdot \nabla\varphi$ to each component φ, as well as the Surface Divergence Theorem [49, 50] to the first term of the volume integral, that is:

$$\iint_{S'} \nabla' \cdot (\hat{\mathbf{n}}' \times \mathbf{H}) \nabla' G \, ds' = \sum_{\alpha = x', y', z'} \hat{\mathbf{e}}_\alpha \iint_{S'} \left[\nabla' \cdot \left((\hat{\mathbf{n}}' \times \mathbf{H}) \frac{\partial G}{\partial \alpha'} \right) - (\hat{\mathbf{n}}' \times \mathbf{H}) \cdot \nabla' \frac{\partial G}{\partial \alpha'} \right] ds' =$$

$$= \sum_{\alpha = x', y', z'} \hat{\mathbf{e}}_\alpha \left\{ - \oint_{C'} \hat{\mathbf{n}}_\mathbf{s} \cdot \left[(\hat{\mathbf{n}} \times \mathbf{H}) \frac{\partial G}{\partial \alpha'} \right] dl' - \iint_{S'} (\hat{\mathbf{n}}' \times \mathbf{H}) \cdot \nabla' \frac{\partial G}{\partial \alpha'} ds' \right\}$$

where $\hat{\mathbf{n}}_\mathbf{s}$ is a unit vector normal to the contour C' pointing outwards on the surface S' as illustrated in figure 2.4. It is given by $\hat{\mathbf{n}}_\mathbf{s} = \hat{\mathbf{l}} \times \hat{\mathbf{n}}'$, where $\hat{\mathbf{l}}$ is the unit vector tangent to the contour C at each point and $\hat{\mathbf{n}}'$ is the normal to the surface S pointing inwards toward the volume V. The negative sign in front of the line integral $\oint$ in the first term of the last equation comes from the fact that $\mathbf{H}$ needs to be taken as the difference between the magnetic field on the outer side of the contour C minus the magnetic field on the inner side of the contour C, that is, $\mathbf{H} = \mathbf{H}_{out} - \mathbf{H}_{in}$, and we assumed $\mathbf{H}_{out} = 0$. The relation between these vectors is plotted in figure 2.4. Grouping all three vector components of $\nabla' G$ we finally find equation (2.8) that relates Franz and Stratton-Chu formulations as derived in section 2.1.1:

$$\iint_{S'} \nabla' \cdot (\hat{\mathbf{n}}' \times \mathbf{H_o}) \nabla' G \, ds' = - \iint_{S'} \left((\hat{\mathbf{n}}' \times \mathbf{H_o}) \cdot \nabla' \right) \nabla' G \, ds' - \oint_{C'} \left[(\hat{\mathbf{n}}' \times \mathbf{H_o}) \nabla' G \right] \cdot \hat{\mathbf{n}}_\mathbf{s} dl' \quad \text{(B.10)}$$

APPENDIX C

PTD Study of Rectangular Aperture

C.1 PTD on Perfect Electric Conductor

For the study of the field diffracted by a rectangular aperture on a perfect electric conducting plate, Babinet's principle was applied to the results derived by Ufimtsev for a conducting wedge [74]. In a similar fashion, results by the Method of Equivalent Currents could have been applied to this problem, however the asymptotic results available in literature for the edge fringe currents, $\mathbf{J}^{FR}$ and $\mathbf{M}^{FR}$, become singular when the observation direction is the continuation of a glancing incident ray, that is, at the shadow boundary [75, 76, 77]. As in Geometrical Theory of Diffraction, the asymptotic expansion losses its physical meaning in the vicinity of these boundaries and rather, Ufimtsev's formulation, which is free of singularities on the shadow boundaries, will be used. According to these results, the total field diffracted by a conducting object is the sum of the field due to the Physical Optics approximation plus the *fringe* fields arising from the excitation of some *fringe* currents in the vicinity of the object edges, that is:

$$\mathbf{E}^t = \mathbf{E}^{PO} + \mathbf{E}^{FR} \qquad\qquad \mathbf{H}^t = \mathbf{H}^{PO} + \mathbf{H}^{FR} \tag{C.1}$$

with

$$\mathbf{E}^{FR} = \int_{C'} 2\mathbf{E}^{EW}(\mathbf{r}, \mathbf{r}') G(\mathbf{r}, \mathbf{r}')\, dl' \tag{C.2}$$

and

$$\mathbf{E}^{EW}(\mathbf{r}, \mathbf{r}') = (\mathbf{E}_{inc} \cdot \hat{\mathbf{l}}')\mathbf{D}_{\mathbf{s}}^{EW}(\mathbf{r}, \mathbf{r}') + \eta\, (\mathbf{H}_{inc} \cdot \hat{\mathbf{l}}')\mathbf{D}_{\mathbf{h}}^{EW}(\mathbf{r}, \mathbf{r}') \tag{C.3}$$

In equation (C.2), $\mathbf{E}^{EW}$ is the elementary wave radiated by each differential arc-length dl' of C', and the unit vector $\hat{\mathbf{l}}'$ is tangent to the edge at every point. $E_{l'}^{inc}$ and $H_{l'}^{inc}$ represent the incident field component parallel to the boundary. Asymptotic expressions for the functions $\mathbf{D_s}^{EW}$ and $\mathbf{D_h}^{EW}$ valid at every point in space were provided by Ufimtsev in terms of the incident and observation angles [74]. In particular, when only one side of a half plane (wedge with $\alpha = 2\pi$) is illuminated at normal incidence ($\theta_{inc} = \frac{\pi}{2}$, $\varphi_{inc} = \frac{\pi}{2}$), and based on the notation of figure C.1, which is local to each edge point, we have:

$$\begin{aligned} \mathbf{D_s}^{EW} &= \hat{\mathbf{e}}_{\mathbf{z}_{edge}} D_s^{EW} = \hat{\mathbf{e}}_{\mathbf{z}_{edge}} \big(U(\sigma, \varphi_{inc}) + U(\sigma, 2\pi - \varphi_{inc})\big) \\ \mathbf{D_h}^{EW} &= \hat{\mathbf{e}}_{\mathbf{x}_{edge}} D_h^{EW} = \hat{\mathbf{e}}_{\mathbf{x}_{edge}} \big(V(\sigma, \varphi_{inc}) + V(\sigma, 2\pi - \varphi_{inc})\big) \end{aligned} \tag{C.4}$$

where

$$U = U^{total} - U^{PO} \qquad V = V^{total} - V^{PO}$$

$$V^{PO} = (cos\ \beta)^{-1} \qquad U^{PO} = -(cos\ \beta)^{-1}$$

$$U^{total} = \frac{1}{4}\Big[cot\frac{\sigma + \varphi_{inc}}{4} - cot\frac{\sigma - \varphi_{inc}}{4}\Big]$$

$$V^{total} = \frac{1}{4sin\ \sigma}\Big[cot\frac{\sigma + \varphi_{inc}}{4} + cot\frac{\sigma - \varphi_{inc}}{4}\Big]$$

and

$$cos\ \beta = sin\ \theta\ cos\ \varphi \qquad cos\ \sigma = -cos\ \beta$$

Ufimtsev also showed that in the asymptotic limit, $kr \longrightarrow \infty$, the main contribution at a given observation point arises from those rays lying on the diffraction cone, in which case the functions $\mathbf{D_s}^{EW}$ and $\mathbf{D_h}^{EW}$ simplify to the well-known GTD coefficients D_s^{FR} and D_h^{FR}, collected in equation (3.18).

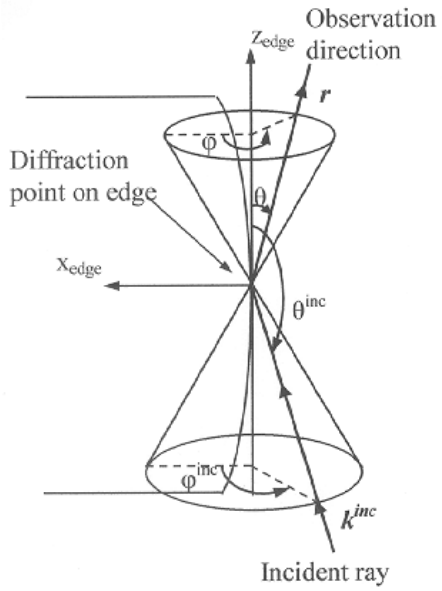
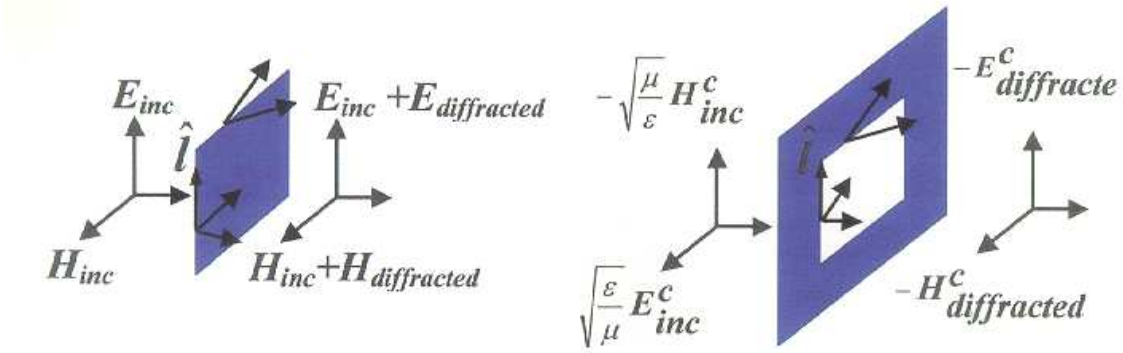


Figure C.1: Angle relation for edge diffraction

C.2 Babinet's Principle

Babinet's principle states that the fields scattered by complementary surfaces are also complementary. In particular, if the electric and magnetic fields scattered by the conducting plate of figure C.2(a), illuminated by $\mathbf{E}_{inc}$ and $\mathbf{H}_{inc}$, are $\mathbf{E}_{diffracted}$ and $\mathbf{H}_{diffracted}$, respectively, then the fields diffracted by the complementary aperture of figure C.2(b), illuminated by $-\sqrt{\frac{\mu}{\epsilon}}\ \mathbf{H}^c_{inc}$ and $\sqrt{\frac{\epsilon}{\mu}}\ \mathbf{E}^c_{inc}$, will relate to their complementary as $-\sqrt{\frac{\mu}{\epsilon}}\ \mathbf{H}^c_{diffracted} = -\mathbf{E}_{diffracted}$ and $\sqrt{\frac{\epsilon}{\mu}}\ \mathbf{E}^c_{diffracted} = -\mathbf{E}_{diffracted}$. Application of this principle to equations (C.2) to (C.3) yields the following fringe field in the far-field zone diffracted by apertures,

$$\mathbf{E^{FR}} = -\eta\ \hat{\mathbf{r}} \times \mathbf{H^{FR}} = -\eta\ \hat{\mathbf{r}} \times \int_{C'} 2\Big[(-\mathbf{H}_{inc} \cdot \hat{\mathbf{l}}')\mathbf{D}^{EW}_{\mathbf{s}} + \eta^{-1}(\mathbf{E_{inc}} \cdot \hat{\mathbf{l}}')\mathbf{D}^{EW}_{\mathbf{h}}\Big] G(\mathbf{r}, \mathbf{r}')\ dl' \quad \text{(C.5)}$$



(a) (b)

Figure C.2: Babinet's principle

C.3 Rectangular Aperture

Equation (C.5) can now be applied to the aperture of figure 3.10, assuming normal incidence. The local incident and observation angles at each edge point need to be expressed in terms of the global coordinate system by means of the appropriate coordinate transformations [97]. The following fringe fields at the entrance pupil of the $4X$ optical system were found:

$$\mathbf{E^{FR}} = -\eta\ \hat{\mathbf{r}} \times \hat{\mathbf{x}}$$

$$\Big\{ -2H_o \frac{e^{-jkr}}{4\pi r}\Big[e^{jkr_y \frac{w}{2}} \int_{-\frac{h}{2}}^{\frac{h}{2}} D_s^{EW}(\theta,\varphi)\ e^{jkr_x x'}\ dx' - e^{-jkr_y \frac{w}{2}} \int_{\frac{h}{2}}^{-\frac{h}{2}} D_s^{EW}(\theta,\varphi)\ e^{jkr_x x'}\ dx'\Big]$$

$$-\frac{2E_o}{\eta}\frac{e^{-jkr}}{4\pi r}\Big[e^{jkr_x \frac{h}{2}} \int_{\frac{w}{2}}^{-\frac{w}{2}} D_h^{EW}(\theta,\varphi)\ e^{jkr_y y'}\ dy' - e^{-jkr_x \frac{h}{2}} \int_{-\frac{w}{2}}^{\frac{w}{2}} D_h^{EW}(\theta,\varphi)\ e^{jkr_y y'}\ dy'\Big]\Big\} \quad \text{(C.6)}$$

where $\mathbf{E}^{inc} = \hat{\mathbf{y}}E_o$ and $-\mathbf{H}^{inc} = -\hat{\mathbf{x}}H_o$. The coefficients $D_{s,h}^{EW}(\theta,\varphi)$ need to be evaluated at each point on the edge according to the following expressions:

$$sin\ \varphi = -\frac{z}{\sqrt{(x \mp \frac{h}{2})^2 + z^2}} \qquad sin\ \theta = \frac{\sqrt{(x \mp \frac{h}{2})^2 + z^2}}{\sqrt{(x \mp \frac{h}{2})^2 + (y-y')^2 + z^2}} \quad \text{(C.7)}$$

on the upper and lower edges, and:

$$sin\ \varphi = -\frac{z}{\sqrt{(y \mp \frac{w}{2})^2 + z^2}} \qquad sin\ \theta = \frac{\sqrt{(y \mp \frac{w}{2})^2 + z^2}}{\sqrt{(y \mp \frac{w}{2})^2 + (x-x')^2 + z^2}} \quad \text{(C.8)}$$

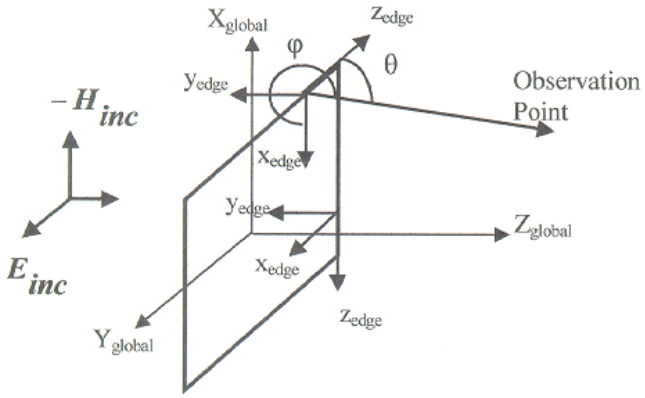


Figure C.3: Transformation from local to global coordinates.

on the right and left edges, which relate to the notation sketched in figure C.3.

This result is then assumed to propagate throughout a $4X$ optical projection system of $NA_o = \frac{0.85}{4} \sim 0.2$ at the entrance pupil, and the corresponding peak amplitude is measured at the center of the image plane as illustrated in figure C.4. The propagation distance r to the entrance pupil, assumed well into the far field region of the antenna, is much larger than the wavelength, while the aperture dimensions considered here are of the order of the wavelength. This allows us to approximate $\theta \approx \frac{\pi}{2}$ and $\varphi \approx \frac{\pi}{2}$, and along this observation direction, the singularities of both U^{total}, V^{total} and U^{PO}, V^{PO} cancel out, which yields the approximation:

$$D_s^{EW} \approx -\frac{1}{2}$$
$$D_h^{EW} \approx \frac{1}{2}$$

Therefore, with the following change of variables, $f = \frac{r_x}{NA}$ and $g = \frac{r_y}{NA}$ and dropping the constant phase factors, the image fields at the focal plane will be given by:

$$\mathbf{E}^{\mathbf{FR}}_{\mathbf{image}}(x,y) = -\eta\hat{\mathbf{r}} \times \hat{\mathbf{x}}\Big\{H_o\ 2h\ NA^2 \iint\limits_{f^2+g^2\leq 1} cos(\pi NAgw)sinc(\pi NAfh)e^{jk(fx+gy)}\ dfdg$$

$$\frac{E_o}{\eta}\ 2w\ NA^2 \iint\limits_{f^2+g^2\leq 1} cos(\pi NAfh)sinc(\pi NAgw)e^{jk(fx+gy)}\ dfdg\Big\} \quad \text{(C.9)}$$

To first order approximation, one can consider $cos(\pi NAgw)$, $sinc(\pi NAfh)$, $cos(\pi NAfh)$ and $sinc(\pi NAgw) \sim 1$ within the integration limits of equation (C.9), which is a reasonable approximation for small Numerical Apertures such as the one on the objet side for imaging systems of high reduction factors. This approximations yielded the following image field expression:

$$\mathbf{E}^{\mathbf{FR}}_{\mathbf{image}}(x,y) \simeq -\eta(\hat{\mathbf{r}} \times \hat{\mathbf{x}}) H_o \, \frac{NA^2 J_1\big(2\pi k\sqrt{x^2+y^2}\big)}{kr} \, (2h+2w) \tag{C.10}$$

where $J_1(\alpha)$ represents the Bessel function of the first kind, order 1, such that at the center of the image, $\frac{J_1(x)}{x} = \frac{1}{2}$ with $x = 0$.

Finally, extracting the main polarization component, E_y, from equation (C.10) and noting that with $4X$ optical reduction, one can assume $r_z \approx 1$, one obtains:

$$E^{FR}_{image} \simeq -E_o \, \frac{NA^2 J_1\big(k\sqrt{x^2+y^2}\big)}{kr} \big(2h+2w\big) \tag{C.11}$$

A similar analysis is carried out on the Physical Optics or Kirchhoff component of the field, which is equivalent to Kirchhoff integral for apertures. In particular, by means of either Babinet's principle or Image Theory, the far-field diffracted by an aperture on a perfect conducting plane can be expressed as [98]:

$$\mathbf{H}^{\mathbf{PO}} = -j\omega\epsilon \frac{e^{-jkr}}{4\pi r} (\hat{\hat{\mathbf{I}}} - \hat{\mathbf{r}}\hat{\mathbf{r}}) \cdot \iint_{A'} (2\hat{\mathbf{n}} \times \mathbf{E}^{inc}) e^{jk\hat{\mathbf{r}}\mathbf{r}'} dx'dy' =$$

$$j\omega\epsilon \frac{e^{-jkr}}{4\pi r} (\hat{\hat{\mathbf{I}}} - \hat{\mathbf{r}}\hat{\mathbf{r}}) \cdot \hat{\mathbf{x}} \; 2E_o \; h \; sinc(\pi NAfh) \; w \; sinc(\pi NAgw) \tag{C.12}$$

The electric field can be obtained at the entrance pupil (far-field region) by means of $\mathbf{E} = -\eta\hat{\mathbf{r}} \times \mathbf{H}$. The same approximations can then be applied to the physical optics field in order to derive the image field expression on the focal plane, which results in:

$$\mathbf{E}^{\mathbf{PO}}_{\mathbf{image}} \simeq -jk\hat{\mathbf{r}} \times (\hat{\hat{\mathbf{I}}} - \hat{\mathbf{r}}\hat{\mathbf{r}}) \cdot \hat{\mathbf{x}} \; 2E_o \; h \; w \; \frac{NA^2 J_1\big(k\sqrt{x^2+y^2}\big)}{kr} \tag{C.13}$$

For the main polarization component along the y-axis,

$$E^{PO}_{image} \simeq -jk \; r_z \; 2E_o \; h \; w \; \frac{NA^2 J_1\big(k\sqrt{x^2+y^2}\big)}{kr} \approx -jk \; 2E_o \; h \; w \; \frac{NA^2 J_1\big(k\sqrt{x^2+y^2}\big)}{kr} \tag{C.14}$$

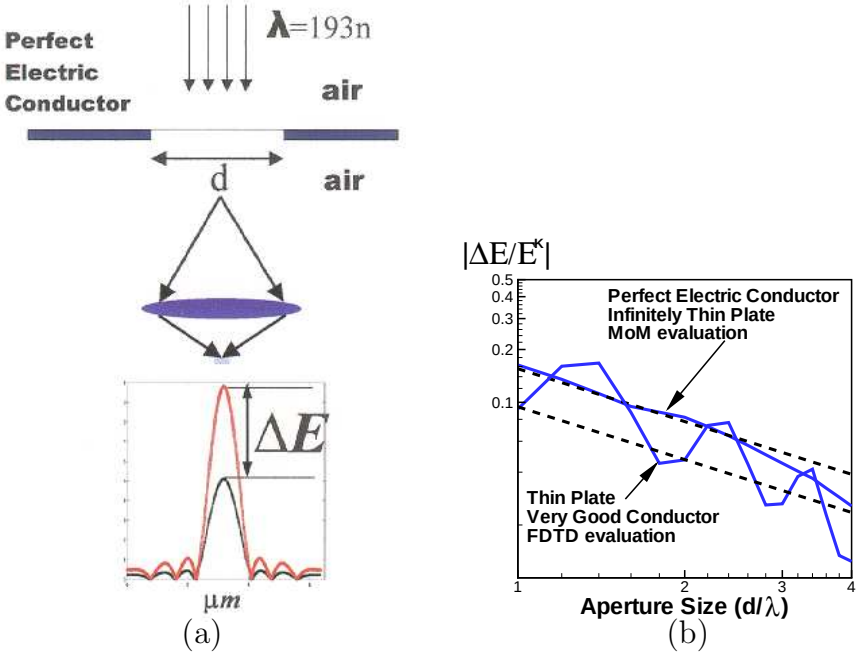


Figure C.4: (a)Application of PTD to 2D openings on perfect electric conducting plates and imaging through $4X$, $NA = 0.85$ optical system. Relative error on the field amplitude due to edge diffraction is measured at the peak of the image.(b) Relative amplitude error follows inverse law with opening size on both perfect electric thin conductors and very good thin conductors.

The final relative error due to the fringe fields at the image focal plane as measured at the peak amplitude turns out to obey the following relation with the opening dimensions:

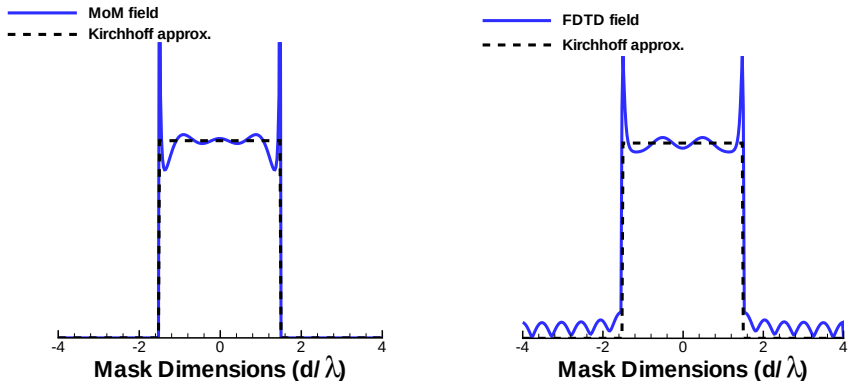
$$\frac{\Delta E}{E} = \frac{E^{FR}}{E^{PO}} = \frac{E^{FR}}{E^{K}} \approx \frac{-(2h+2w)}{-j2k\ h\ w} = -\frac{j}{2k}\frac{4}{\frac{2hw}{w+h}} \tag{C.15}$$

This effective mean size of the aperture dimensions turns out to be the harmonic mean of the width and the height, defined as $\frac{2wh}{w+h}$. In particular, when the aperture dimensions are of the order of the wavelength, its effective size can be

expressed as $\frac{d}{\lambda}\lambda$, with $\frac{d}{\lambda}$ no larger than 10, and equation (C.15) turns into:

$$\frac{\Delta E}{E} = \frac{E^{FR}}{E^{PO}} = \frac{E^{FR}}{E^{K}} \approx -\frac{j}{\pi}\frac{1}{\frac{d}{\lambda}} = -\frac{j0.32}{\frac{d}{\lambda}} \quad \left(-\frac{j0.16}{\frac{d}{\lambda}} \quad for\ 2D\ features \right) \tag{C.16}$$

This reciprocal law of the relative error due to Kirchhoff or the Physical Optics approximation is exactly reproduced by the results of figure C.4(b), which correspond to 2D apertures illuminated with a TM polarized electric field (lying on the plane of the page, normal to the aperture edges) at 193nm wavelength and $NA = 0.85$. The total electric field (Kirchhoff plus fringe fields) due to an aperture on a perfect electric conductor was evaluated by the Method of Moments [98, 99]. An example of the fields on the mask surface obtained by this method is plotted in figure C.5(a). Figure C.5(b), on the other hand, correspond to the fields on the surface of a thin layer of very good conductor ($\sigma \longrightarrow \infty$) as evaluated by the Finite-Difference Time-Domain software Tempest 6.0 [55].



(a) (b)

Figure C.5: Electric field amplitude on a perfect conductor surface with an aperture as evaluated by the Method of Moments (MoM). (b) Electric field amplitude on a very good thin conductor with an aperture as evaluated by FDTD software Tempest.

Slight differences can be appreciated between the solutions for a perfect conductor and good conductor, mainly on the opaque region where the field is consid-

ered zero on a perfect conductor. These differences presumably give place to the resonant effects on the relative error on a non-perfect conducting plate as plotted in figure C.4(b). As the opening size increases, the approximation $cos(\pi NAgw)$, $sinc(\pi NAfh)$, $cos(\pi NAfh)$ and $sinc(\pi NAgw) \sim 1$ breaks down and the error starts to deviate from the inverse law.

References

[1] G. E. Moore, "Cramming more components onto integrated circuits," *Electronics*, vol. **38, No 8**, pp. 114–117, April 19, 1965.

[2] A. K.-K. Wong, *Resolution Enhancement Techniques in Optical Lithography*, SPIE Press, 2001.

[3] F. Schellenberg, ed., *Selected Papers in Resolution Enhancement Techniques*, SPIE Press, 2004.

[4] P. J. Silverman, "The Intel lithography roadmap," *Intel Technology Journal*, vol. **6, Issue 2**, May 16, 2002.

[5] *International Technology RoadMap for Semiconductors*, http://public.itrs.net/, 2003.

[6] *http://www.asml.com*.

[7] H. J. Levinson, *Principles of Lithography*, SPIE Press, 2001.

[8] M. Born and E. Wolf, *Principles of Optics*, Pergamon Press, 6th ed., 1987.

[9] D. S. Goodman and A. E. Rosenbluth, "Condenser aberrations in kohler illumination," *Proceedings of the SPIE*, vol. **922**, pp. 108–134, 1988.

[10] Y. Borodovsky and et. al., "Lithography strategy for 65nm node," *Proceedings of the SPIE*, vol. **4754**, pp. 1–14, 2002.

[11] M. D. Levenson, N. S. Viswanathan, and R. A. Simpson, "Improving resolution in photolithography with phase shifting-mask," *IEEE Trans. Electr. Dev.*, vol. **ED-29(12)**, pp. 1828–1836, 1982.

[12] B. J. Lin, "The attenuated phase-shifting mask," *Solid State Technology*, vol. **35,No.1**, pp. 43–47, 1992.

[13] D. Lammers, "Intel drops 157 litho from roadmap," *EEtimes* , May 23, 2003.

[14] B. W. Smith, A. Bourov, L. Zavyalova, N. Lafferty, Y. Fan, F. Cropanese, M. Gower, N. Rizvi, and J. Webb, "Approaching the numerical aperture of H2O: Immersion lithography at 193nm," *Proc. of the SPIE*, vol. **5377**, pp. 273–284, 1998.

[15] B. Streefkerk, J. Baselmans, W. G. van A., J. Mulkens, . Hoogendam, M. Hoogendorp, D. Flagello, P. Graupner, and C. Zeiss, "Extending optical lithography with immersion," *Proc. of the SPIE*, vol. **5377**, pp. 285–305, 2004.

[16] E. Abbe, "Beitrage zur theorie des mikroskops und der mikroskopischen wahrnehmung — Contributions to the theory of the microscope and the nature of microscopic vision," *Archiv f. Mikroskopische Anat.*, vol. **9**, pp. 413–468, 1873.

[17] J. W. Goodman, *Statistical Optics*, Wiley, New York, 1985.

[18] H. Hopkins, "The concept of partial coherence in optics," *Proc.Royal Soc.London*, vol. **A208**, pp. 263–277, 1951.

[19] C. Pierrat and A. Wong, "The MEF revisited: Low k1 effects versus mask topography effects," *Proceedings of the SPIE*, vol. **5040**, pp. 193–202, 2003.

[20] A. K. Wong and A. R. Neureuther, "Mask topography effects in projection printing of phase-shifting masks," *IEEE Trans. on Electron Devices*, vol. **41, No 6**, pp. 895–902, 1994.

[21] M. S. Yeung and E. Barouch, "Limitation of the kirchhoff boundary conditions for aerial image simulation in 157 nm optical lithography," *IEEE Electron Device Lett.*, vol. **21 No 9**, pp. 433–435, Sep. 2000.

[22] M. Lam, K. Adam, and A. Neureuther, "Domain decomposition methods for simulation of printing and inspection of phase defects," *Proceedings of the SPIE*, vol. **5040**, pp. 1492–1501, 2003.

[23] P.-Y. Yan, "Understanding Bossung curve asymmetry and focus shift effect in euv lithography," *Proc. of SPIE*, vol. **4562**, pp. 1–9, 2001.

[24] K. Adam and A. R. Neureuther, "Simplified models for edge transitions in rigorous mask modeling," *Proc. of the SPIE*, vol. **4346 pt.1-2**, pp. 331–344, 2001.

[25] K. S. Yee, "Numerical solution of initial boundary value problems involving Maxwell equations in isotropic media," *IEEE Antennas and Propagation*, vol. **14, No 3**, pp. 302–307, 1966.

[26] A. Taflove and S. C. Hagness, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*, Artech House, 2nd ed., 2000.

[27] A. K. Wong and A. R. Neureuther, "Rigorous three-dimensional time-domain finite-difference electromagnetic simulation for photolithographic applications," *IEEE Trans. on Semiconductor Manufacturing*, vol. **8, No 4**, pp. 419–431, Nov. 1995.

[28] K. A. Michalski and J. R. Mosig, "Multilayered media Green's function in integral formulations," *IEEE Trans Antennas and Propagation*, vol. **45**, pp. 508–519, 1997.

[29] M. S. Yeung, "Three-dimensional mask transmission simulation using a single integral equation method," *Proc. of the SPIE*, vol. **3334**, pp. 704–713, 1998.

[30] M. S. Yeung, "Fast and rigurous three-dimensional mask diffraction simulation using Battle-Lemarie wavelet based multiresolution time-domain method," *Proc. of the SPIE*, vol. **5040**, pp. 69–77, 2003.

[31] D. Cole, E. Barouch, U. Hollerbach, and S. Orszag, "Derivation and simulation of higher numerical aperture scalar aerial images," *J. Appl. Phys.*, vol. **31**, pp. 4110–4119, 1992.

[32] D. Cole, E. Barouch, U. Hollerbach, and S. Orszag, "Extending scalar aerial image calculations to higher numerical apertures," *J. Cac. Sci. Technol. B*, vol. **10(6)**, pp. 3037–3041, 1992.

[33] M. S. Yeung, D. Lee, R. Lee, and A. R. Neureuther, "Extension of the Hopkins theory of partially coherent imaging to include thin-film interference effects," *proceedings of the SPIE*, vol. **1927**, pp. 452–463, 1993.

[34] M. S. Yeung, "Modeling high numerical aperture optical lithography," *Proceedings of the SPIE*, vol. **922**, pp. 149–167, 1988.

[35] D. G. Flagello and A. E. Rosenbluth, "Lithographic tolerances based on vector diffraction theory," *J.Vac.Sci.Technol.*, vol. **B10(6)**, pp. 2997–3003, Nov/Dec 1992.

[36] D. G. Flagello, T. Milster, and A. E. Rosenbluth, "Theory of high-na imaging in homogeneous thin films," *J. Opt. Soc. Am. A*, vol. **13, No 1**, pp. 53–64, 1996.

[37] Q.-D. Qian and F. A. Leon, "Fast algorithms for 3D high na lithography simulation," *Proceedings of the SPIE*, vol. **2440**, pp. 372–380, 1995.

[38] C. A. Mack, "A comprehensive optical lithography model," *Proceedings of SPIE*, vol. **538**, pp. 207–220, 1985.

[39] K. Adam, Y. Granik, A. Torres, and N. Cobb, "Improved modeling performance with an adapted vectorial formulation of the hopkins imaging equation," *Proceedings of the SPIE*, vol. **5040**, pp. 78–90, 2003.

[40] F. Dill, W. Hornberger, P. Hauge, and J. Shaw., "Characterization of positive photoresist," *IEEE Trans. Electron. Dev.*, vol. **ED-22(7)**, pp. 445–452, July 1975.

[41] J. W. Goodman, *Introduction to Fourier Optics*, McGraw-Hill, 1996.

[42] J. J. Stamnes, *Waves in Focal Regions: Propagation, Diffraction and Focusing of Light, Sound and Water Waves*, Institute of Physics Publishing. Adam Hilger Series on Optics and Optoelectronics, 1986.

[43] E. Wolf, "Electromagnetic diffraction in optical systems I. an integral representation of the image field," *Proc. Roy. Soc. A*, vol. **253**, pp. 349–357, 1959.

[44] E. Wolf and Y. Li, "Conditions for the validity of the Debye integral representation of focused fields," *Optics Communications*, vol. **39, No 4**, pp. 205–210, 1981.

[45] H. A. Eide and J. J. Stamnes, "Exact and approximate solutions for focusing of two-dimensional waves. II. numerical comparisons among exact, Debye, and Kirchhoff theories," *J. Opt. Soc. Am. A*, vol. **15**, pp. 1292–1307, 1998.

[46] J. A. Stratton, *Electromagnetic Theory*, McGraw-Hill, 1941.

[47] C. T. Tai, "Direct integration of field equations," *Progress In Electromagnetic Research, PIER*, vol. **28**, pp. 339–359, 2000.

[48] R. Mittra, ed., *Computer Techniques for Electromagnetics*, Hemisphere Publishing Co., 1987.

[49] W. C. Chew, *Waves and Fields in Inhomogeneous Media*, IEEE Press, New York, 1995.

[50] C. T. Tai, *Dyadic Green Functions in Electromagnetic Theory*, IEEE Press, New York, 2nd ed., 1994.

[51] A. Ishimaru, *Electromagnetic Wave Propagation, Radiation and Scattering*, Prentice Hall, 1991.

[52] C. Balanis, *Advanced Engineering Electromagnetics*, John Wiley & Sons, 1989.

[53] B. Richards and E. Wolf, "Electromagnetic diffraction in optical systems II. structure of the image in an aplanatic system," *Proc. Roy. Soc. A*, vol. **253**, pp. 358–379, 1959.

[54] M. Mansuripur, "Distribution of light at and near the focus of high-numerical aperture objectives," *J. Opt. Soc. Am. A*, vol. **3, No 12**, pp. 2086–2093, 1986.

[55] A. K. Wong and A. R. Neureuther, "Polarization effects in mask transmission," *Proc. of SPIE*, vol. **1674**, pp. 193–200, 1992.

[56] K. Adam, S. Hotta, and A. R. Neureuther, "Characterization of phase defects in phase shift masks," *J. Vac. Sci. Technol. B*, vol. **18(6)**, pp. 1–5, 2000.

[57] S. Hotta, T. V. Pistor, K. Adams, and A. E. Rosenbluth, "Effects of shifter edge topography on through focus performance," *Proceedings of the SPIE*, vol. **4186**, pp. 827–837, 2001.

[58] O. M. Bucci and G. Pelosi, "From wave theory to ray optics," *IEEE Antennas and Propagation Magazine*, vol. **36**, p. 4, 1994.

[59] K. Miyamoto and E. Wolf, "Generalization of the Maggi-Rubinowicz theory of the boundary diffraction wave – part i," *J. Opt. Soc. Am.*, vol. **52, No 6**, pp. 615–625, 1962.

[60] A. Rubinowicz. *Ann. Physik*, vol. **53**, p. 257, 1917.

[61] K. Miyamoto and E. Wolf, "Generalization of the Maggi-Rubinowicz theory of the boundary diffraction wave – part ii," *J. Opt. Soc. Am.*, vol. **52, No 6**, pp. 626–637, 1962.

[62] E. W. Marchand and E. Wolf, "Boundary diffraction wave in the domain of the Rayleigh-Kirchhoff diffraction theory," *J. Opt. Soc. of Am.*, vol. **52, No 7**, pp. 761–767, 1962.

[63] J. Boersma, "Simple solution of Sommerfeld's half-plane diffraction problem," *J. Applied Science and Engineering A*, vol. **2**, pp. 187–193, 1977.

[64] E. T. Copson, "On an integral equation arisin in the theory of diffraction," *Quart. J. Math. Oxford Ser.*, vol. **17**, p. 19, 1946.

[65] D. S. Jones, "A simplifying technique in the solution of a class of diffraction problems," *Quart. J. Math. Oxford Ser.(2)*, vol. **3**, p. 189, 1952.

[66] P. C. Clemmow, "A method for the exact solution of a class of two-dimensional diffraction problems," *Proc. R. Soc. London Ser. A*, vol. **205**, p. 286, 1951.

[67] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, Dover Pub. Inc., 1965.

[68] A. Kalashnikov. *Journal of Russian Physical and Chemical Society. Physical Division*, vol. **44, No 3**, pp. 137–144, 1912.

[69] A. Rubinowicz. *Ann. Physik*, vol. **73**, p. 339, 1924.

[70] A. Sommerfeld, *Lectures on Theoretical Physics IV. Optics*, Academic Press.

[71] P. Y. Ufimtsev, "Rubinowicz and the modern theory of diffracted rays," *Electromagnetics*, vol. **15**, pp. 547–565, 1995.

[72] J. Keller, "Geometrical theory of diffraction," *J. Opt. Soc. Am. A*, vol. **52, No 2**, pp. 116–130, 1962.

[73] P. Y. Ufimtev, *Method of Edge Waves in the Physical Theory of Diffraction*, Foreign Technology Division, Air Force Systems Command, 1971.

[74] P. Y. Ufimtsev, "Elementary edge waves and the physical theory of diffraction," *Electromagnetics*, vol. **11**, pp. 125–160, 1991.

[75] A. Michaeli, "Elimination of infinities in equivalent edge currents, part I: Fringe current components," *IEEE Transactions on Antennas and Propagation*, vol. **AP-34, No. 7**, pp. 912–918, 1986.

[76] M. Ando, "Radiation pattern analysis of reflector antennas," *Electronics and Communications in Japan, Part I*, vol. **68, No 4**, pp. 93–102, 1985.

[77] K. M. Mitzner, "Incremental length diffraction coefficients," *Aircraft Division Northrop Corp. Tech. Rep.*, vol. **AFAL-TR-73-296**, 1974.

[78] D.-W. Duan, Y. Rahmat-Samii, and J. P. Mahon, "Scattering from a circular disk: A conparative study of PTD and GTD techniques," *Proc. IEEE*, vol. **79, No 10**, pp. 1472–1480, 1991.

[79] A. Michaeli, "Incremental diffraction coefficients for the extended physical theory of diffraction," *IEEE Transactions on Antennas and Propagation*, vol. **43, No. 7**, pp. 732–734, 1995.

[80] J. Tirapu-Azpiroz, P. Burchard, and E. Yablonovitch, "Boundary layer model to account for thick mask effects in photolithography," *Proceedings of the SPIE*, vol. **5040**, pp. 1611–1619, 2003.

[81] J. Tirapu-Azpiroz and E. Yablonovitch, "Modeling of near-field effects in sub-wavelength deep ultraviolet lithography," in *Future Trends of Microelectronics 2003*, S. Luryi, J. Xu, and A. Zaslavsky, eds., Wiley Interscience.

[82] G. Wojcik, J. J. Mould, R. Ferguson, R. Martino, and K. K. Low, "Some image modeling issues for i-line, 5X phase shifting masks," *Proc. of the SPIE*, vol. **2197**, pp. 455–465, 1994.

[83] H. Hopkins, "The concept of partial coherence in optics," *Proc.Royal Soc.London*, vol. **A217**, pp. 408–432, 1953.

[84] A. K.-K. Wong, *Rigorous Three-Dimensional Time-Domain Finite-Difference Electromagnetic Simulation*. PhD thesis, University of California at Berkeley, 1994.

[85] T. V. Pistor, A. R. Neureuther, and R. J. Socha, "Modeling oblique incidence effects in photomasks," *Proceedings of the SPIE*, vol. **4000**, pp. 228–237, 2000.

[86] T. V. Pistor, *Electromagnetic Simulation and Modeling with Applications in Lithography*. PhD thesis, University of California at Berkeley, 2001.

[87] W. Oldham, S. Nandgaonkar, A. Neureuther, and M. O'Toole, "A general simulator for VLSI lithography and etching processes: Part I–application to projection lithography.," *IEEE Trans.Electron Devices*, vol. **ED-26(4)**, pp. 717–722, April 1979.

[88] K. Toh and A. Neureuther, "Identifying and monitoring effects of lens aberrations in projection printing," *In Proc.SPIE Optical Microlithography VI*, vol. **722**, pp. 202–209, 1987.

[89] *http://cuervo.eecs.berkeley.edu/Volcano/*.

[90] *http://www.kla-tencor.com/products/bysol-litho-simulation.html*.

[91] *http://www.sigma-c.de/solid-c/solid-c.html*.

[92] K. Adam, *Domain Decomposition Methods for the Electromagnetic Simulation of Scattering from Three-Dimensional Structures with Applications in Lithography*. PhD thesis, University of California at Berkeley, 2001.

[93] E. W. Marchand and E. Wolf, "Consistent formulation of Kirchhoff's diffraction theory," *J. Opt. Soc. of Am.*, vol. **56**, pp. 1712–1722, 1966.

[94] E. Wolf and E. W. Marchand, "Comparison of the Kirchhoff and the Rayleigh-Sommerfeld theories of diffraction at an aperture," *J. Opt. Soc. of Am.*, vol. **54, No 5**, pp. 587–594, 1964.

[95] M. Totzeck, "Validity of the scalar Kirchhoff and Rayleigh-Sommerfeld diffraction theories in the near field of small phase objects," *J. Opt. Soc. Am.*, vol. **8, No. 1**, pp. 27–32, 1991.

[96] C. T. Tai, *Dyadic Green Functions in Electromagnetic Theory*, Intext Educational Publishers, 1st ed., 1971.

[97] Y. Rahmat-Samii, "Useful coordinate transformations for antenna applications," *IEEE Trans. Antennas and Propagation*, vol. **AP-27, No 10**, pp. 571–574, 1979.

[98] A. F. Peterson, S. L. Ray, and R. Mittra, eds., *Computational Methods for Electromagnetics*, Wiley-IEEE Computer Society Pr., 1997.

[99] J. J. H. Wang, *Generalized Moment Methods in Electromagnetic*, Wiley-Interscience, 1991.