Diffractive Optics for Maskless Lithography and Imaging

by

Rajesh Menon

B. E. Electrical Engineering, Nanyang Technological University (1998)
S. M. Electrical Engineering, Massachusetts Institute of Technology (2000)

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2003
May 2003

© Massachusetts Institute of Technology 2003. All rights reserved.

Author ....
Department of Electrical Engineering and Computer Science
May 21, 2003

Certified by ....
Henry I. Smith
Keithley Professor of Electrical Engineering
Thesis Supervisor

Accepted by ....
Arthur C. Smith
Chairman, Department Committee on Graduate Students

BARKER

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
AUG 0 6 2004
LIBRARIES

Diffractive Optics for Maskless Lithography and Imaging

by

Rajesh Menon

Submitted to the Department of Electrical Engineering and Computer Science
on May 21, 2003, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

Semiconductor industry has primarily been driven by the capability of lithography to pattern smaller and smaller features. However due to increasing mask costs and complexity, and increasing tool costs, the state-of-the-art technology in lithography is accessible only to a select few.

Zone-plate array lithography (ZPAL) is a novel method of maskless lithography that aims to alleviate some of these issues while offering a solution that can be extended to the limits of nanolithography. In ZPAL, an array of diffractive lenses is used to form an array of spots on the substrate. Each spot is modulated independently by means of spatial-light modulators. This essentially creates a "parallel laserwriter". In addition, this lithography system can be converted into a parallel-confocal microscope, which enables fast, high-resolution imaging.

This thesis addresses the performance of diffractive lenses, particularly high-numerical aperture zone plates for lithography and imaging using a combination of experimental and theoretical studies. A novel proximity-effect correction algorithm that was implemented effectively in a ZPAL system is also described. Variations to another diffractive lens known as the photon sieve are proposed. The first ever lithography results performed using these new elements are presented in this thesis.

Thesis Supervisor: Henry I. Smith
Title: Keithley Professor of Electrical Engineering

Acknowledgments

Firstly, I would like to thank Professor Henry Smith for giving me the opportunity to work on this project. To understand with academic rigor is one of the important lessons I have learnt from him. I thank him for giving me the freedom to direct the research based on my intellectual curiosity.

Dr. Darío Gil has been a wonderful friend throughout my stay at MIT. His perseverance, technical prowess, and intellectual conscientiousness in science as well in social issues will always be an inspiration for me. I thank Dr. David Carter for being a patient mentor, and succeeding in transforming me from a theorist into an experimentalist! Darío and David have been instrumental in the success of this research.

I have great respect for Professor George Barbastathis, who has been a champion of this work. He has been a constant source of encouragement and suggestions throughout the course of this research. I am also extremely grateful for his friendship.

Several past and present members of the NanoStructures laboratory namely Dr. Todd Hastings, Dr. Thomas Murphy, Dr. Farhan Rana, Dr. Juan Ferrera, Feng Zhang, Michael Walsh, Mark Finlayson, Minghao Qi, Tim Savas, Thomas O'Reilly, Dr. James Goodberlet, Dr. Mitchell Meinhold, Tymon Barwicz, Amil Patel, Dr. Mark Schattenburg, and Dr. Paul Konkola have helped in shaping this research project.

I am very grateful to James Daley for showing me the ropes for almost everything inside the clean-room. His patience, good-nature, and hard-work has always impressed me. Much thanks is owed to Jimmy Carter for helping me with the scanning-electron microscope.

For a long time, Mark Mondol was my only connection to India at the NanoStructures laboratory. I thank him for this and all his help with electron-beam lithography.

Professor Dennis Prather and his students at the University of Delaware were extremely helpful with several aspects of the simulation tools used in this work.

I thank Cindy Lewis for help with all administrative work during my stay at MIT.

Euclid Moon has been my officemate, and friend since I first joined the NanoStructures laboratory. I greatly apreciate our discussions on science and philosophy. I will always treasure the adventures we have shared, especially the time spent admiring music and nature

at the *Fusion Festival*!

Paul Herz was my first and truest friend at MIT. I am forever grateful to him for his support through all the changes in my life.

I would like to express appreciation to my psy-trance family including Sasha, Selim and Yannis, I will miss our "Krysalis" gatherings.

Thanks also goes to Juliet Gopinath and Joyce Wu for being my steadfast friends from the very beginning.

I would like to extend my gratitude to the Hewetts especially Mother (Sheila), Laurie, Rebekah, and Patrick for welcoming me with open arms into their family. I can hardly wait for our future family reunions!

My family has been instrumental in my success. My only sister, Rakhee and her husband, Vijay have supported me through everything. I owe all my success to my wonderful parents, who have been patient with me through my long absences, and have been supportive of all my decisions. I would like to honor my grandparents by dedicating this thesis to them.

Finally, I would like to thank my wife, Marianne. She is my best friend and my soulmate. She always brings hope into the darkest moments. I thank her for being so kind and understanding with me. I am forever grateful to her for bringing such joy and hope into my life.

Contents

List of Figures

Chapter 1

Introduction to Maskless Lithography

The rapid technological growth since the invention of integrated circuits has been driven primarily by lithography. Lithography refers to the technology used to produce patterns on a substrate such as a silicon wafer. These patterns form the building blocks of integrated circuits and subsequently of the electronic revolution of the 20th century.

There are two main flavors of lithography: mask-based and maskless.

1.1 Mask-based lithography

In mask-based lithography, the master-pattern on a mask is imaged onto the substrate. This technology is illustrated using the example of the optical stepper.

The optical stepper is the workhorse of lithography for the semiconductor industry. Figure 1-1 shows the simplified schematic of an optical stepper. The mask consists of a fused silica substrate with the pattern defined by patches of opaque material such as chromium. The mask is manufactured by means of one of the direct-write (maskless) lithography methods. The pattern on the mask is imaged using optics onto the surface of the substrate. The substrate is coated with a layer of photoresist. Photoresists are materials that undergo photochemical reactions when exposed to light. The photoresist is thus able to record the projected image. Further processing then transfers this image into the silicon wafer. One major drawback of this technique is that the pattern to be printed is fixed by the master pattern on the mask.

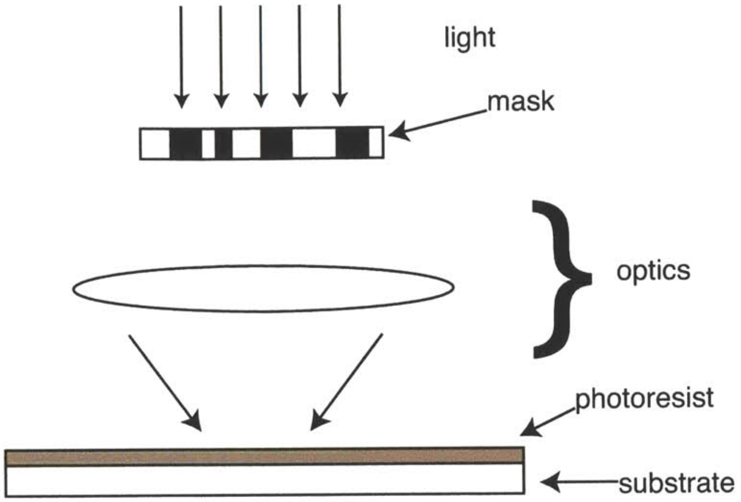


Figure 1-1: *Simplified schematic of an optical stepper. The mask contains the master pattern as depicted by the dark and white regions. The pattern on the mask is imaged onto the surface of the substrate using optics. The substrate surface has a layer of photoresist which records the image of the pattern.*

1.2 Maskless Lithography

Maskless lithography, as its name implies, needs no mask. An example will help illustrate the concept. Scanning-electron-beam lithography is the most often used maskless lithography technique. Figure 1-2 shows a simplified schematic of a scanning-electron-beam-lithography system. A beam of electrons is focused using magnetic fields into a spot. This focused spot can be deflected by means of magnetic or electric fields to *write* desired patterns directly on the resist. In this case the resist is sensitive to the electron-beam. The pattern is thus built up from discrete spots. The major difference from the mask-based approach is that in the electron-beam system, the same focused spot can be used to write any arbitrary pattern, just by changing the deflection pattern of the beam.

One common theme in maskless lithography is the fact that the pattern information, instead of being stored in the form of a physical mask (which is a very efficient but expensive method of storage; more on that later), is stored in the form of digital information in a hard

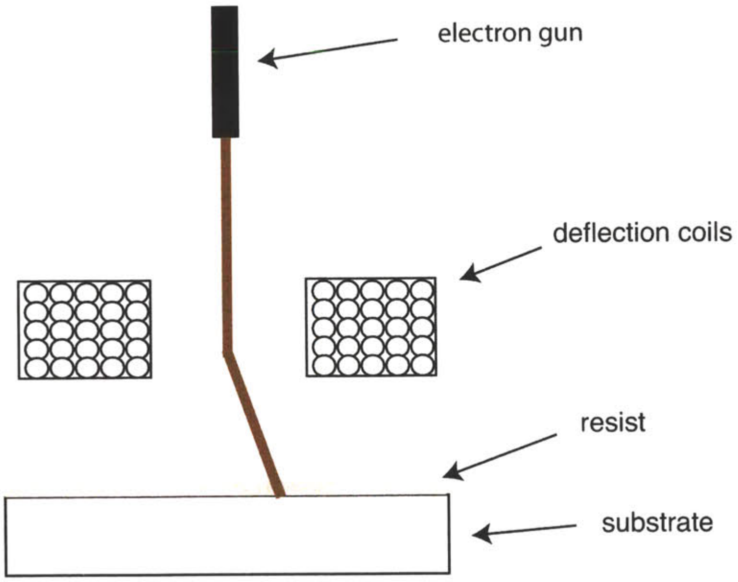


Figure 1-2: *Simplified schematic of a scanning-electron-beam-lithography tool. The electron gun emits a beam of electrons which are focussed by magnetic lenses. The deflection coils create fields which move the focused electron beam on the substrate. Electron-sensitive resist is coated on the surface of the substrate. Thus, arbitrary patterns are printed spot by spot.*

drive. This enables the quick modification of such patterns in software, when necessary.

1.3 The Case for Maskless Lithography

The reasons for considering maskless lithography as a viable manufacturing alternative are threefold:

1. Complexity and Cost of Mask-based tools.
2. Complexity and Cost of Masks.
3. Lack of flexibility of mask-based tools.

Complexity and Cost of Mask-based tools

Lithography tools for the semiconductor industry have developed in an evolutionary fashion driven by the ever-increasing demand for smaller linewidths and more features on printed patterns. The principle of operation of lithography tools has changed very little since their first implementation. The level of complexity of these tools has, however increased dramatically. The tendency has been to *squeeze more out* of each generation of tools in terms of the sizes of printed features, the number of pixels per field, and speed of the production process. This has made leading-edge tools extremely expensive (see figure 1-3).

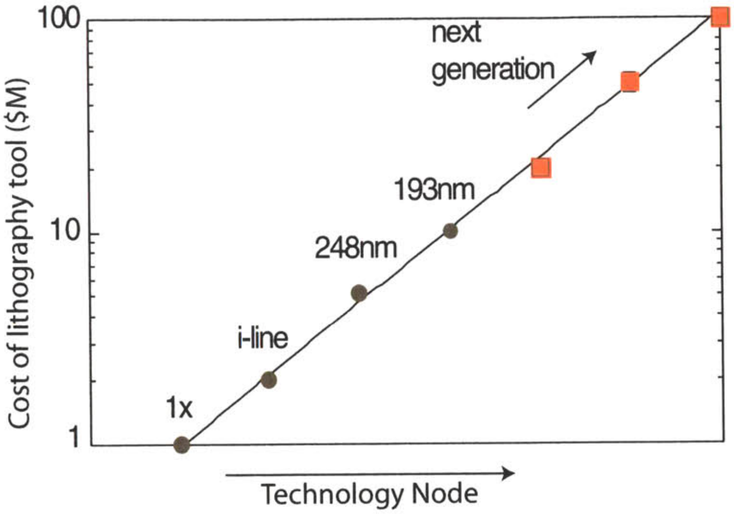


Figure 1-3: *Escalating cost trend of optical lithography tools. The cost for optical lithography tools has increased in an exponential manner with every successive generation. Note that the cost of a single tool is slated to hit $100 Million within 3 generations [1].*

This exponential increase in the cost of the tool is mostly due to the increase in its complexity. In order to understand the reasons for this increase, we start with the well-known equation relating the exposure wavelength to the minimum feature size.

$$minimum\ feature\ size \propto \frac{Wavelength}{Numerical Aperture} \tag{1.1}$$

The numerical aperture is a measure of the resolution of the optical system (we will discuss this in more detail in later chapters). Thus, the minimum feature size (MFS) decreases with decreasing wavelength and increasing numerical aperture (NA). The trend in the semiconductor industry has been to increase the NA of the system, while also increasing the field size until it becomes cost-effective to decrease the wavelength. This is because a change in the wavelength requires appropriate changes in the materials for the optical system and materials for the photoresist. Hence, changing the wavelength is more complicated than changing the NA of the optical system. This is true until the limit of NA of the system is reached. This is illustrated in figure 1-4, where both the MFS and the wavelength are plotted for various generations of lithography tools. Note that for the DUV (248 nm) generations, the MFS decreases mainly due to increase in the NA of the system.

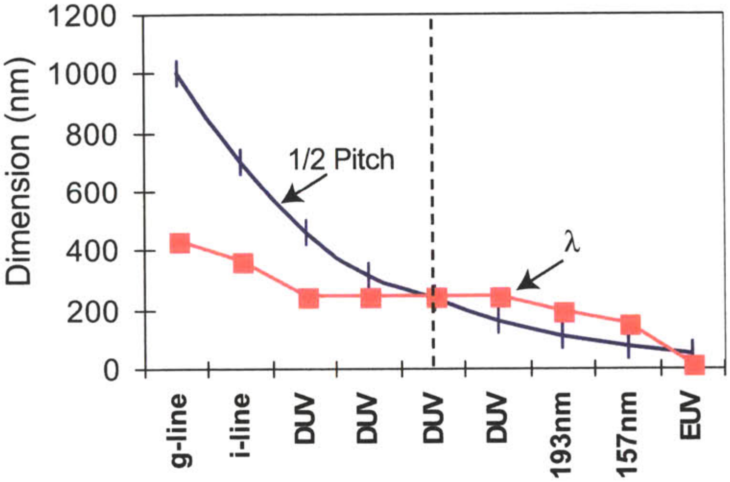


Figure 1-4: *Trend towards subwavelength lithography. The minimum feature size and the exposure wavelength are both plotted for lithography tools from various generations. Note that to the right of the vertical dotted line, the smallest printed features are subwavelength. This adds significant complexity to the optical projection system. Data from the Intel Lithography roadmap [2].*

Thus, the requirements of increasing NA and decreasing wavelength for smaller printed features has led to exponential increases in the complexity and cost of the lithography tools. The semiconductor industry has evolved in such a manner that the high-volume output of such expensive and complicated tools justify their existence. However, the explosive growth in custom-designed integrated circuits and the increasing prevalence of nanotechnology in emerging sciences such as nanoelectromechanical systems (NEMS), integrated optics, biomedical applications etc. poses a clear need for simpler, less-expensive lithography tools.

Complexity and Cost of Masks

The cost of masks for optical steppers has increased significantly as well. Figure 1-5 shows the relative cost of a typical mask for various generations. Note that the cost increases in an exponential fashion, very similar to the increase in the cost of the tool.

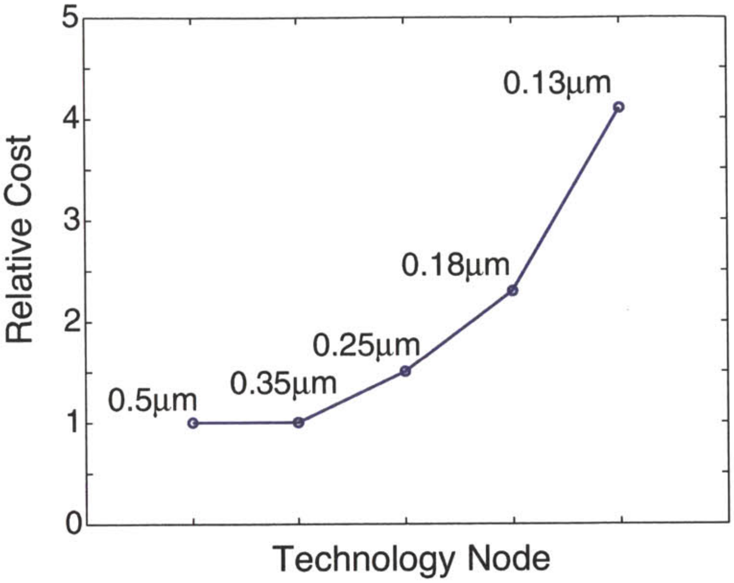


Figure 1-5: *Escalating cost of masks. The relative cost of a typical mask is plotted for several lithography generations. The trend shows an exponential increase in the cost of masks. It is projected that the cost of an EUV mask-set could be as much as $1 million [3].*

This huge increase in the cost of the mask can be attributed to four main factors.

1. *Writing*

   The mask is typically manufactured using a direct-write method such as scanning-electron-beam lithography. The cost of an electron-beam lithography tool is very high (more than $10 million). In addition, such electron-beam tools are very slow and it takes more than 30 hours to write a typical mask. The high capital cost and the long turn-around time contributes to very high mask costs.

2. *Inspection and Repair*

   For advanced masks, the presence of defects on the mask is detrimental to the printing of good patterns. Hence, all such masks need to be inspected several times during their manufacture. Inspection tools for advanced masks are extremely expensive (more than $10 million). If defects are present, it is relatively cheaper to fix these errors than to rewrite the entire mask. However, these processes of inspection and repair tend to be very slow, adding to the final cost of the mask.

3. *Complicated Optical Proximity Correction*

   In order to print features much below the exposure wavelength, the master pattern on the mask needs to incorporate optical proximity correction (OPC). OPC works by adding assist features to the master pattern. These features compensate for corner rounding, line shortening and other image imperfections introduced by the subwavelength features. Moreover, these assist features themselves are generally as small as the smallest printed features. This entails significant cost and time investment in preprocessing of the pattern and fabrication of the mask. Due to the small feature sizes on the mask, the yield decreases as well. Hence, the need for OPC drives up the cost of the mask significantly.

4. *Phase Shifting Masks*

   Phase shift works in such a manner as to increase the contrast between light and dark regions of the printed image. This causes the images printed on the photoresist to have sharper edges. Since this involves an extra processing step, it increases the cost of the mask. Moreover, this adds significant complexity to the inspection-repair cycle.

Decreasing feature sizes on the wafer has led to extreme resolution-enhancement techniques to be utilized on the mask. This is the primary reason for the exponential increase in the cost of masks. Except in the case of large volume manufacturing, these high mask costs cannot be sustained. Getting rid of the masks altogether is the natural alternative for low-volume manufacturing.

Lack of Flexibility

The cost and complexity factors discussed above create an enormous barrier to using conventional optical projection lithography for prototyping purposes. The reality of a **fixed** pattern on the mask and the high costs of obtaining new masks serve as a strong impediment to experimentation during device development. This situation is compounded by a steeper learning curve, inaccurate simulation tools and increased number of design iterations for advanced chip prototyping. In short, this *lack of flexibility* leads to over-conservative designs and an ineffficient use of available resources.

Device development cycles are naturally longer in emerging markets of nanotechnology and microelectromechanical systems (MEMS). Thus, a major hurdle for many nanotechnology and MEMs device manufacturers is to find a lithography solution that meets both technology and production requirements. Since nanotechnology (excluding the semiconductor industry) is a nascent industry, high tool and mask costs cannot be supported. Moreover, nanotechnology requires higher developmental effort resulting in higher prototyping requirements. It is highly inefficient and expensive to use lithography technology developed for the mature semiconductor industry for the nanotechnology space. Maskless lithography of various forms could very well be the choice of lithography for such new devices.

1.4 Forms of Maskless Lithography

A myriad of ideas for maskless lithography have been proposed. Almost all the proposals can be divided into two types. The first uses a variable mask and images the mask-pattern onto the substrate. Since the pattern on the mask is variable, arbitrary patterns can be printed. The second forms a focused spot on the substrate and writes arbitrary patterns by scanning this spot or/and the substrate. In this section, we briefly analyze these two types.

Variable Mask Technology

A schematic of a system using variable mask technology is shown in figure 1-6. The system is identical to the optical stepper (figure 1-1) except for the variable mask. The pattern on the variable mask is controlled by digital data from a computer. Arbitrary patterns can be printed by imaging the pattern on the mask onto the wafer.

In general, the variable mask is a spatial light modulator. It is a programmable device comprised of a number of pixels, whose light transmissions can be individually controlled.

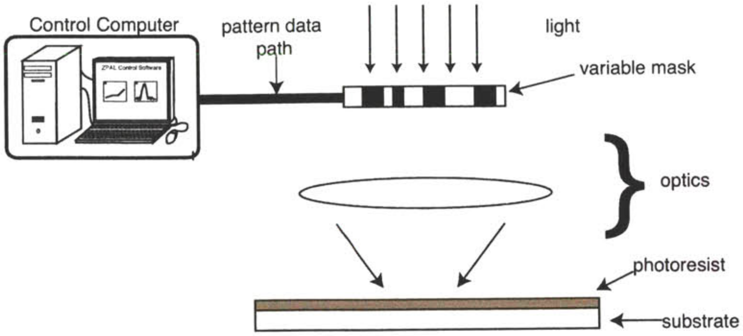


Figure 1-6: *Schematic of a system using variable masks. The system is the same as an optical stepper except for the variable mask. The pattern on the mask can be changed by the computer. There is typically a digital data path between the computer and the hardware for the variable mask.*

There are various forms of this variable mask. An example will illuminate this idea. A maskless lithography tool, currently being commercialized by Micronic Systems employs such a variable mask [4]. In their tool, the variable mask is used in a reflection configuration as opposed to the transmission configuration depicted in figure 1-6. The variable mask in their tool consists of an array of moveable micromirrors. A simplified schematic of the Micronic system is shown in figure 1-7(a). Light from a laser is directed onto the variable mask as shown. The variable mask is an array of micromirrors which can tilt about one axis. Since the micromirrors are close packed, appropriate tilts in the micromirrors can produce grating patterns on the mask. Light reflecting from these grating patterns gets redirected into diffraction orders as shown in figure 1-7(b). By controlling the tilt of the micromirrors on the variable mask, the intensity of the light directed into the zero order can be modulated. Only the zeroth diffraction order is allowed to pass through the Fourier filter and form pixels of the image on the substrate. Thus variable image formation is achieved through diffractive-spatial modulation of light. A scanning electron micrograph of the micromirror array is shown in figure 1-7(c). By scanning the substrate on a stage and by exposing overlapping fields, the entire substrate can be patterned accurately.

The optical projection system in the Micronic tool is very similar to that of a stepper. However, the size of the image projected onto the substrate is smaller than that used in an optical stepper. Larger images are built up by scanning the substrate and "stitching"

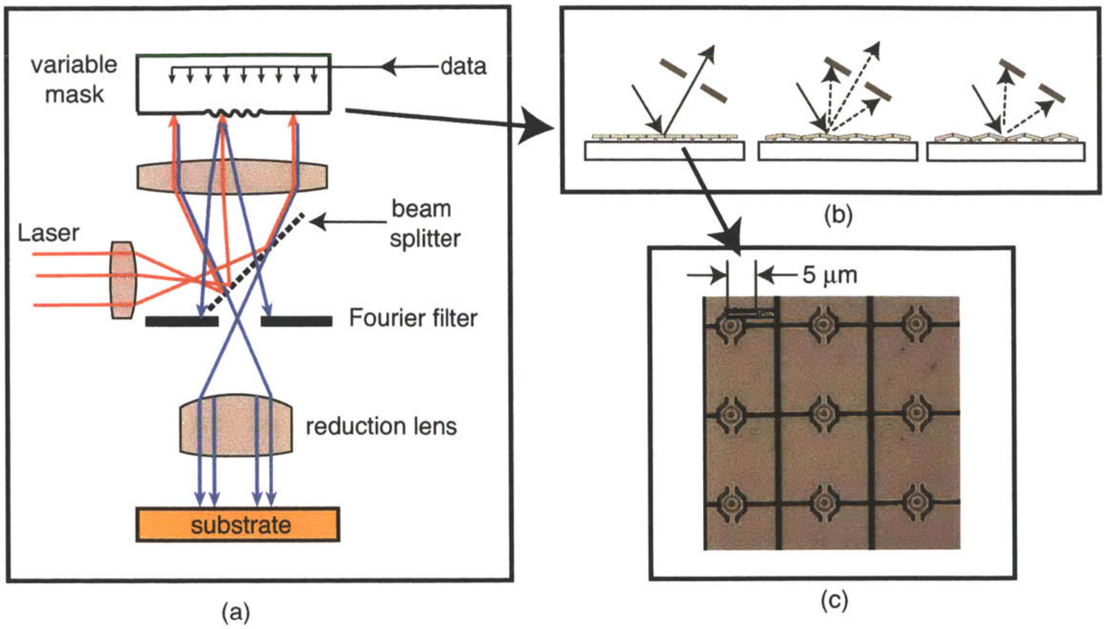


Figure 1-7: *Schematic of the Micronic Systems maskless lithography tool. (a) Full schematic. (b) Working principle of variable mask. The mask is an array of micromirrors, which can tilt about one axis to create a diffraction grating pattern on the mask. Light reflecting from this grating pattern gets redirected into diffraction orders. The height of the grating can be controlled by the tilt of the micromirrors and diffraction into the zero order can be controlled. Only the zeroth diffraction order is allowed to propagate past the Fourier filter and form pixels of the image on the substrate. (c) Scanning electron micrograph of the micromirror array. Note that the micromirrors are close packed and they can tilt about the vertical axis.*

the smaller images. Off-axis illumination techniques can be used in this case to improve the resolution of printed features. However, the efficacy of these techniques tend to be very pattern dependent. Changes in the off-axis illumination require significant changes to the system. Therefore, off-axis illumination techniques effectively defeat the purpose of being able to print arbitrary patterns easily. Other resolution-enhancement techniques such as phase-shifting cannot be realized in such systems in a straightforward manner.

Variable mask systems are an evolutionary step towards maskless lithography from optical projection lithography. These have promise where maskless lithography is important. However, several key challenges still remain to be solved.

Scanning Spot Technology

As the name implies, this involves scanning a focused spot on the substrate to *write* arbitrary patterns. The substrate may also be scanned while the spot is held stationary or, both the substrate and the spot may be scanned. One example of such a system is the scanning-electron-beam-lithography system. The pattern to be printed is broken up into *dots*. The focused spot is then scanned over the location of the *dots* to build up the pattern *dot* by *dot*. This is illustrated in figure 1-8.

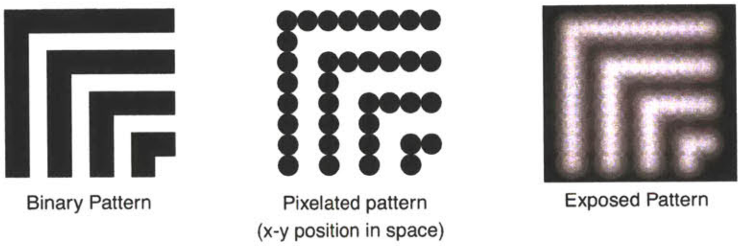


Figure 1-8: *Scanning Spot Maskless Lithography. (a) The binary pattern is drawn using CAD software. (b) This binary pattern is broken up into dots. These dots represent the positions in space where the focused spot must deliver dose into the resist on the substrate. (c) A fictitious final image in resist is depicted.*

This is the most general form of maskless lithography in the sense that any arbitrary pattern can be composed of *dots*, the only limitation being the size of the *dot* itself. One can envision using resolution-enhancement techniques to improve the size and quality of the spot. This improved spot can then be used to *write* arbitrary patterns.

In general, there are several appealing features of a scanning-spot system. However, using a single spot to *write* an entire pattern makes this a very slow process. This is the single most important disadvantage of such a system. In this thesis, we explore a maskless lithography technique which aims to overcome this limitation by using a large number of focused spots to *write* in parallel. This technique is called Zone-Plate-Array Lithography (ZPAL).

Chapter 2

Introduction to Zone-Plate-Array Lithography

Maskless lithography has grown in significance recently as research and industry leaders recognize its potential impact. Zone-Plate-Array Lithography (ZPAL) is a scanning-spot maskless lithography technique developed at the Nanostructures Laboratory at MIT. In this chapter, we provide a system-level description of ZPAL. The important components of the system are described as well.

2.1 System Description

ZPAL is a maskless lithography technique which uses an array of zone plates to create an array of spots on the substrate. A zone plate is a diffractive lens that focuses incident light into a spot. The light incident on each zone plate is controlled by a multiplexing element such as a spatial light modulator, where one pixel of the modulator controls the light incident on one zone plate. By scanning the substrate on a stage and properly timing the light modulation on each zone plate, arbitrary patterns are printed in a "dot-matrix" fashion. A schematic of the system is shown in figure 2-1.

Patterning using ZPAL involves breaking a large pattern into smaller "unit-cells". A unit-cell is that portion of the pattern, which one zone plate is responsible for writing. The

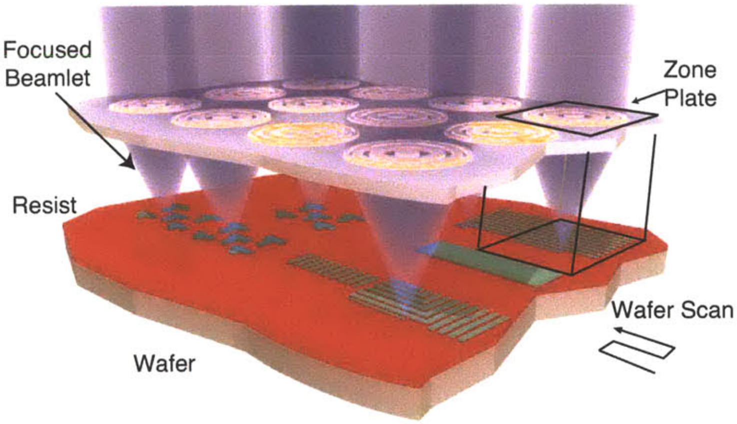


Figure 2-1: *Schematic of a Zone-Plate-Array Lithography system. Light is incident on an array of zone plates. These zone plates, being diffractive lenses, focus light into an array of spots on the substrate. A multiplexing element (not shown here) controls light incident on each zone plate. By scanning the substrate on a stage, arbitrary patterns are printed in a "dot-matrix" fashion.*

unit-cell covers an area on the substrate directly beneath one zone plate[1]. At any instant, a zone plate writes one spot. Hence, within each unit-cell, the pattern is built up "spot-by-spot" by scanning the substrate. As will be shown later, each such "spot" can be very small and hence, high resolution patterning is possible. Moreover, since all the unit-cells are patterned at the same time, the trade-off between field-size and resolution is overcome. Thus, ZPAL combines the advantages of optical maskless lithography with the high throughput of a parallel-beam system.

The major components of a ZPAL system are depicted in figure 2-2. Light from a source is "cleaned" using a spatial filter and collimated before being directed onto the multiplexing device. Light from the device is then imaged onto the zone-plate array. Since the period of the pixels on the multiplexing device may be different from the period of the zone plates in the zone-plate array, magnifying optics may be required. The zone-plate array forms an array of spots on the substrate, which sits on a scanning stage. A central computer is responsible for synchronization of the motion of the stage and the pattern on the multiplexer. A prototype ZPAL system was built in the course of this research. The major components

---

[1] This is true when the zone plates are tight packed. In general, the unit-cell is the total area of the array divided by the number of zone plates.

of this system are discussed in detail next.

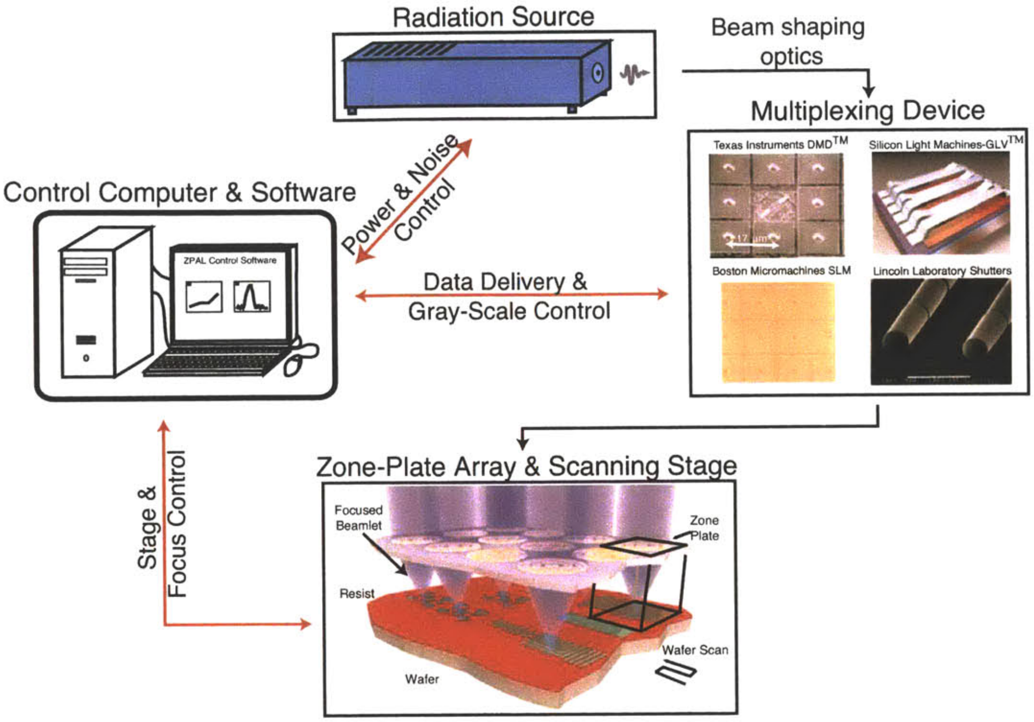


Figure 2-2: *Schematic of the components of a ZPAL system.*

2.1.1 Radiation Source

The wavelength of the radiation source is the defining factor for resolution in ZPAL. Since diffractive lenses exhibit strong chromatic aberrations, it is important that the source has sufficiently small bandwidth. The effect of the source bandwidth on lithographic patterning is discussed in chapter ???. The power of the source is one of the factors affecting throughput of the lithography system. Light from the source should be collimated in order to attain diffraction-limited focusing of the zone plates.

A helium-cadmium gas laser at a wavelength of 442 nm was initially used in the ZPAL system at MIT. Currently, the system uses a gallium-nitride diode laser at a wavelength of 400 nm. This laser is very compact (38mm by 144mm) and has very quick startup times compared to traditional gas lasers. In addition, the diode laser has an excellent temperature coefficient (quoted value is 0.05 nm/°C).

Figure 2-3: *Photograph of the diode laser. This is the radiation source used in the ZPAL prototype built during the course of this research. This laser diode has a central wavelength of 400 nm and a output power of 25 mW. This laser module was manufactured by Power Technology, Inc [5].*

2.1.2 Micromechanical Multiplexers

In ZPAL, the micromechanical multiplexer is a device comprising several pixels, such that one pixel controls the light incident on one zone plate. The rate at which all the pixels can be refreshed is an important factor in determining the patterning speed of the system. There is also an associated data delivery subsystem, which transfers the pattern data from the control computer to the multiplexing element. The rate of this data transfer is another factor affecting the patterning speed. In order to obtain good pattern fidelity, the intensity of light incident on each zone plate must be variable; this is called *grayscaling*. If real-time grayscaling is not possible, multiple passes over the same area of the pattern are required. This again affects the patterning speed of the system. In addition, the multiplexers must be compatible with the wavelength of the source. In this section, we briefly describe several candidates for the multiplexing element in ZPAL.

Texas Instruments Digital Micromirror Device (TI DMD$^{\text{TM}}$)

The TI DMD$^{\text{TM}}$ is an array of micromirrors [6]. Each micromirror can be independently activated by electrostatic forces. A scanning-electron micrograph of 9 micromirrors of this device is shown in figure 2-4(a). Each micromirror has electronics built directly beneath it. When not actuated, the micromirror is mechanically centered. When a voltage is applied between one electrode and the micromirror, electrostatic force causes the micromirror to tilt.

The schematic in figure 2-4(b) shows the configuration in which this device can be used

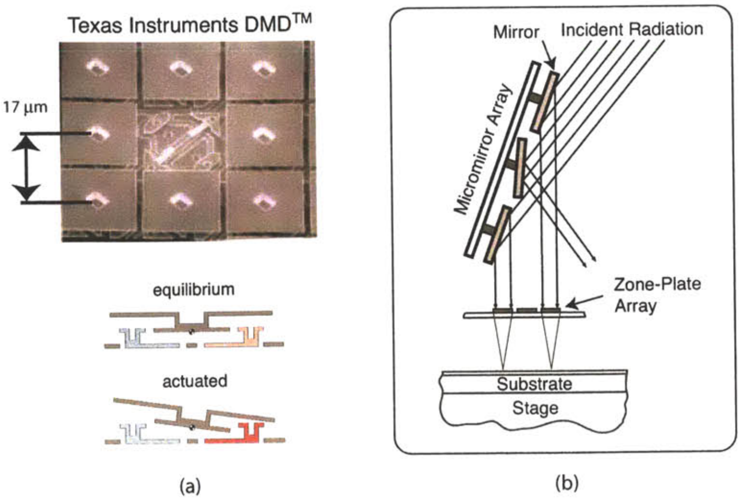


Figure 2-4: *Texas Instruments Digital Micromirror Array. (a) Scanning electron micrograph of the TI DMD*$^{\text{TM}}$*. Note that the central mirror is removed to show the structure underneath. The mirror is hinged about one diagonal and is actuated electrostatically as shown in the bottom. (b) Schematic of the TI DMD*$^{\text{TM}}$ *being used in ZPAL. When the micromirror is not actuated, light is directed into the zone plate. Since the pitch of the micromirrors and that of the zone plates may be different, a telescopic system is required to transfer light from the micromirror array to the zone-plate array (not shown in this schematic).*

in ZPAL. When the micromirror is in the "ON" position, light is directed into the zone plate. When the micromirror is in the "OFF" position, light is directed away from the zone plate. In this manner, "ON-OFF" states can be achieved. Grayscale control is achieved by using the micromirror in a dynamic mode. The micromirror is driven by a pulsed signal which alternates its state between "ON" and "OFF". By changing the duty cycle of the signal, the ratio of time the micromirror stays "ON" to the time it stays "OFF" can be controlled. This, in turn, provides grayscale control of the light incident on the zone plate. This form of "time-multiplexing" grayscale control, however requires more time to pattern every spot and hence decreases patterning speed.

The TI micromirrors were initially used in the ZPAL system because of their early commercial availability. Figure 2-5 show scanning electron micrographs of patterns written using the TI DMD$^{\text{TM}}$ in a ZPAL system.

The TI device is packaged in a hermetically sealed glass case. This package is almost opaque to all wavelengths below about 300 nm. This is a significant barrier to scaling the

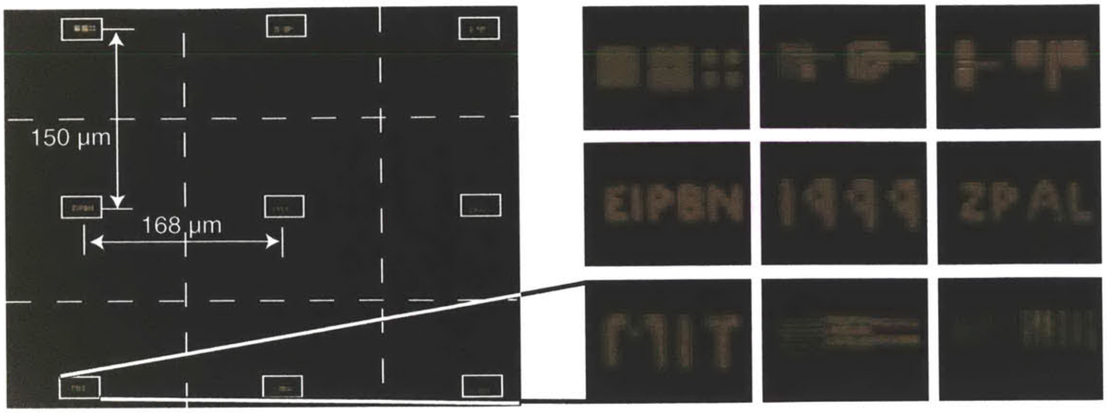


Figure 2-5: *Patterns written in parallel using the TI DMD*$^{\text{TM}}$*. Scanning electron micrographs of patterns written in photoresist after development in a ZPAL system at* $\lambda = 442\,nm$*, and zone plate NA=0.7. This experiment was performed with a 3X3 array of zone plates. Hence 9 different patterns were written simultaneously.*

system to lower wavelengths. In addition, the rate at which all the pixels could be refreshed is limited to about 10 kHz. As mentioned earlier, in order to attain diffraction-limited focusing, the wavefront incident on the zone plate must be a uniform plane wave. Since there is a small indent at the center of each micromirror[2], it is not very clear what the wavefront incident on the zone plate might look like. Aberrations in this wavefront will translate into blurring of the focused spot.

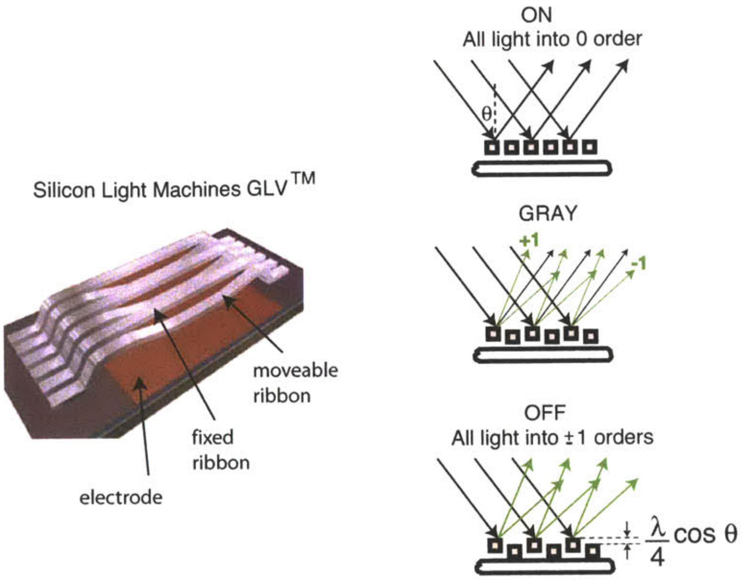
Silicon Light Machines Grating Light Valve (SLM GLV$^{\text{TM}}$)

The SLM GLV$^{\text{TM}}$ consists of a linear array of ribbons. Alternate ribbons are movable and can be electrostatically actuated. A group of 6 ribbons are controlled in unison to form a single pixel. A schematic of this device is shown in figure 2-6(a).

The working principle of the device is illustrated in figure 2-6(b). When the ribbons are not actuated, light reflects in a specular manner. When alternate ribbons are pulled down, light sees a reflection grating and hence, diffracts into the ±1 and the 0 orders. The amount of light into each order depends on the height of this grating. When this height corresponds to a multiple of $\lambda/4\cos(\theta)$, then the light reflecting off the actuated ribbon and

---

[2] This is due to the post that attaches the micromirror to the substrate.



(a) (b)

Figure 2-6: *Silicon Light Machines Grating Light Valve. (a) Schematic of the* GLV$^{\text{TM}}$*. The device consists of a row of ribbons, of which a group of 6 ribbons composes one pixel (shown here). Alternate ribbons are moveable and can be electrostatically pulled down. (b) Cross-section through one pixel of the* GLV$^{\text{TM}}$*. Top: The ribbons are not actuated. Incident light reflects totally into the 0 order. Center: Alternate ribbons are pulled down slightly. Incident light sees a reflection grating and some light gets redirected into the* ± 1 *orders. Bottom: Alternate ribbons are pulled down to completely extinguish the 0 order. All light is diffracted into the* ± 1 *orders.*

that reflecting of the fixed ribbon interfere destructively into the 0 order. Hence, all the light is switched into the ±1 orders. Thus, "ON-OFF" states and any states in between can be achieved or, in other words, the grayscaling capability is built into the device. This avoids multiple passes and hence, makes the patterning faster.

Figure 2-7 shows a simple schematic illustrating one configuration using the GLV as the multiplexing device in ZPAL. One pixel on the GLV controls the light incident on one zone plate. In this configuration, the 0 order light is used to illuminate the zone plate. The light efficiency is high but the "ON-OFF" contrast may not be very high due to spurious reflections which scatter light into the 0 order. On the other hand, if the +1 or −1 order is used to illuminate the zone plate, the "ON-OFF" contrast is maximized, but the light efficiency is lowered. There are optical techniques to collect both the +1 and the -1 orders to increase the light efficiency.

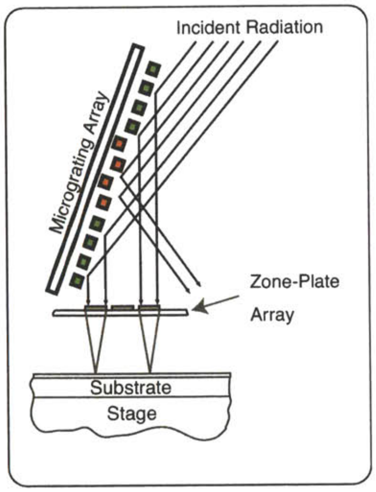
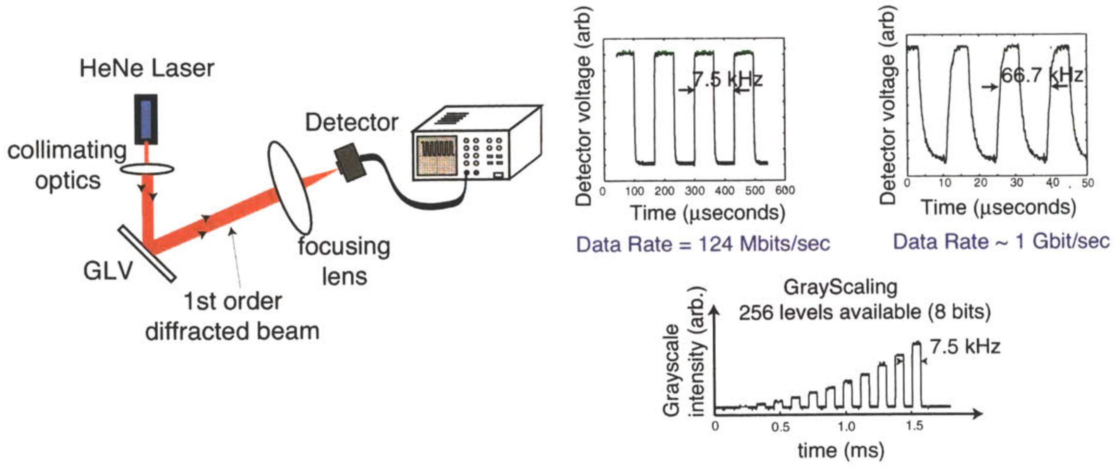


Figure 2-7: *GLV$^{\mathrm{TM}}$ in ZPAL. One pixel on the GLV$^{\mathrm{TM}}$ controls light incident on one zone plate. In this configuration, the 0 order from the GLV$^{\mathrm{TM}}$ is used to illuminate the zone plate. The higher diffraction orders are directed away by means of a fourier transform lens and a spatial filter, which are not shown in the figure.*

Since the working principle of the GLV is based on diffraction, it is possible to use this device at shorter wavelengths. The actuation of the device involves a small displacement of the ribbons and therefore, the refresh rate of all the pixels can be quite high. This was experimentally verified in our laboratory. The experimental setup is illustrated in figure 2-8(a). A helium-neon laser ($\lambda = 633\,\mathrm{nm}$) was used to illuminate the device. First order light reflected off the device was collected by a lens and directed onto a detector. All the pixels on the device were turned “ON” and “OFF” at different frequencies and the signal on the detector was measured. The results are shown in figure 2-8(b). The top row is a plot of the detector signal as a function of time. The device worked reliably up to a refresh frequency of 67 kHz (beyond which the detector was unable to reliably measure the modulated optical signal). The plot on the bottom row indicates that real-time grayscaling is indeed possible at high speeds.

Since the GLV is a linear device, the number of spots projected onto the substrate is lower than when using a 2-D device such as a TI DMD$^{\mathrm{TM}}$. This adversely affects the patterning



(a) (b)

Figure 2-8: *Experimental verification of the performance of the GLV*$^{\text{TM}}$. *(a) Experimental setup. The data delivery system from the computer to the GLV is not shown here. (b) Top row: Optical signal modulated by the GLV at 7.5 kHz and 66.7 kHz. Bottom: Optical signal modulated in time and intensity (grayscaled) by the GLV.*

speed due to the increased number of "turns" of the stage.[3]

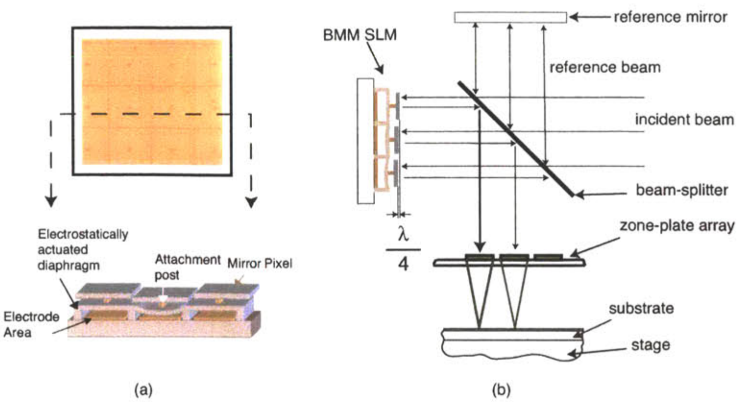
Boston MicroMachines Spatial Light Modulator (BMM SLM)

The BMM SLM is an array of micromirrors which can move up and down in a piston fashion. An optical micrograph and a schematic of the device are shown in figure 2-9(a). Each micromirror is mechanically attached to a diaphragm by a post. The diaphragm and the substrate form a pair of electrostatic parallel-plate actuators. This device was invented at Boston University [7] and is currently being commercialized by Boston Micromachines [8].

This device is employed in an intereferometric manner in order to accomplish spatial light modulation. Light reflected from the device is interfered with a reference beam. If a micromirror is pulled down by $\lambda/4$, light reflected off it will destructively interfere with its corresponding portion of the reference beam, producing a dark pixel. By pulling the micromirror down by heights between 0 and $\lambda/4$, grayscale modulation of the pixel can be achieved. Thus, this device provides grayscaling while writing, which avoids multiple

---

[3] Most of the time is spent changing directions (or turning) while scanning. More turns are needed for a 1-D array of spots as compared to that for a 2-D array of spots.



(a) (b)

Figure 2-9: *Boston Micromachines SLM. (a) Top: Optical micrograph of a portion of the device. Bottom: Schematic cross-section through the dashed line. Each mirror can be electrostatically actuated by means of electrodes below. (b)Working principle of the BMM SLM in ZPAL. Light reflected from the micromirrors is interfered with a reference beam as shown. When a mirror is pulled down by $\lambda/4$, the reflected light is phase shifted by π with respect to the reference beam and that portion of the beam shows a dark pixel. One micromirror controls light incident on one zone plate. By controlling the heights to which the micromirror is pulled down, the intensity of light incident on the zone plate can be grayscaled.*

writing passes. One pixel (micromirror) illuminates one zone plate. This working principle is illustrated in figure 2-9(b).

Since the on-axis reflected light is used in this configuration, spurious reflections from the device could reduce contrast between the "ON" and "OFF" states. Nevertheless, this is an attractive and simple solution if large-pixel-count devices at high refresh rates become commercially avaliable.

Microshutters

All the multiplexers discussed so far are reflective or diffractive in their function. These devices scatter light and therefore, reduce contrast between the "ON" and the "OFF" states. Moreover, they will not operate at shorter wavelengths such as EUV (13 nm) or X-ray, since materials with sufficient reflectivity are not readily available[4]. In order to address these disadvantages a multiplexer that works by transmission is required.

---

[4] Although glancing angle reflection is still possible at these wavelengths, this presents challenging design issues for the system.

Such a device, comprising an array of microshutters, is being developed for astronomical applications at the NASA Goddard Space Flight Center [9]. Each pixel on this device consists of a silicon nitride membrane shutter attached to a frame by a torsion beam. The shutter is capable of opening 90 degrees out of the plane of the device by rotating about the axis of the torsion beam. Scanning electron micrographs of this device are shown in figure 2-10.

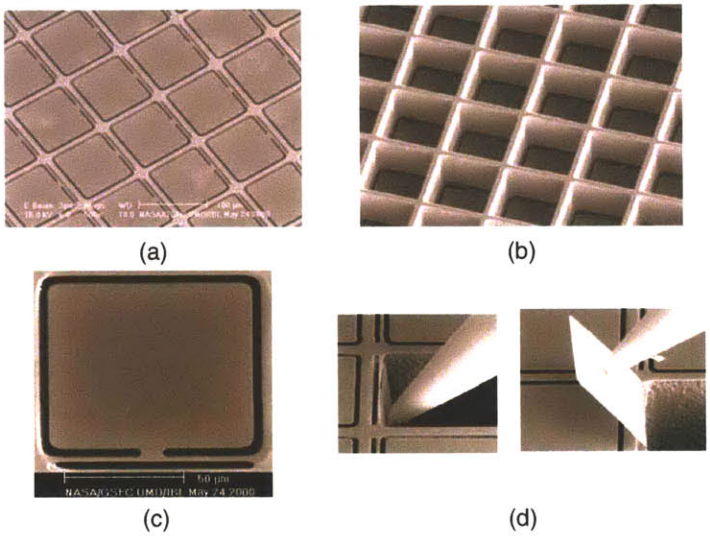


Figure 2-10: *Microshutter arrays being developed at NASA. (a) Scanning electron micrograph of an array of microshutters. (b) Backside of the array showing the support structures. (c) Single microshutter. (d) The microshutter being actuated manually using a micromanipulator. The shutter hinges about the torsional hinge and can rotate 90 degrees out of the plane of the array. Note that each microshutter is about* $100\mu m \times 100\mu m$ *in dimension.*

Two methods of actuation for this device has been proposed. The first method uses a combination of magnetic and electrostatic methods. A moving electromagnet beneath the microshutter array produces a force on a magnetic pad on the microshutter causing it to bend out of plane. An electrode on the frame of the microshutter, when actuated holds the microshutter open. The microshutter array is placed in close proximity to a transparent substrate which has an array of electrodes aligned to each microshutter. Voltage applied to an electrode holds the corresponding microshutter closed while the magnet is scanned. The operating sequence is illustrated in figure 2-11 (a). The second method involves mechancal and electrostatic actuation with a double shutter mechanism. Two arrays of microshutters are aligned to each other and brought in close contact. Microshutters on both arrays have

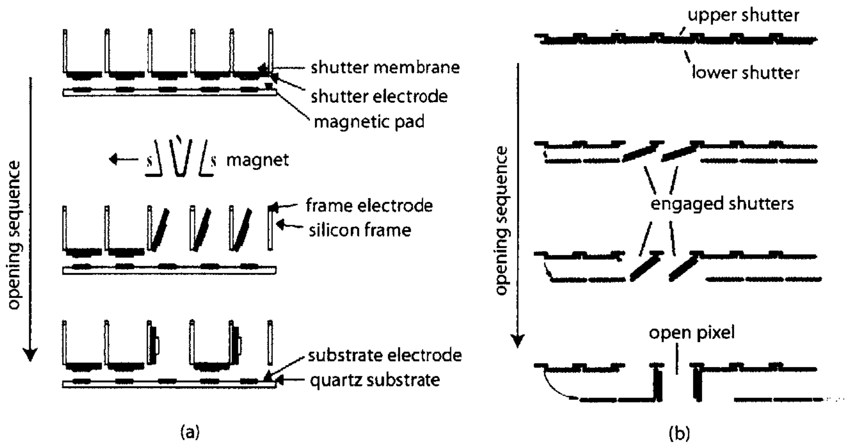


Figure 2-11: *Microshutter actuation methods. (a) Magnetic and electrostatic actuation. As the magnet is scanned from left to right underneath the device, the microshutter experiences a force out of the plane of the device. If the microshutter is selected to be opened, a voltage on the frame electrode aids the magnetic force and holds the microshutter open. The microshutter can be held closed by a voltage applied to the substrate electrode. (b) Mechanical and electrostatic actuation. Two microshutter arrays are aligned and brought in contact. The lower pixel is moved as shown. When a microshutter pair (lower and upper) is engaged by applying a voltage between their electrodes, the pixel is opened.*

electrodes. A voltage is applied to engage the upper and lower microshutters. The lower array is driven by a precision stage in a direction as shown in figure 2-11 (b). This motion causes the engaged pixels to open.

A different microshutter array device invented at the MIT Lincoln Laboratories and commercialized by Axsun Technologies, Inc. [10] is shown in figure 2-12. An actuating signal causes a thin, reflective, and conductive film to unfurl itself over a hole in the substrate by electrostatic attraction. When the signal is not applied, the stresses left in the film cause it to roll itself back up.

The main disadvantage of a microshutter array is the fact that real-time grayscaling is not possible. Hence, multiple passes during writing would be needed. This will increase the patterning time and reduce throughput. The frame rate of both the microshutter devices is quite low. In addition, the devices presented here are in the early stages of their development. However, this technology is very promising to meet the multiplexing needs for ZPAL

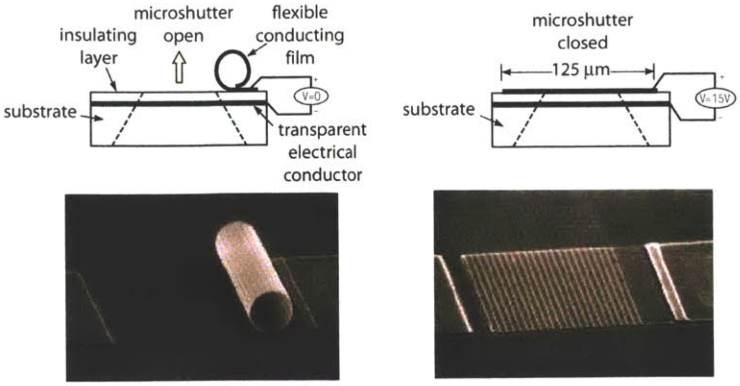


Figure 2-12: *Microshutter array device invented at MIT Lincoln Laboratories. Top left: Schematic cross-section through a single microshutter. When not actuated, stresses in the microshutter film cause it to roll itself up and the hole in the substrate is open. Top right: The film microshutter is electrostatically actuated to unroll and cover the hole. Bottom row: Scanning electron micrographs of the microshutter in open (left) and closed (right) positions. Each microshutter is* $125\mu m \times 125\mu m$ *in dimension.*

especially at shorter wavelengths.

2.1.3 Zone-Plate Array

The zone-plate array is perhaps the most novel and important component of the entire system. Hence, we devote all of next chapter to discuss it.

2.1.4 Scanning Stage

As mentioned earlier, patterns are written in ZPAL by scanning the substrate on a stage. Hence, very high precision scanning stages are necessary to ensure high quality patterns. The speed of the stage is an important factor affecting the throughput of the system. Control software synchronizes the stage with the multiplexers in order to produce the desired pattern on the substrate. To obtain smooth lines, and good linewidth and edge-placement control, the stage placement has to be controlled to a small fraction of the smallest feature in the pattern.

In the laboratory implementation of ZPAL, we use a "piezo flexure nanopositioning scanner" from Physik Instrumente (model P-770) [11]. Since the stage uses piezoelectric drives, less than 10 nm resolution in stage position is possible with closed loop control. The stage also has an X-Y scanning range of $200\mu m \times 200\mu m$. A photograph of the stage is shown in figure 2-13.

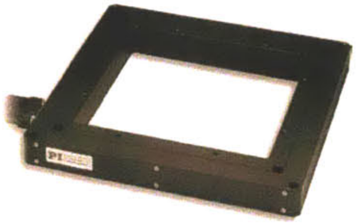


Figure 2-13: *Scanning stage from Physik Instrumente. This is an X-Y scanning stage with a clear aperture equipped piezielectric drives. Closed loop operation provides a scan range of* $200\mu m \times 200\mu m$ *and position resolution less than 10 nm.*

There is a second stage which maintains the gap between the zone-plate array and the substrate. It is important to hold this gap steady while the pattern is being written. This is achieved using a novel device from New Focus, Inc. called the picomotor$^{\mathrm{TM}}$. The picomotor$^{\mathrm{TM}}$ is a piezoelectric actuator which turns a screw. This device uses static friction to turn and hold the screw. Thus, this device "holds" the gap between the zone-plate array and the substrate. The operation relies on the difference between static and dynamic friction, and provides placement resolution less than 30 nm. As will be explained in a later chapter, we use a confocal signal from the zone plates to set the desired gap.

2.2 Advantages of ZPAL

Zone-plate-array lithography is not only a novel concept, it is also the result of the simultaneous maturing of several sister technologies including microelectromechanical devices (for multiplexing), computer (for data processing and control), high speed data channels, and developments of microfabrication techniques (for fabrication of zone-plate arrays). Within maskless technologies, ZPAL enjoys many important advantages.

ZPAL is a massively parallel, maskless optical lithography technique. Being an optical system avoids the obvious disadvantages of charged particles such as sensitivity to stray and inter-particle fields. Moreover, photoresist technology for optical lithography is a very mature field. A large number of zone plates writing at the same time increases the writing speed tremendously.

ZPAL is perhaps the only maskless technology that is scaleable to shorter wavelengths with minimal changes in its core architecture. Since zone plates can be designed to focus any wavelength, if suitable sources and multiplexers are available, there is a very clear path of wavelength-scaling in ZPAL.

ZPAL employs a large array of zone plates for patterning. Zone plates being single-level diffractive lenses can be fabricated using reliable planar fabrication techniques. It is important to note that zone plates at extremely high numerical apertures (greater than 0.9) can be fabricated. This is a far-cry from refractive lenses, which require very tedious and careful grinding and polishing steps. Since zone plates are used on-axis, image aberrations which arise in projection systems are totally avoided as well.

In the next chapter, we will discuss all aspects of zone plates including their fabrication and focusing performance.

Chapter 3

Zone Plates for Lithography

In this chapter, the working principle of a zone-plate lens is described. Different methods of simulating diffraction by the zone plates are discussed next. Arguments for the necessity of vector diffraction models to accurately simulate high-NA zone plates are provided. Finally, fabrication techniques to manufacture zone-plate arrays are also explained.

3.1 Zone-plate as a Diffractive Lens

A zone plate is a diffractive element which focuses an incident plane wave to a point. Hence, it is a diffractive lens. The zone plate consists of concentric-circular gratings, whose period decreases with increasing radius, such that the first order of diffraction converges at the same point on the optical axis. This point is called the *focus* of the zone plate.

The zones of the zone plate are arranged such that alternate zones have the same optical path length to the focus and hence, interfere constructively at the focus. When alternate zones are blocked, as shown in figure 3-1(a), a portion of the transmitted light is focussed into a spot. This implementation is called an *amplitude* zone plate. Only about half of the incident light is transmitted through the amplitude zone plate. It will be shown shortly that about 10% of the incident light is focussed into the focal spot of the first order. The opaque zones are out of phase from the transparent zones by half a wavelength. So, it was proposed [12] that these opaque zones also be made transparent, but with a phase shift of π, which corresponds to half a wavelength, as shown in figure 3-1(b). Then almost all the incident light is transmitted and a larger portion of the light interferes constructively at the focus. In this *phase* zone plate, about 40% of the incident light is focussed into the first order.

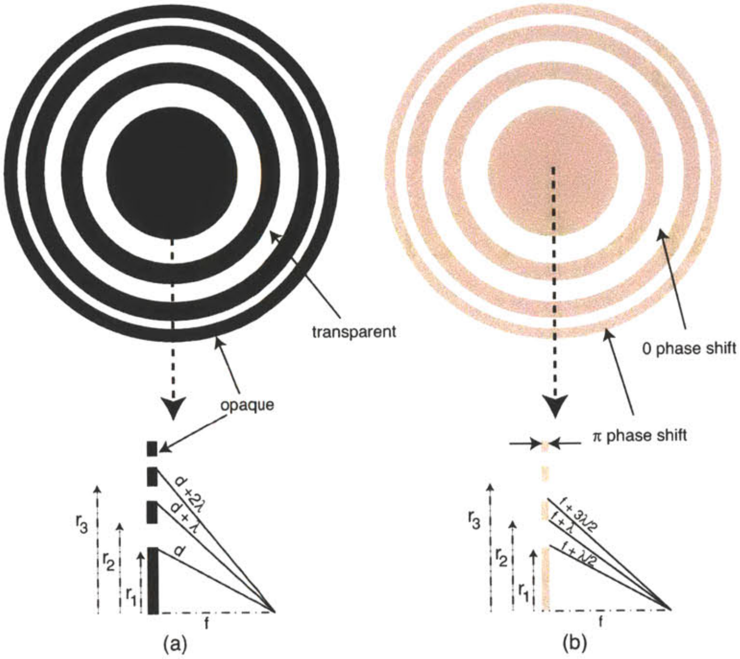


Figure 3-1: *Schematic of a zone plate. Top: Top-view of a zone plate. Note that the zones form a circular grating whose period decreases with increasing radius. Bottom: (a) Cross-section through an amplitude zone plate. Alternate zones are transparent (in white) and light from these zones interferes constructively at the focus. Opaque zones are shown in black. (b) Cross-section through a phase zone plate. Alternate zones are phase shifted by π (shown in gray) and light from all zones interferes constructively at the focus. Zones in white are transparent with no phase shift.*

3.1.1 Design of a Zone Plate

For both amplitude and phase zone plates, the expression for the radii of the zones can be derived from the contraints imposed on the optical-path lengths to the first-order focus, and by using the Pythagorean theorem [13]. The radius of the n^{th} zone is given by

$$r_n^2 + f^2 = (f + n\frac{\lambda}{2})^2 \tag{3.1}$$

$$r_n = \sqrt{n\lambda f + (n\frac{\lambda}{2})^2} \tag{3.2}$$

This equation defines the geometry of the zone plate.

The zone plate design begins with the choice of the wavelength, the numerical aperture (NA) and the focal length. From these 3 parameters, one can calculate the number of zones as follows. The numerical aperture, NA is a measure of the size of the focussed spot produced by the zone-plate lens. NA is defined as the sin of the half-angle, θ subtended by the zone plate at the focus (see figure 3-2).

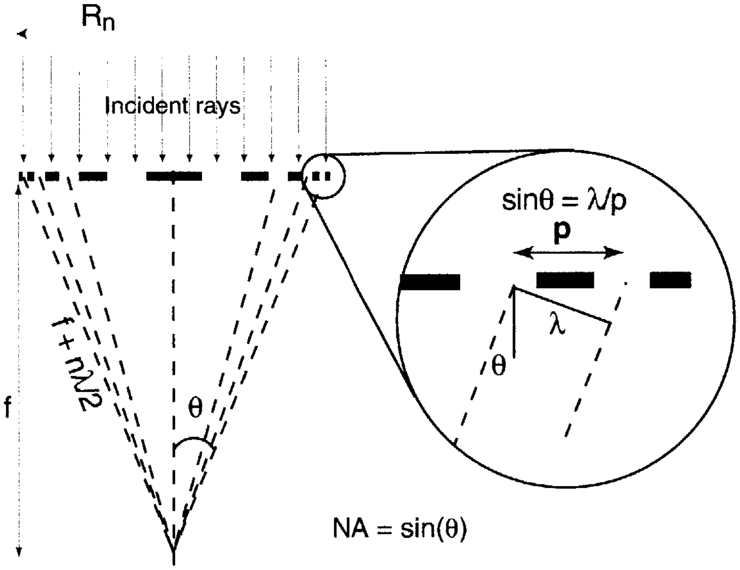


Figure 3-2: *Numerical Aperture of a zone plate. The schematic shows the cross-section of an amplitude zone plate (black zones are opaque). Normally incident rays are diffracted into the first order. The inset shows diffraction from the local grating formed by the outermost zones. The diffraction angle, θ depends on the local period of the outermost zones, P, and the wavelength of light, λ. This angle of diffraction, θ defines the numerical aperture of the zone plate.*

Figure 3-2 shows a ray diagram of light diffracting into the first order, when a normal-plane wave is incident on an amplitude zone plate. The inset shows the outermost zones of the zone plate. This circular grating diffracts light at an angle, θ. Since the path-length difference between the two rays (from the outermost and next to outermost transparent zones) is λ, the diffraction angle, θ is related to the period, P by

$$NA = \sin(\theta) = \frac{\lambda}{P} \tag{3.3}$$

This indicates that the numerical aperture of the zone plate is related to the size of the smallest zone[1]. It will be shown later that the size of the focussed spot is approximately equal to the size of the smallest zone. The angle, θ can be related to the radius of the outermost zone, R_N:

$$\tan(\theta) = \frac{R_N}{f} = \frac{\sqrt{N\lambda f + (\frac{N\lambda}{2})^2}}{f} \tag{3.4}$$

Solving for N and substituting $NA = sin(\theta)$, we get

$$N = \{\frac{2f}{\lambda}(\sqrt{1 + \tan^2(\sin^{-1}(NA))} - 1)\} \tag{3.5}$$

where $\{\}$ denotes that the value is rounded to the nearest integer. Once N is calculated, the zone radii can be calculated by equation (3.2). This completes the design procedure for the zone plate.

3.1.2 Diffraction Efficiency

Since the zone plate is a diffractive lens, incident light energy is redirected into several diffraction orders just as with a diffraction grating. In this section, we carry out a simple analysis to estimate the fraction of incident light redirected into the various orders. To simplify the mathematics, we begin with this approximation of equation (3.2).

$$r_n = \sqrt{n\lambda f} \tag{3.6}$$

---

[1] The smallest zone is also the outermost zone from equation (3.2).

This approximation is valid when the following condition is met

$$n << \frac{4f}{\lambda} \tag{3.7}$$

This condition is met for zone plates of long focal lengths and small NAs. However, even for zone plates of high NA, this condition is typically met for several of the inner zones. The presence of the term quadratic in N in equation (3.2) will modify the efficiency values somewhat. We will discuss this subject later in this chapter.

Equation (3.6) implies that r_n^2 is a periodic function with a period of $2\lambda f$. This periodicity causes incident light to be diffracted into various orders. In order to analyze the distribution of energy into the various diffraction orders, we study the far-field diffraction pattern of a zone plate that is illuminated by a normal-plane wave. The far-field diffraction pattern is given by the Fourier transform of the transmission function of the zone plate[14]. This transmission function, T is periodic in r^2 and can be expressed as a fourier series in r^2. With a coordinate transformation, $u = r^2$,

$$T(u) = \sum_{m=-\infty}^{\infty} c_m e^{i2m\pi \frac{u}{2\lambda f}} \tag{3.8}$$

The Fourier transform of a periodic function is a sum of delta functions. Each delta function represents one diffraction order and the corresponding efficiency is given by the square of the amplitude of the delta function. By fourier transforming equation (3.8), it can be shown that the amplitudes of the delta functions are the corresponding fourier-series coefficients. Thus, the diffraction efficiency into the $\mathrm{m^{th}}$ diffraction order is given by $|c_m|^2$[15].

For an amplitude zone plate, the coefficients c_m are given by

$$c_m = \begin{cases} \frac{1}{2} & m = 0 \\ \frac{1}{m\pi} & m \text{ odd} \\ 0 & m \text{ even} \end{cases} \tag{3.9}$$

Thus, the amplitude zone plate diffracts 25% of the incident light into the 0 order. Half of the incident energy is absorbed by the opaque zones. In the approximation of equation (3.6), the areas of all zones are equal, and half the area of the zone plate is opaque. The remaining 25% of the incident light is diffracted into the odd orders. About 10% of the incident light is diffracted into the first order. Even orders are absent due to the symmetry of the zones (50% duty cycle).

In the case of a phase zone plate, the coefficients are given by

$$c_m = \begin{cases} 0 & m = 0, \text{ even} \\ \frac{2}{m\pi} & m \text{ odd} \end{cases} \tag{3.10}$$

Since the area of the zones which are phase shifted is equal to the area of the zones which are not phase shifted, the 0 order is cancelled. The diffraction efficiency into the first order is about 40%.

The transformation, $u = r^2$ in the above analysis changes the structure of the diffracted orders from plane waves (such as in simple diffraction gratings) into spherical waves focussed at different points along the optical axis. Figure 3-3 shows the diffraction orders of a zone plate. Note that the negative orders are diverging spherical waves (virtual foci) and the positive orders are converging spherical waves (real foci). For a given focal plane such as the first order, other diffraction orders contribute to a background intensity.

In the above calculations, we assumed N, the number of zones to be infinite. For finite N, the Fourier series is multiplied by a window function, defining the aperture of the zone plate. This implies that the sum of delta functions in the fourier domain is convolved with a *sinc* function[2]. In this case, the diffracted spherical waves will be aberrated. It must be noted that the above calculations are valid only for thin, low NA zone plates. At high NAs and especially for the outer zones, the condition in equation (3.7) is not satisfied. In that case, the areas of all zones are no longer equal. Then, we would not expect the 0 order for a phase zone plate to be totally cancelled. As will be described later in this chapter, high NA

---

[2] A sinc function is the Fourier transform of a rectangular window function given by $a\frac{\sin\frac{\pi v a}{2}}{\frac{\pi v a}{2}}$, where $2a$ is the size of the window.

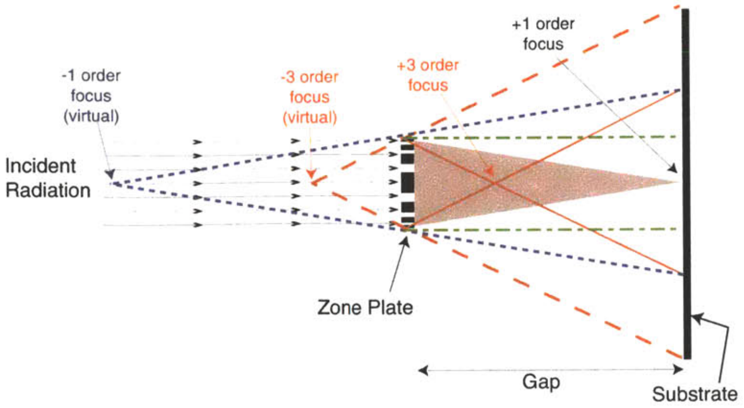


Figure 3-3: *Diffraction orders of a zone plate. Normally-incident-plane wave is diffracted by the zone plate into the zero order and odd orders. In a phase zone plate, there is no energy diffracted into the zero order.*

and finite thickness of the zones lead to effects which tend to reduce the overall diffraction efficiency.

3.2 Simulation of Zone Plates

In the previous section, the diffraction efficiency of a zone plate into various orders was calculated. In order to quantify the focusing performance of zone plates, the spatial structure of the focused spot has to be calculated as well. This requires the use of diffraction theories. These theories can be broadly divided into scalar and vector. Although electromagnetic fields are vector in nature, it has been shown that scalar diffraction theories are adequate for many cases. In this section, we begin by presenting the Fresnel-Kirchoff diffraction theory as the scalar diffraction theory of choice for simulating the focusing performance of zone plates.

3.2.1 Fresnel-Kirchoff Diffraction Theory

The fundamental principle behind scalar diffraction is Huygens' principle which states that every point on a wavefront acts as a secondary source, emitting spherical wavelets. The envelope of all these secondary wavelets constitute the propagating wavefront. This principle is illustrated in figure 3-4. This forms the qualitative basis of wave propagation.

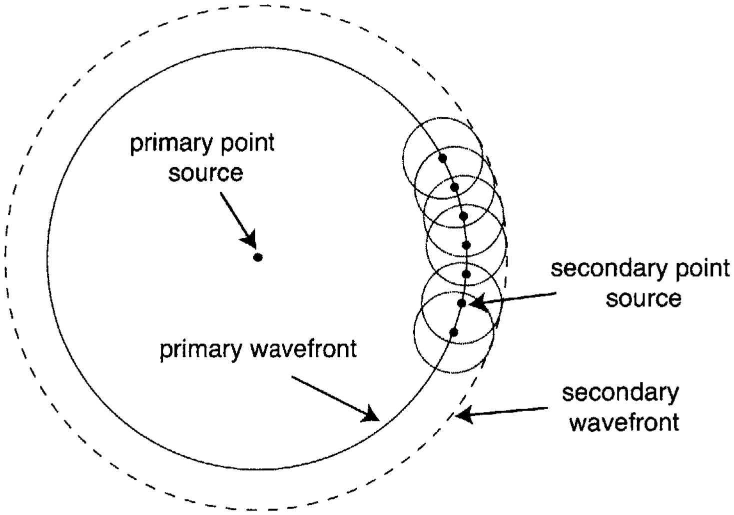


Figure 3-4: *Huygens' Principle. The primary point source produces the primary-spherical wavefront. Each point on this wavefront acts as a secondary-point source, producing secondary-spherical wavelets. The envelope of all these wavelets constitutes the secondary wavefront.*

Optical wave propagation is further governed by the Helmholtz wave equation, which in the scalar regime is given by

$$\nabla^2 u - \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} = 0 \tag{3.11}$$

where u is the scalar field, t is time, c is the speed of light, and

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{3.12}$$

Using the wave equation and Huygens' principle, one can derive the field after diffraction by a general aperture. This derivation is carried out in Appendix A. In the Fresnel-Kirchoff formulation, the field at a point P_1 due to a point source at P_0 in the presence of an aperture Σ (see figure 3-5) is given by

$$U(P_0) = \frac{A}{i\lambda}\iint_{\Sigma} \frac{e^{ik(r_{21}+r_{01})}}{r_{21} \times r_{01}}\left[\frac{\cos(\vec{n},\vec{r_{01}}) - \cos(\vec{n},\vec{r_{21}})}{2}\right]d\Sigma \tag{3.13}$$

where P_1 is a point on the aperture. Equation (3.13) is the summation of the secondary-point

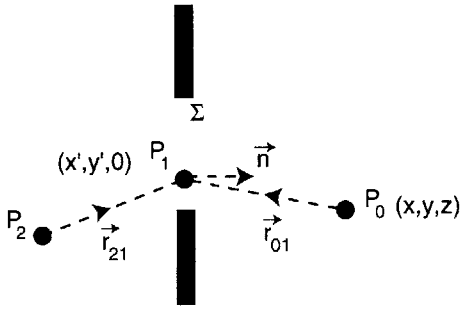


Figure 3-5: *Fresnel-Kirchoff formulation for diffraction from a general aperture.*

sources on the wavefront incident on the aperture Σ.

The main assumption in this formulation is that the presence of the aperture does not perturb the field or its derivative. Hence, the zone plate can be decomposed into the constitutive apertures (circular zones) and their field contributions calculated separately. Since, the wave equation is linear, the field at any point beyond the zone plate is the sum of the fields diffracted from the constitutive apertures. In other words, we can replace the simple aperture Σ in equation (3.13) with the zone-plate-transmission function. If T is the zone-plate-transmission function,

$$U(P_0) = \iint_{-\infty}^{\infty} T \times U(P_1)\frac{e^{ikr_{01}}}{r_{01}}d\Sigma \tag{3.14}$$

Thus equation (3.14) denotes a general expression for calculating diffraction by a zone plate using the Fresnel-Kirchoff formulation. For a normally-incident-plane wave, equation (3.13) can be further simplified as:

$$U(P_1) = \frac{1}{i\lambda}[\frac{\cos(\vec{n}, \vec{r_{01}}) - 1}{2}] \tag{3.15}$$

The numerical implementation of equation (3.15) was the subject of earlier work[16]. Using cartesian coordinates, (x, y, z) for P_0 and $(x', y', 0)$ for P_1, we can rewrite equation (3.14) as

$$U(P_0) = \frac{1}{2i\lambda}\iint_{-\infty}^{\infty} T(x', y') \times H(x - x', y - y')dx'dy' \tag{3.16}$$

where

$$H(x-x',y-y') = \frac{e^{ik\sqrt{(x-x')^2+(y-y')^2+z^2}}}{\sqrt{(x-x')^2+(y-y')^2+z^2}}(1+\frac{z}{\sqrt{(x-x')^2+(y-y')^2+z^2}}) \quad (3.17)$$

Equation (3.16) is the convolution of the zone-plate-transmission function and the Fresnel-Kirchhoff diffraction kernel, H.

$$U(P_0) = \frac{1}{2i\lambda}T \otimes H \quad (3.18)$$

This formulation of the fresnel-kirchoff diffraction integral was implemented in software. The results of simulations using this software was validated against known solutions. This is described in detail in appendix A.

In order to understand the scope of the above formulation, the approximations leading to the Fresnel-Kirchoff formulation are listed below.

1. The presence of the aperture does not affect the field or its derivative.

2. The field and its derivative are both 0 in the shadow of the aperture.

Both these conditions are not exactly true at the edges of the aperture. When light encounters sharp transitions such as at the edges of the zones, it scatters into the shadow of the aperture. Sharp transitions at the zone boundaries represent subwavelength features and scalar theory is not sufficient to accurately model their effect. In addition, scalar theory fails to account for the polarization state of light. At high numerical apertures, zone plates have more subwavelength zones and these can be fully taken into account using a vector formulation of diffraction. In the next section, we describe a method of calculating electromagnetic fields in a rigorous manner and their applications to simulating zone-plate diffraction.

3.2.2 Vector Diffraction Theory

The behavior of electromagnetic fields in all media is governed by Maxwell's equations. These equations in differential form are listed below.

Faraday's Law

$$\frac{\partial \vec{B}}{\partial t} = -\vec{\nabla} \times \vec{E} - \vec{M} \tag{3.19}$$

Ampere's Law

$$\frac{\partial \vec{D}}{\partial t} = \vec{\nabla} \times \vec{H} - \vec{J} \tag{3.20}$$

Gauss' law for the electric field

$$\vec{\nabla} \cdot \vec{D} = 0 \tag{3.21}$$

Gauss' law for the magnetic field

$$\vec{\nabla} \cdot \vec{B} = 0 \tag{3.22}$$

The symbols and their MKS units are listed below.

$\vec{E}$: Electric field (volts/meter)

$\vec{D}$: Electric flux density (coulombs/meter^2)

$\vec{H}$: Magnetic field (amperes/meter)

$\vec{B}$: Magnetic flux density (webers/meter^2)

$\vec{J}$: Electric current density (amperes/meter^2)

$\vec{M}$: Equivalent magnetic current density (volts/meter^2)

Since all the media considered in this research are linear, isotropic and nondispersive, we can write the following constitutive relations.

$$\vec{D} = \epsilon \vec{E} \tag{3.23}$$

$$\vec{B} = \mu \vec{H} \tag{3.24}$$

where ϵ is the electric permittivity (in farads/meter) and μ is the magnetic permeability (in henrys/meter). In materials which have isotropic and non-dispersive electric and magnetic losses, we can write constitutive relations for the current densities.

$$\vec{J} = \sigma \vec{E} \tag{3.25}$$

$$\vec{M} = \sigma^* \vec{H} \tag{3.26}$$

where σ and σ^* are the electric conductivity (in siemens/meter) and equivalent magnetic loss (in ohms/meter) respectively. These losses attenuate electric and magnetic fields, and convert energy to heat. By substituting equations (3.23) and (3.25) into (3.19), and equations (3.24) and (3.26) into (3.20), we get

$$\frac{\partial \vec{H}}{\partial t} = -\frac{1}{\mu} \vec{\nabla} \times \vec{E} - \frac{1}{\mu} \vec{M} \tag{3.27}$$

$$\frac{\partial \vec{E}}{\partial t} = \frac{1}{\epsilon} \vec{\nabla} \times \vec{H} - \frac{1}{\epsilon} \vec{J} \tag{3.28}$$

Writing out the components of the above vector equations we get

$$\frac{\partial H_x}{\partial t} = \frac{1}{\mu}[\frac{\partial E_y}{\partial z} - \frac{\partial E_z}{\partial y} - \sigma^* H_x] \tag{3.29}$$

$$\frac{\partial H_y}{\partial t} = \frac{1}{\mu}[\frac{\partial E_z}{\partial x} - \frac{\partial E_x}{\partial z} - \sigma^* H_y] \tag{3.30}$$

$$\frac{\partial H_z}{\partial t} = \frac{1}{\mu}[\frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial x} - \sigma^* H_z] \tag{3.31}$$

$$\frac{\partial E_x}{\partial t} = \frac{1}{\epsilon}[\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} - \sigma E_z] \tag{3.32}$$

$$\frac{\partial E_y}{\partial t} = \frac{1}{\epsilon}[\frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} - \sigma E_y] \tag{3.33}$$

$$\frac{\partial E_z}{\partial t} = \frac{1}{\epsilon}[\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} - \sigma E_z] \tag{3.34}$$

This set of six equations form the basis of the finite-difference time-domain (FDTD) method. This method is a numerical technique to solve for the components of the electromagnetic field, given by equations (3.29) to (3.34). The partial derivatives in equations (3.29)-(3.34) are approximated by finite differences. This implies that space and time have to be discretized. However, this discretization must be performed in a careful manner so as to satisfy all the Faraday's and Ampere's laws at all points. A Cartesian finite-difference-time domain (FDTD) grid with a cubic-space lattice is shown in figure 3-6. The figure also shows the position of the $\vec{E}$ and $\vec{H}$ fields in 3-dimensional space. For every $\vec{E}$ field component, the corresponding $\vec{H}$ components form a magnetic current loop and vice-versa. This implies that in the FDTD method, both the local differential and the macroscopic integral forms of Maxwell's equations are calculated simultaneously [17].

In this method, the electric and magnetic fields are calculated at the same time. In each time step, there are two sub-steps. In the first time-sub step, the $\vec{E}$ fields are calculated and stored in memory based upon previously calculated $\vec{H}$ fields. In the second time-sub step, the $\vec{H}$ fields are updated using the newly computed $\vec{E}$ fields. This procedure is repeated until the time-stepping is stopped. This, so called *leapfrog computation* of $\vec{E}$ and $\vec{H}$ fields in time and space is illustrated in figure 3-7. A one dimensional wave propagation example is

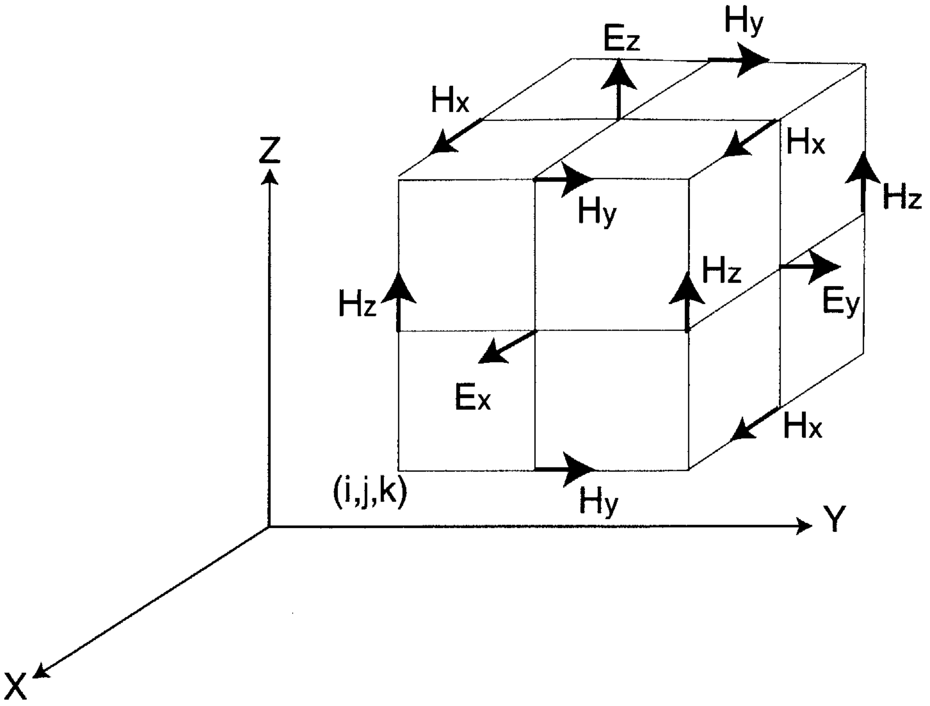


Figure 3-6: *Cartesian grid for FDTD. A cubic cell of the 3-D FDTD grid is shown. Note that each $\vec{E}$ field component is surrounded by $\vec{H}$ field components and vice-versa.*

shown in this figure. At each time-sub step, t, the $\vec{E}$ field at any point is calculated based on the $\vec{H}$ fields of the neighboring points and the previous time-sub step, t-0.5Δt. In the succeeding time-sub step, t+0.5Δt, the $\vec{H}$ field at any point is calculated based on the $\vec{E}$ fields at the neighboring points and the preceding time-sub step, t. This illustrates the idea of *central differences* in space and leapfrogging in time.

Using central difference in space (to the second order) and leapfrogging in time, the three-dimensional FDTD equation for the x component of the $\vec{E}$ field (from equation (3.34)) at the cell, (i, j, k) (see figure 3-6) and the time step, n is given by

$$\begin{gathered}\frac{E_x|^{n+1/2}_{i,j+1/2,k+1/2} - E_x|^{n-1/2}_{i,j+1/2,k+1/2}}{\Delta t} = \\ \frac{1}{\epsilon_{i,j+1/2,k+1/2}}\left(\frac{H_z|^n_{i,j+1,k+1/2} - H_z|^n_{i,j,k+1/2}}{\Delta y} - \frac{H_y|^n_{i,j+1/2,k+1} - H_y|^n_{i,j+1/2,k}}{\Delta z}\right),\end{gathered} \tag{3.35}$$

where $(\Delta x, \Delta y, \Delta z)$ are the cell dimensions and Δt is the time step. This can easily be

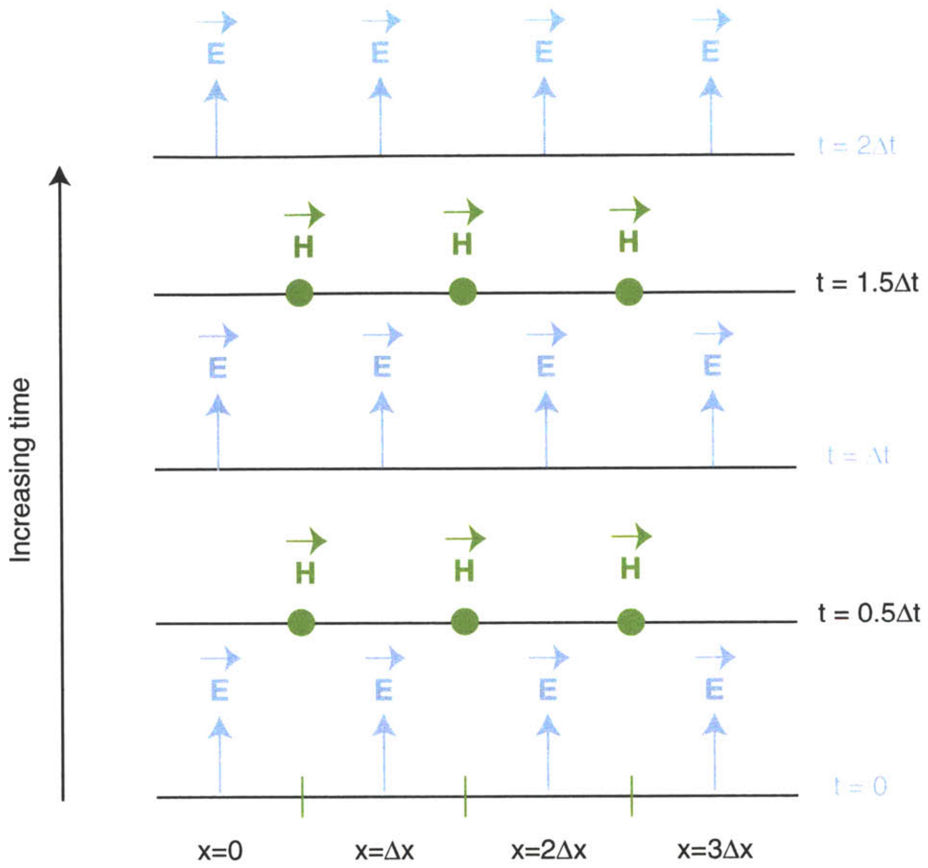


Figure 3-7: *Spacetime diagram of 1-D wave solution using FDTD. 1-D wave propagation is used to illustrate the use of central differences for space and leapfrog stepping for time.*

extended to the other field components. Thus, all the fields in space can be computed by stepping through time. The criterion for stopping the time-stepping depends on the particular situation. In the case of continuous illumination by a monochromatic-plane wave, the time-stepping stops when fields in all space become time harmonic (assuming linear media). In the case of a pulsed illumination, the time-stepping is stopped when the transient fields due to the pulse dies down. A powerful aspect of the FDTD method is that since time-stepping is explicit, it is possible to study the time-evolution of electromagnetic fields.

Since FDTD formulations involve the discretization of space, boundaries which simulate unbounded space are necessary for many practical situations. This is also important to keep the number of FDTD cells (and hence, the memory used) manageable. An artificial absorbing medium was proposed by Berenger [18], which could perfectly absorb plane waves of all angles of incidence, polarization and frequency. Such a medium was termed the *perfectly matched layer (PML)*. Thus, the boundaries of the computational domain were padded with a PML medium (which is a fictitious mathematical construct) whose thickness can be adjusted to almost perfectly simulate unbounded space.

The computational burden of the FDTD method is mainly governed by the number of cells, N used in the computation. Since the electromagnetic field components are updated at each lattice cell at every instant of time, an order(N) computation is implied. Moreover, the six field components at all the lattice cells have to be stored in memory. Thus in many cases, the hardware memory limitations pose the biggest barrier to very large scale FDTD computation. The time steps, t_{max} required for a stable solution is another factor which affects the computational burden of the FDTD method. The simulation must proceed through enough time steps such that an electromagnetic wave can causally connect every point of the structure of interest. Thus, t_{max} is a function of the size of the computational domain and in turn, of the number of lattice cells, N. This is approximately an order($N^{\frac{1}{3}}$) computational process[19]. In addition, there must be more time steps to account for the transient behavior of the fields for pulsed sources. Other algorithms complimentary to FDTD, for example progressive mesh refinement and numerical-propagation error compensation can further increase the overall computational burden of the FDTD method. These would typically contribute by increasing the number of lattice cells. Thus, the computational order for large scale FDTD problems is about $N^{\frac{4}{3}}$ [20].

In this research, we used two software implementations of the core FDTD algorithm.

TEMPEST is an FDTD simulator developed at the University of California at Berkeley [21]. The acronym stands for "Time-domain Electromagnetic Massively Parallel Evaluation of Scaterring from Topography". This is a general 3-D FDTD solver. The diffracting structures including the boundary conditions, and source-excitation conditions are specified through input-text files. This is extremely useful as one can use the same core FDTD algorithm for many different structures.

BORFDTD is an FDTD simulator developed at the University of Delaware [22]. The acronym stands for "Body Of Revolution FDTD". This FDTD engine was built to simulate radially symmetric structures. Using this symmetry property, the three dimensional FDTD problem can be reduced to a two dimensional one. This considerably reduces the computational time and memory requirements for a given problem. Therefore, this software provided the most efficient means for simulating zone plates, which are radially symmetric.

Figure 3-8 shows the FDTD domain used for simulating one zone plate. The domain is bounded by a PML of finite thickness. The excitation is a collimated-plane wave incident on the zone plate from above. The dashed line shows the axis of revolution of the domain. Note that the zone plate is radially symmetric. The FDTD time-stepping terminates when the fields inside the domain reach a steady time-harmonic state. The fields are then read out from a plane below the zone plate such as AA'. Near-to-far-field transformation techniques are used to compute the fields in the focal plane[22]. This approach ensures that the number of lattice cells within the domain stays within manageable bounds.

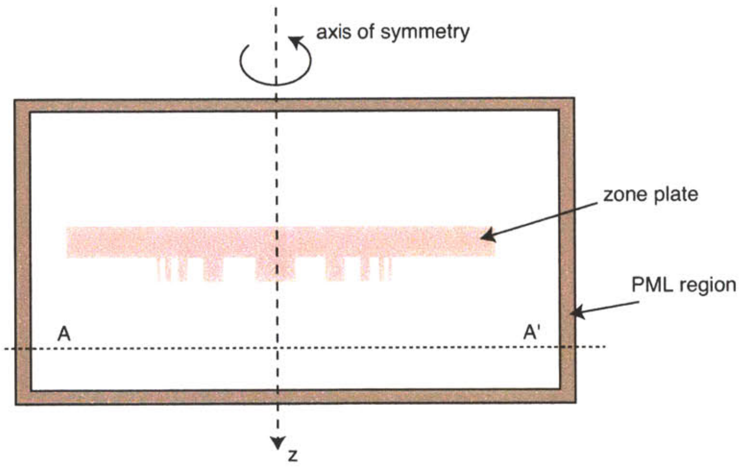


Figure 3-8: *Spatial FDTD domain for simulating one zone plate. The domain is surrounded by a layer of Berenger's PML. This simulates unbounded space. After the FDTD time-stepping is terminated, the steady state field at the plane AA' is stored. The fields in the focal plane are then calculated using a near-to-far-field transformation (i.e. propagation).*

3.3 High-Numerical-Aperture Effects in Zone Plates

As mentioned earlier, vector models for electromagnetic diffraction and propagation are invaluable to understand the performance of high numerical aperture (NA) zone plates. In order to analyze the design of a high-NA zone plate, the design equations for a zone plate are repeated here. Given a wavelength, λ, focal length, f, and NA,

$$N = \{\frac{2f}{\lambda}(\sqrt{1 + \tan^2(\sin^{-1}(NA))} - 1)\} \tag{3.36}$$

$$r_n = \sqrt{n\lambda f + (n\frac{\lambda}{2})^2} \text{ for } n = 1 \text{ to } N \tag{3.37}$$

For a fixed focal length and wavelength, as the NA increases, the number of zones increases. The widths of the zones decrease with increasing radius. Hence, high-NA zone plates have more zones which are smaller. Vector-simulation methods such as FDTD are required to take into account diffraction from features that are of the size of the wavelength or smaller.

Zone width is given by

$$zw_n = r_n - r_{n-1} \tag{3.38}$$

for all $n > 1$.

In figure 3-9, the zone width (in microns) is plotted as a function of the zone number. A wavelength of 400 nm and a focal length of 50μm were used in this calculation. It can be seen that all zones beyond zone 40 have widths less than the wavelength. This implies that zone plates with NA greater than about 0.5 have subwavelength zones[3] and hence, require full-vector simulation tools for accurate modeling.

There are three main effects due to these subwavelength features:

---

[3] This NA value is dependent on the focal length and wavelength. The numbers used here are typical for ZPAL.

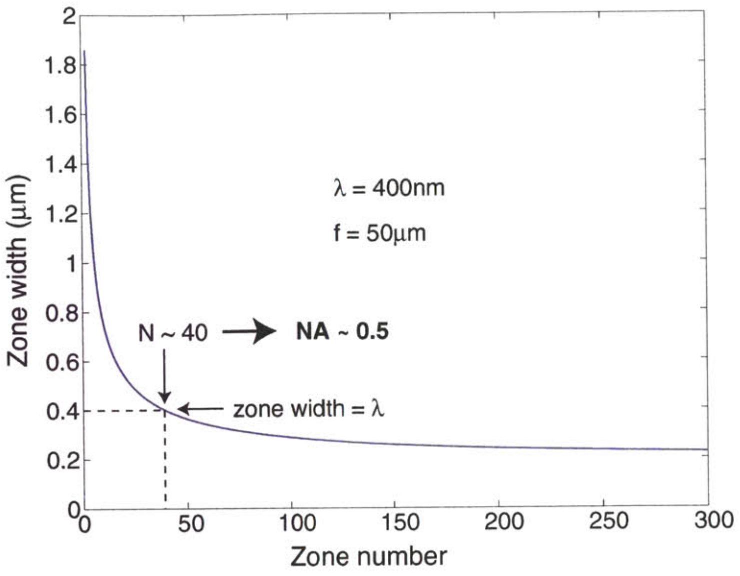


Figure 3-9: *Zone width as a function of zone number. Assuming a wavelength of 400 nm and focal length of 50μm, the width of each zone is plotted as a function of the zone number. The zone width is less than the wavelength for all zones beyond n=40. This implies that zone plates with NA greater than 0.5 have subwavelength zones (for λ=400 nm and f=40μm) and require full-vector models.*

1. The effect of the polarization of incident light.
2. The shadow effect.
3. The effect of inter-zone-field coupling.

These are discussed in the following sections.

3.3.1 Polarization

Scalar theory, as formulated earlier in this chapter does not take polarization of the incident light into account. It is well known that subwavelength structures (such as gratings) exhibit polarization-sensitive diffraction. This behavior is termed *form birefringence*. This has been exploited to create integrated optical devices such as polarizing beam-splitters [23]. When linear polarized light is incident on a high-NA zone plate, different portions of the zone plate diffract the incident light differently depending on whether the zones are aligned with the direction of polarization or not. In figure 3-10(a), linear-polarized light is incident on a zone plate. The insets show magnified cross-sections of two different regions of the zone plate, and the corresponding directions of the incident $\vec{E}$ field. It is apparent that in the two regions, the incident $\vec{E}$ field sees very different diffracting structures. Hence, the different portions of the zone plate diffract the linear-polarized-incident light with different effciencies into the first order. This gives rise to a squeezing of the spot and results in an astigmatic spot. A simulation of the field intensity in the focal plane of a NA=0.9 zone plate when a plane wave is normally incident on it with a $\lambda = 400$ nm and $f = 50\mu$m is shown in figure 3-10(b). The focussed spot shows significant astigmatism.

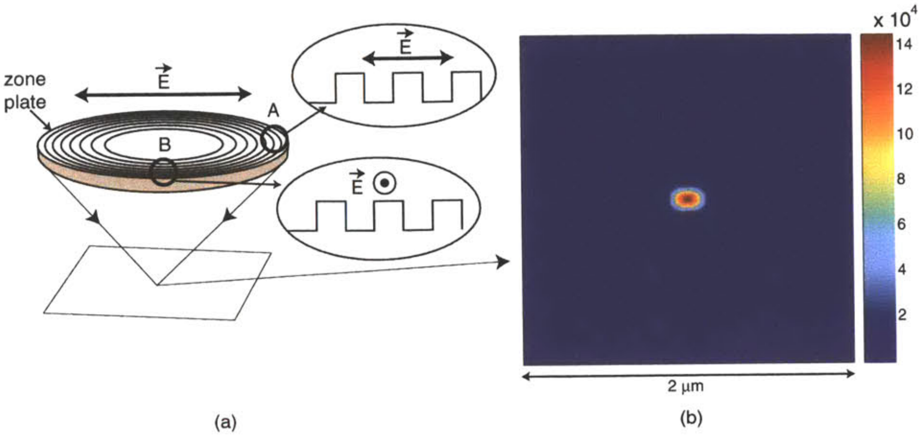


Figure 3-10: *Form birefringence in a high-NA zone plate. (a)Incident light has linear polarized electric field as shown. A and B show different regions of the zone plate and the corresponding cross-section of the zones relative to the direction of the $\vec{E}$ field. In B, the $\vec{E}$ field points into and out of the plane of the paper. The $\vec{E}$ fields in the two regions see very different diffracting structures and hence, they diffract the linear-polarized-incident light differently. (b)Simulated intensity in the focal plane of a zone plate of NA=0.9. The spot is astigmatic because of the linear polarization of the incident light. The values on the colorbars are in arbitrary units. The exposure wavelength was 400 nm and the focal length was 50μm.*

This astigmatic spot is problematic in scanning-spot lithography. In order to obtain a round spot, we use circular polarized light to illmuinate the zone plates in the experiments. Circular polarized light is a linear superposition of two orthogonal-linear-polarized components, which are phase shifted by half a wavelength with respect to one another. Figure 3-11 illustrates this idea. In figure 3-11(a), the $\vec{E}$ field vectors are directed along the Y-axis.

In figure 3-11(b), the $\vec{E}$ field vectors are directed along the X-axis. In both cases, the light propagates along the Z-axis. Note that the X-polarized light has a phase delay with respect to the Y-polarized light equal to half the wavelength. Linear superposition of these two polarizations results in circularly polarized light, as shown in figure 3-11(c). The photoresist, which records the focussed spot, is insensitive to polarization. Since, it takes finite time for the recording in photoresist to take place, the effect of circular polarized light is just to time-average the intensity of the 2 linear polarizations.

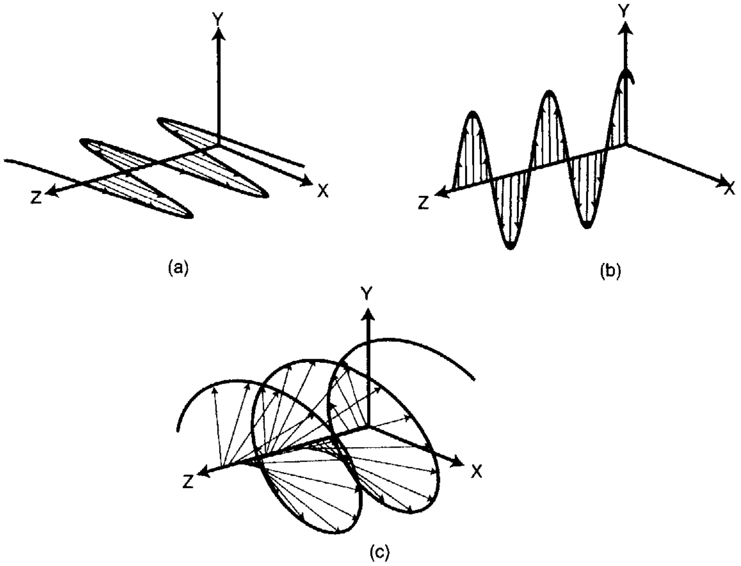


Figure 3-11: *Circular Polarized Light. (a) Light polarized with the $\vec{E}$ field in the Y-direction. (b) Light polarized with the $\vec{E}$ field in the X-direction. The wave in (b) has a phase shift of half a wavelength with respect to the one in (a). (c) Linear superposition of the waves in (a) and (b) resulting in circular polarized light. The $\vec{E}$ field vectors lie on a helix which spirals about the direction of propagation. The net effect on photoresist is a time-averaging of the intensities of (a) and (b).*

The FDTD software computes the fields in the focal plane of a zone plate for linear-polarized-incident light. In order to simulate circular-polarized-incident light, we begin by calculating the corresponding fields due to incident light with orthogonal-linear polarization. Since the zone plate is radially symmetric, rotating the polarization of incident light simply rotates the co-ordinate system (see figure 3-12). Hence, the resulting fields can be computed by a simple rotation and coordinate transformation. In the rotated coordinate system, every point gets transformed as

$$(x, y) \rightarrow (-y, x) \tag{3.39}$$

and the field components at that point gets transformed as

$$E_x \quad \rightarrow E_y \tag{3.40}$$

$$E_y \quad \rightarrow -E_x \tag{3.41}$$

$$E_z \quad \rightarrow E_z \tag{3.42}$$

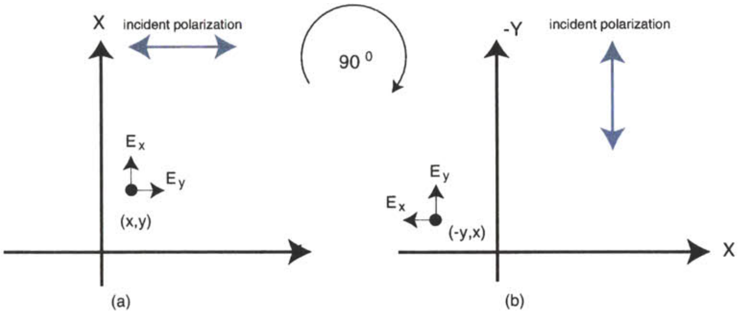


Figure 3-12: *Rotation of coordinate system. (a) The incident $\vec{E}$ field is polarized in the X-direction. The resulting diffracted fields at a point (x, y) are shown. (b) The incident $\vec{E}$ field is polarized in the Y-direction. The diffracted fields at a point $(-y, x)$ is equivalent to the fields at the point (x, y) in (a). Thus rotating the polarization of incident light simply rotates the coordinate system. This is true because the diffracting structure (zone plate) is cylindrically symmetric.*

The total field due to circular polarized light is then obtained by linear superposition of the two linear polarizations but also including a phase delay of half a wavelength to one.

$$\vec{E}_{circ} = \vec{E}^1_{linear} + e^{i\pi}\vec{E}^2_{linear} \tag{3.43}$$

The intensity in the focal plane of a zone plate of NA=0.9 is shown in figure 3-13 for linear polarized incident light (left) and circular polarized incident light (right). The circularly polarized light produces a round focused spot. Circular-polarized-incident light also changes the size of the focussed spot. As mentioned earlier, the size of the focussed spot determines

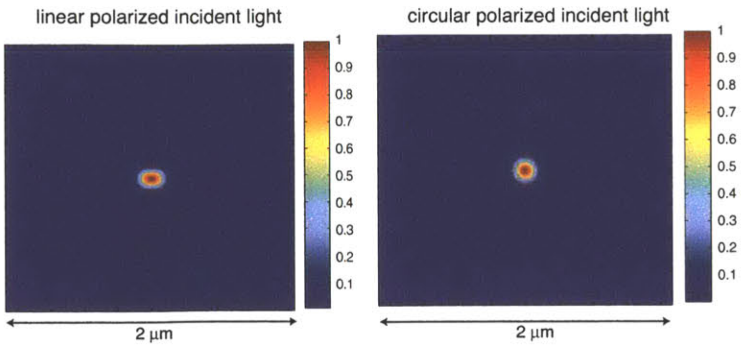


Figure 3-13: *Focussed spot with linear and circular polarized light. (a) Linear polarized incident light produces an elongated (astigmatic) spot. (b) Circular polarized incident light produces a round spot. The size of this spot is slightly larger. These simulations were performed for a phase zone plate with $\lambda = 400nm$, $f = 40\mu m$, $N = 324$ zones, which corresponds to an NA of 0.9.*

the resolution of a scanning-spot-patterning system such as ZPAL. The focussed spot for zone plates of increasing NA was simulated with linear and circular polarized incident light. The size of the focussed spot is measured as the *full-width-at-half-maximum (FWHM)*. In the case of linear-polarized-incident light, the size of the spot was measured in two orthogonal directions. As expected, the spot-size decreases with increasing NA. For NA greater than about 0.5, the spot for linear-polarized-incident light exhibits astigmatism (seen by the difference in the spot-size in the two orthogonal directions). This is in agreement with our earlier assertion that for $NA > 0.5$, subwavelength zones exist and hence, vector effects become significant. The spot-size for circular-polarized-incident light is approximately an average of the orthogonal spot-sizes for linear-polarized-incident light. This is expected since circular polarization is a time-averaged version of orthogonal linear polarizations.

In order to avoid the problem of astigmatism, it was proposed to use incident light whose polarization was radially symmetric [24, 25]. It was shown that such special polarization states produce spots which are smaller than that formed by linear or circular polarized light. This could be very useful if such special polarization could be generated in a practical manner.

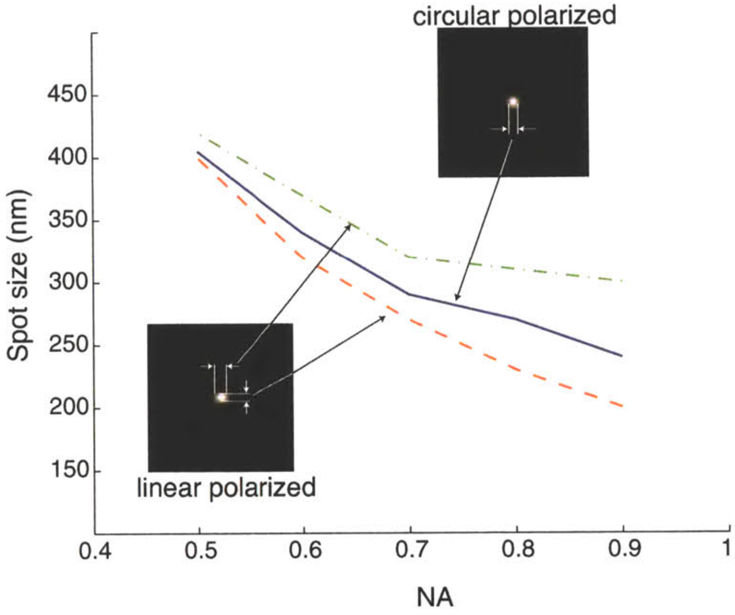


Figure 3-14: *Spot size as a function of NA. The solid line is the plot of FWHM versus NA for circular-polarized-incident light. The dashed and the dot-dashed lines are plots of FWHM in orthogonal directions for linear-polarized-incident light. These simulations were performed for a phase zone plate with* $\lambda = 400nm$, *and* $f = 40\mu m$. *The number of zones was increased to increase the NA. Note that subwavelength zones exist for NA greater than 0.5 and form-birefringence behavior becomes significant, shown by the splitting of the three curves.*

3.3.2 Shadow Effect

The *shadow effect* refers to the phenomenon of blocking of light rays diffracted from subwavelength zones by adjacent zones. This is significant when the aspect ratio of the zones in a zone plate are large i.e. the width of the zones are small compared to the height of the zones. This always occurs in high-NA zone plates. The phenomenon can be explained using a simple ray diagram as shown in figure 3-15. As the NA of a zone plate increases, light from the outermost zones has to be diffracted through larger angles to reach the focal point. These rays have to travel through other zones. This causes scattering at the zone edges and loss of diffraction efficiency. The scattered light contributes to background in the focal plane, which can deteriorate image contrast. The effect of background on image contrast in discussed in the following chapters.

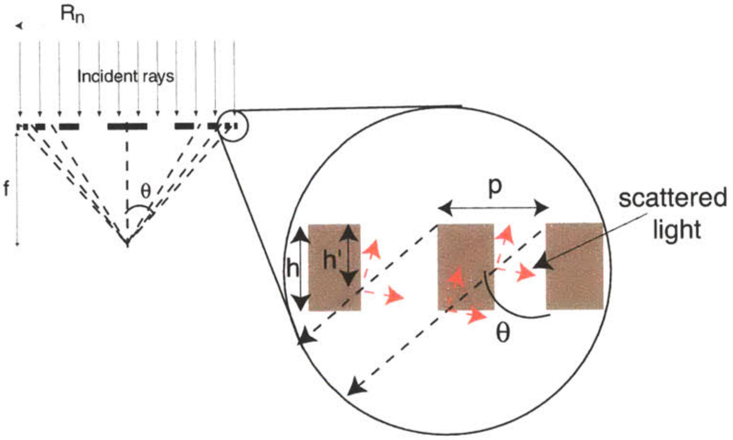


Figure 3-15: *Shadow effect in high-NA zone plates. For zone plates of high NA, the outer zones have small widths compared to their heights. The focal point lies in the geometrical shadow of light diffracted from these small zones. In other words, light from the outer zones have to be diffracted through large angles and encounter neighboring zone-edges. This causes scattering and a reduction in the focusing efficiency.*

From simple geometry, the height, h', where the diffracted ray from the outermost zone hits the adjacent zone is given by

$$h' = \frac{\lambda}{2}\sqrt{1-(NA)^2} \tag{3.44}$$

If this height is less than the height of the zones, h, then shadow effects will occur. For a π phase shift, the height of the zones is given by

$$h = \frac{\lambda}{n-1} \tag{3.45}$$

where n is the refractive index of the zone-plate material. From the above equations it is clear that shadow effects occur for all NA, provided that the refractive index of the zone-plate material is between 0 and 2. In scalar theory, the phase-shifts are modeled as a simple phase shift in the transmission function and the finite thickness of the zones are ignored. Hence, scalar theory does not take into account this shadow effect. It is interesting to note that this effect can be reduced by increasing the refractive index due to a reduction in the zone-height. Increasing the refractive index, however, increases Fresnel reflection losses at the interfaces. Furthermore, it has been pointed out that for zone widths smaller than half the wavelength, the shadow effect can be somewhat overcome due to waveguiding effects [26, 27].

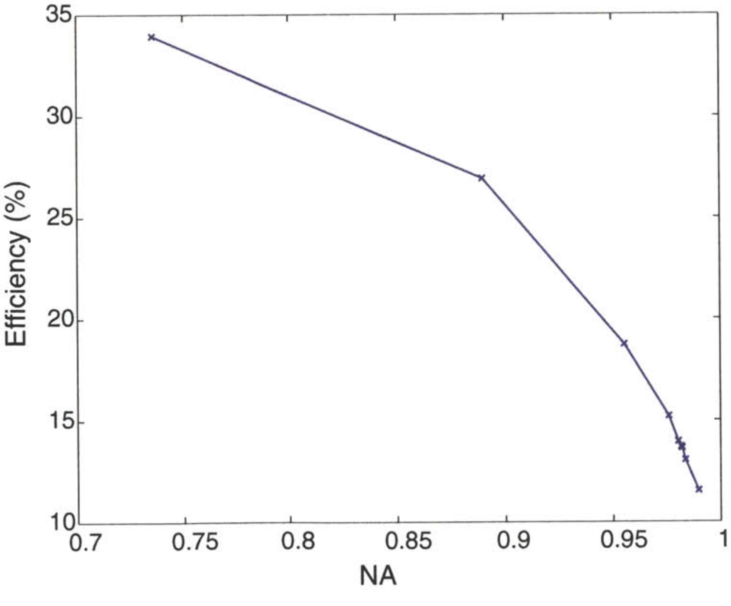


Figure 3-16: *Focussing efficiency as a function of NA. The focussing efficiency is defined as the ratio of the energy within the FWHM to the total energy in the focal plane. The energy is proportional to the light intensity and hence, can be calculated from the fields. The rapid decrease in efficiency with NA is mainly due to subwavelength and shadow effects.*

The focussing efficiency as a function of NA is plotted in figure 3-16. The focussing efficiency is defined as the fraction of the energy focussed within the FWHM to the total energy incident on the focal plane. The rapid decrease in the efficiency is mainly due to the shadow effect and zone-edge reflections. Moreover, the aspect ratio of the zones may not be the optimum for such large angles of diffraction.

3.3.3 Inter-zone coupling

When the widths of the zones become less than the wavelength of light, the coupling of electromagnetic fields between neighboring zones become significant. This is not taken into account in scalar theory, where the presence of a neigboring zone does not affect the fields in a given zone. This is another important reason why vector theory is necessary to accurately simulate subwavelength zones in high-NA zone plates.

In order to illustrate this concept, electromagnetic field propagation through a subwavelength grating was simulated. The line-to-space ratio of the grating was equal to one and the period was equal to the wavelength of the propagating light. A plot of the intensity (square of the magnitude of the $\vec{E}$ field) in one period of this infinite grating is shown in figure 3-17(a). The plot of the intensity in a grating with the grooves spaced further apart (line/space = 1/2) is shown in figure 3-17(b). The outline of the grating is shown in white in the figure. Note that the intensity inside the teeth of the grating is enhanced due to the neighboring teeth. This is fully taken into account in the vector diffraction theories.

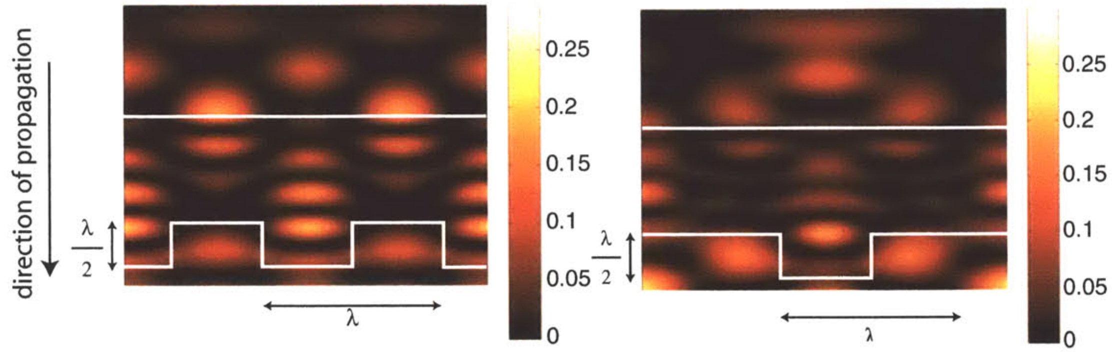


Figure 3-17: *Inter-zone coupling. The square of the magnitude of the steady-state $\vec{E}$ field is plotted for 2 cases. (a) Dense grating, where the width of each groove and trough is $\lambda/2$. (b) Semi-dense grating, where the width of each groove is λ and each trough is $\lambda/2$. Note for the dense grating, the presence of the neighboring teeth enhance the field confinement within each tooth. Thus for subwavelength structures, inter-zone coupling of electromagnetic fields are important. The incident field in this case was a TE plane wave and the simulation was based on a 2-D FDTD method (using TEMPEST).*

3.4 Fabrication of Zone-Plate Arrays

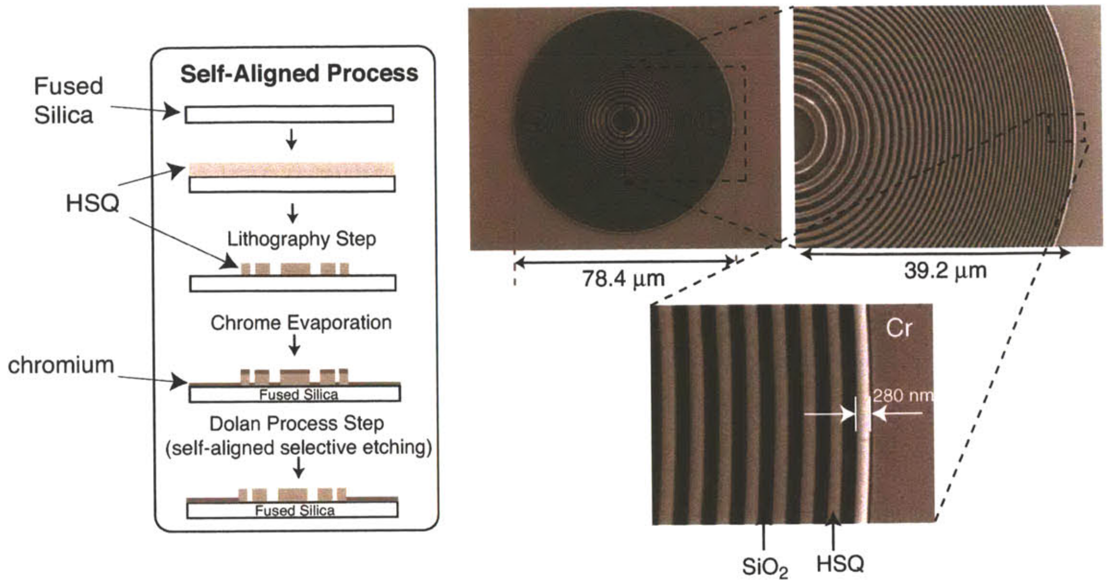
Since the radiation source in the ZPAL system at MIT has a wavelength of $\lambda = 400\,\text{nm}$, the smallest zones in the zone plates range from about 285 nm (for NA=0.7) to about 210 nm (for NA=0.95). In order to access these feature sizes reliably, the zone plates were patterned using scanning-electron-beam lithography (SEBL). The zone plates are fabricated on fused silica, which offers good transparency in the ultraviolet regime ($\lambda \leq 400\,\text{nm}$). The steps involved in the fabriation were developed by Darío Gil and are described in detail in his Ph.D thesis [28]. The major steps are briefly described below.

1. The fused silica substrate is coated with *Hydrogen Silsesquioxane (HSQ)*. HSQ is a spin-on glass and also a negative resist sensitive to electrons. After exposure and appropriate development, the resist remains in the exposed regions, whereas it is removed from all unexposed regions.

2. A thin layer of chromium is evaporated over the substrate. This layer must be thick enough to block all incident radiation from passing through those areas on the substrate, not occupied by the zone plates. The chromium covering the zone plates are electrically isolated from the rest of the chromium on the substrate. Hence, they are maintained at different electrical potentials and a selective-electrochemical-wet etching process is carried out to remove the chromium covering the zone plates. This is called the *brushfire* or *Fulton-Dolan* process [29]. Note that this is a self-aligned process as well.

After these 2 steps, the substrate contains zone-plate patterns defined in HSQ and the rest of the area is covered by chromium. Since, HSQ has a refractive index that is very close to that of fused silica at λ=400 nm, the pattern definition in HSQ on top of a fused silica substrate is sufficient to function as a zone plate. The process is depicted schematically in figure 3-18(a).

It is important that the period of the zones in the zone plate be maintained very accurately because this determines the focal length of the zone plate and hence, any period errors contributes to on-axis aberrations. This issue will be explored in detail in the following



(a) (b)

Figure 3-18: *Fabrication of zone plates. (a) Sequence of steps involved in fabricating the zone plates. The zone-plate pattern is written using SEBL in HSQ, which is spin-coated on a fused silica substrate. After development, the exposed portion of HSQ remains (negative resist) and forms the zone plate because the refractive index of HSQ is very close to that of fused silica at the wavelength of interest, $\lambda = 400\,nm$. A thin layer of chromium is evaporated over the entire substrate. Chromium covering the zone plates is selectively removed by an electrochemical-wet etch. This leaves a substrate with zone plates in open apertures. (b) Scanning-electron micrograph of zone plates fabricated in HSQ on fused silica.*

chapter. The placement of one zone plate with respect to the others in the array is another important factor during fabrication. In ZPAL, each zone plate writes patterns within its unit-cell. By combining large number of unit-cells, the entire pattern is built up. At the boundaries of the unit-cells, it is important that features line up properly. Any pattern-distortions at the boundaries of the unit-cells are termed *stitching errors.* One of the methods of reducing such errors is to place the zone plates in the array with a placement error of less than a small fraction of the spot-size. This is possible in SEBL by means of a technique known as *spatial-phase-locked electron-beam lithography (SPLEBL).* This is a novel technique invented and pursued at the NanoStructures laboratory. The basic idea involves keeping track of the position of the electron-beam by means of a signal from a grid on top of the resist, and using this signal to close the feedback loop to the electron-beam deflectors. Placement accuracy below 2 nm has been experimentally demonstrated [30].

Chapter 4

Lithographic Performance of Zone Plates

In this chapter, we analyze the performance of zone plates with regard to zone-plate-array lithography (ZPAL). The lithographic figures-of-merit for scanning-spot lithography such as ZPAL are introduced. Simulations are used to study the impact of design factors, fabrication errors, and system-level errors on these figures-of-merit. Finally, experimental results are presented to quantify the lithographic performance of zone plates.

4.1 Lithographic Figures-of-Merit

In ZPAL, arbitrary patterns are written by scanning a focussed spot. The figures-of-merit for a scanning-spot system are

1. Contrast and Resolution
2. Depth of focus
3. Throughput

4.1.1 Contrast and Resolution

Scanning-spot lithography is based on the ability to produce sufficient image contrast in photoresist. Background is present in all lithographic systems in one form or another. In optical lithography, there will be some blurring of the image as the light propagates from the mask to the substrate due to diffraction. This, together with flare (light scattered from the optics) causes undesired exposure in areas that were not intended to receive radiation. In electron-beam lithography, forward scattering of electrons in the resist and back-scattering from the substrate tend to distribute electrons in areas that were not intended to receive any electron-exposure. In both these examples, the effect is a reduction of image contrast.

In ZPAL, the problem of background is of particular significance due to the diffractive nature of zone plates. As mentioned in the last chapter, zone plates diffract light into a number of orders, whereas only one order (+1) is useful for patterning. All the other orders contribute to background, which is detrimental to image contrast. The image contrast is easily defined by means of an example. Imagine a pattern consisting of dense lines and spaces that is written using ZPAL. A schematic of the cross-section through the image of such pattern is shown in figure 4-1. The *image contrast* is defined as

$$C = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} \tag{4.1}$$

where I_{max} and I_{min} are the maximum and minimum values of the intensity in the image.

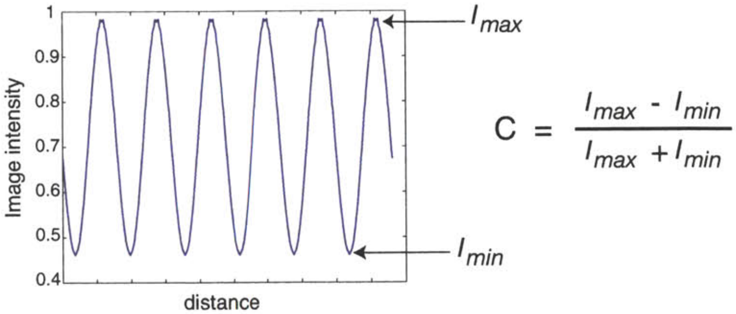


Figure 4-1: *Definition of image contrast. A schematic cross-section through a pattern of dense lines and spaces is shown. The image contrast is determined by the maximum and minimum values of the image intensity.*

The aerial image is built up by scanning the focussed spot. The photoresist records this image. As will be discussed shortly, this interaction is highly nonlinear and, to a first order, can be approximated by a thresholding function, i.e., only those areas of the resist that receive an exposure dose higher than an upper threshold undergo a photochemical reaction. This reaction induces polymeric chain-scission (for positive resist[1]), such that the resist in the exposed areas is completely removed after development. No reaction takes place in those areas of the resist that received a net exposure dose less than a lower threshold, and no resist in these areas is removed during development. Some resist is removed in those areas which receive an exposure dose between the upper and lower thresholds, and this gives rise to sloped resist sidewalls at the transition from being fully exposed to being fully unexposed. With this simple model, we can now explain the relationship between image contrast and *resolution*. The resolution is defined as the size of the smallest feature that can be patterned reliably. An example will help illustrate this idea. The cross-section through the intensity profile of a pattern consisting of dense lines and spaces is shown in figure 4-2. The lines and spaces on the left are resolved. On the right, the lines and spaces are closer together. Since the image intensity on the right falls between the two resist-thresholds, a reliable image of the pattern is not formed, and the image is said to be not resolved. This example is illustrative of the tricky concept of resolution in lithography. The size of patterned features depend not only on the aerial image and the resist characteristics but also on post-processing that is performed on the substrate after lithography. Thus, resolution depends on the imaging system as well as the image-recording system. It is appropriate at this point to discuss the role of the photoresist in pattern formation.

The contribution of Photoresist in Lithography

Photoresist is typically an organic compound that changes its solubility in another chemical (known as the developer) as a result of radiation exposure. In the case of a positive resist, the portion of the resist that is exposed to radiation is dissolved away in the developer. In the case of a negative resist, the opposite occurs. In this thesis, only positive resists are

---

[1] The photochemical reaction causes cross-linking in negative resists so that on development, only the exposed areas remain.

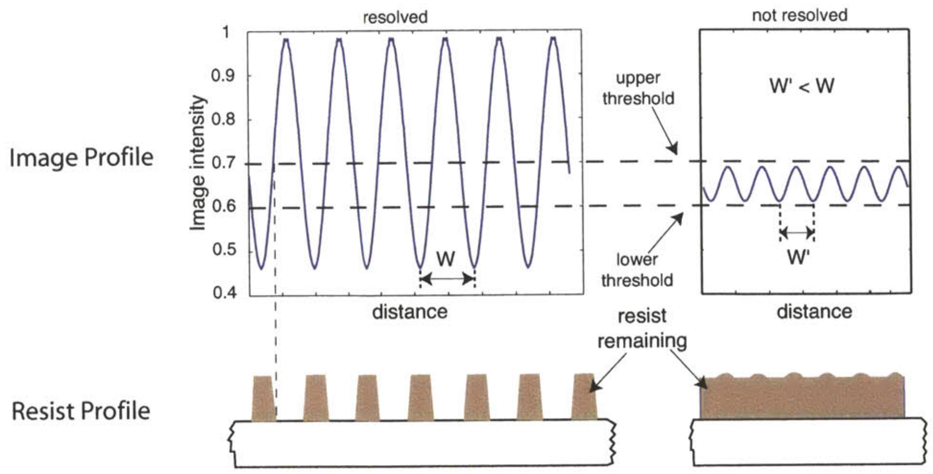


Figure 4-2: *Relationship between image contrast and resolution. Cross-section through line/space patterns od different periods is shown in the top row. The lines and spaces on the left are resolved. On the right, the lines and spaces are much smaller. The image intensity in this case falls between the two thresholds and the image is not reliably recorded (or resolved).*

considered, but without any loss of generality, the arguments are equally valid for negative resists. The performance of any photoresist can be characterized by its contrast curve. Figure 4-3 depicts the contrast curve for a fictitious positive resist. The thickness of resist remaining after development is plotted against exposure dose (in log scale for clarity). Above E_0, all the photoresist is removed. Hence, this is the upper threshold (also called the *clearing dose*). Very little photoresist is removed below an exposure dose of about E_1. Therefore, this is the lower threshold. It is obvious that the photochemical reaction is a highly nonlinear process. Nevertheless, close to E_0, the curve can be approximated as:

$$T = T_0 \gamma \ln \frac{E_0}{E} \tag{4.2}$$

where T is the thickness of resist remaining, T_0 is the initial resist thickness and γ is a factor known as the *resist contrast*. This equation is plotted as a dotted straight line in figure 4-3. In order to relate the image contrast to the resist contrast, consider the minimum of image intensity, I_{min} and the maximum of image intensity, I_{max}. As mentioned earlier, for $I_{min} < E_1$ and $I_{max} > E_0$, the image will be resolved and the resist will produce the desired

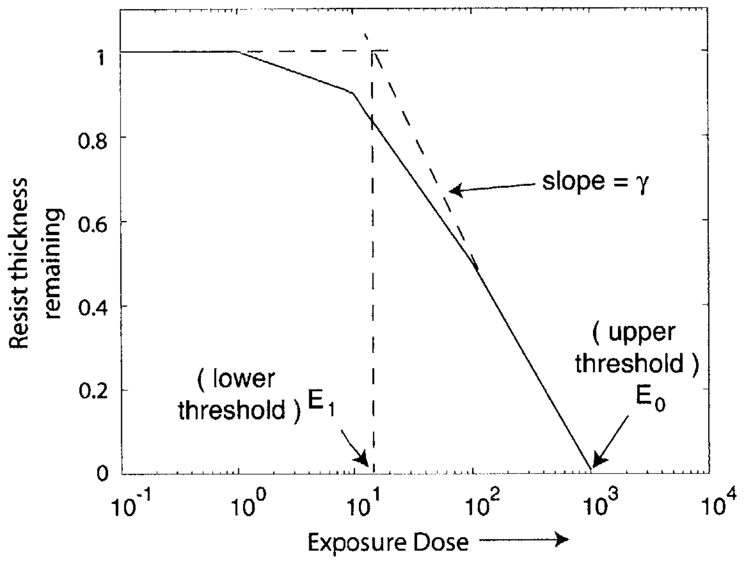


Figure 4-3: *Typical contrast curve of a positive photoresist. Thickness of photoresist remaining is plotted as a function of the exposure dose. The upper and lower thresholds are marked (see figure 4-2).*

binary image. These conditions imply that

$$C = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} > \frac{E_0 - E_1}{E_0 + E_1} \tag{4.3}$$

Using equation (4.2), the smallest image contrast that is resolved by the resist is given by

$$C_{min} = \frac{1 - e^{-\frac{1}{\gamma}}}{1 + e^{-\frac{1}{\gamma}}} \tag{4.4}$$

In figure 4-4, we plot the minimum resolvable image contrast (C_{min}) against the resist contrast (γ). From this plot, we can see that a resist of contrast equal to 2 is capable of resolving aerial images of contrast greater than about 0.25.

Resist technology has improved tremendously, and as illustrated in figure 4-5, it has dramatic effects on the quality of the patterning that can be achieved with any lithographic system. It is probably not unreasonable to claim that historically, the contribution from

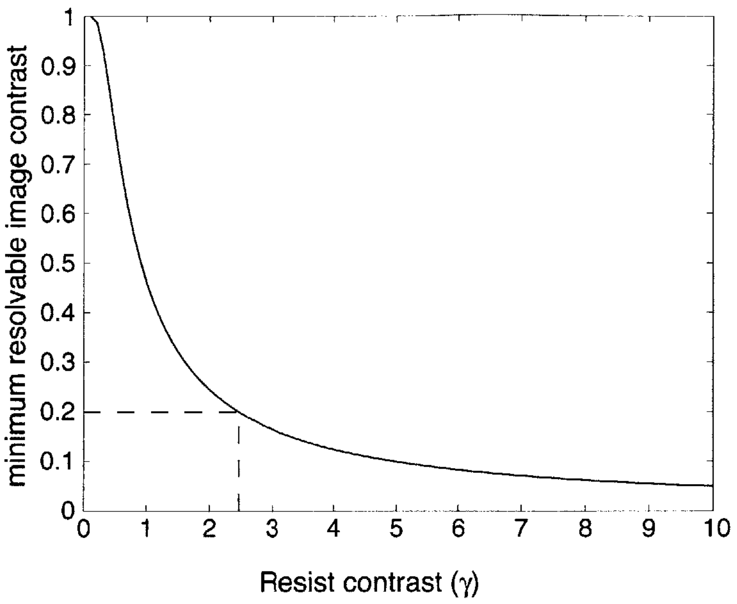
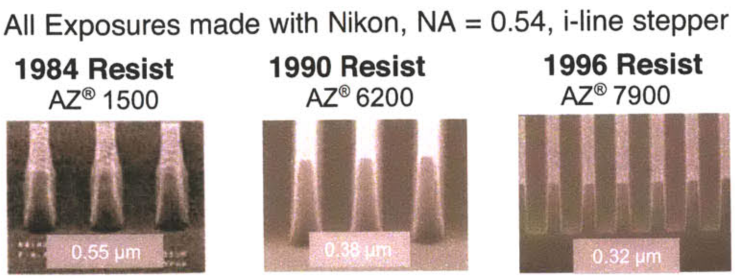


Figure 4-4: *Minimum resolvable-image contrast versus resist contrast. Note that resist of contrast of about 2 can resolve an image of contrast greater than about 0.25.*

resist and resist processing has outweighed stepper improvements [31]. Photoresists for 193 nm have γ values ranging from 4 to 10 [32, 33]. Through discussions with experts on 193nm resists at MIT-Lincoln Laboratory, ASML, as well as with a research scientist from the Advanced Lithography group at IBM, we learned that image contrasts as low as 0.2 are acceptable at 193 nm.

All the lithography results presented in this thesis were performed at λ=400 nm using Shipley S1813 photoresist. We performed an experiment to measure the corresponding resist contrast. Resist was spin-coated on a 3" silicon wafer (spun at 5000 rpm and baked on a hot plate at 90°C for 60 seconds) to a thickness of 150 nm. A layer of 200 nm of anti-reflection coating (BarLi) was spin-coated (spun at 3000 rpm and baked on a hot plate at 175°C for 90 seconds) before spinning the resist. A square pattern of size 5μm was written using a NA=0.9 zone plate. This pattern was written at several different doses. After exposure, the substrate was developed in Shipley 351 developer (diluted with water in the ratio 1:5) for 45 seconds. The thickness of resist remaining in the final patterns was measured using an atomic-force microscope (AFM). Figure 4-6(a) shows an atomic-force micrograph of nine



Source: Ralph Dammel, Clariant

Figure 4-5: *Contribution of resist advancements to high-resolution lithography. Using the same I-line lithography stepper, improved resist chemistry enables finer resolution and better profiles. All three resists were considered state-of-the-art at the time. The experiment was performed by Ralph A Dammel from Clariant.*

squares written at nine different doses. The thickness of resist remaining inside these squares was measured and plotted in figure 4-6(b). The exposure dose (in milliseconds) is plotted on a log scale. The measured γ for this resist process was about 2. From figure 4-4, it can be inferred that this resist process is capable of resolving aerial-image patterns with contrast higher than about 0.25.

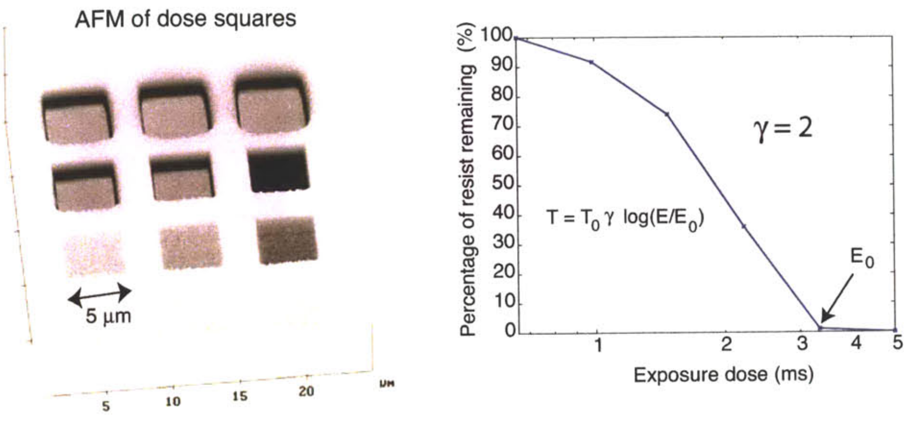
After lithography and development, the remaining resist is typically used as a mask for etching or ion-implantation. These processes transfer the pattern in the resist into the substrate. At the edges of the patterns in resist, there is usually a slope which can create problems for accurate edge-placement during this transfer step. In general, vertical resist sidewalls are desirable. The slope of the resist sidewall can be written as

$$\frac{dT}{dx}\Big|_{x=x_0} = \tan\theta \tag{4.5}$$

where x_0 is the point of intersection of the resist sidewall with the substrate and θ is the sidewall angle. Using equation (4.2),

$$\tan\theta = \frac{dT}{dE}\Big|_{E_0} \times \frac{dE}{dx}\Big|_{x_0} = -T_0\gamma \times \left(\frac{1}{E}\frac{dE}{dx}\right)\Big|_{x=x_0} \tag{4.6}$$

Thus, the factors contributing to the resist-sidewall slope can be separated into that due to the resist and that due to the zone plate. The slope can be increased (the resist profile made



(a)

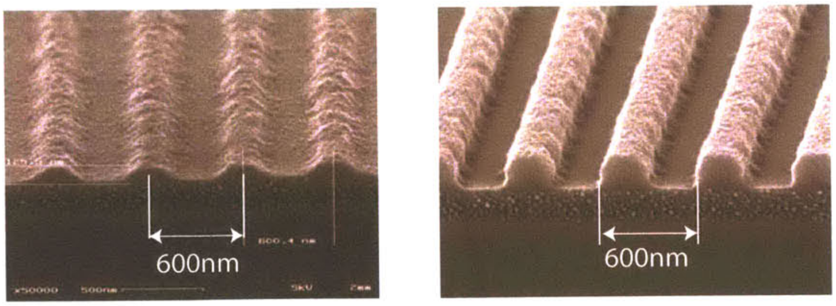
(b)

Figure 4-6: *Experimental determination of resist contrast. (a) Atomic-force micrograph of a square patterns (each of side* 5μm*) written at several doses using a* $\lambda = 400$ *nm, NA=0.9 ZPAL system. (b) Thickness of resist remaining within the squares (as a percentage of the original resist thickness) is plotted against the exposure dose (in log scale). The resist contrast,* γ *is the slope of the linear portion of this plot and was calculated to be about 2.*

more vertical) by increasing the resist contrast. This is illustrated in figure 4-7, where the same grating pattern was printed in two resists of different contrasts.

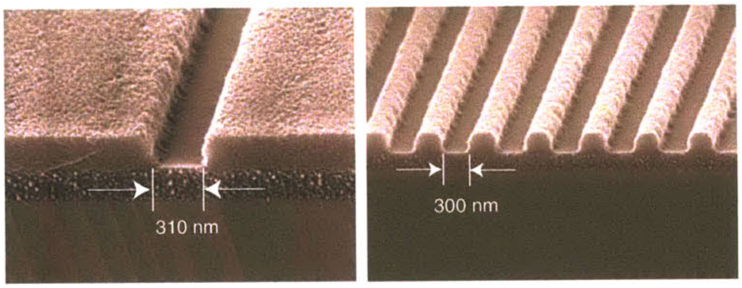
The factor due to the zone plate is a measure of the normalized slope of the intensity of the aerial-image profile at the upper threshold dictated by the resist (E_0). In image-projection systems, this factor is highly dependent on the density of the features in the pattern. However, for scanning-spot systems, the exposure dose can be adjusted from pixel-to-pixel to some extent (determined by the grayscaling capabilities of the system). This means that for any arbitrary pattern, the local dose can be adjusted so as to maximize the normalized slope of the local image at the resist threshold. This is a major advantage of scanning-spot systems such as ZPAL. This is illustrated in figure 4-8, which shows very little change in the sidewall angle for line/space patterns of different densities.

Another important metric to characterize lithography is *exposure latitude*. This is a measure of the change in linewidth of the printed feature due to a change in the exposure dose. When the exposure dose is increased, the edge of the profile moves as shown in figure 4-9, and the width of the line increases. If the exposure dose is increased by a fraction, f,



(a) (b)

Figure 4-7: *Effect of photoresist on sidewall angle. Scanning-electron micrographs of 600 nm period gratings, exposed using a NA=0.7 zone plate at* $\lambda = 400\,nm$ *in (a) Shipley 1813 photoresist and (b) PFI-88 photoresist. A layer of ARC was used between the resist and the silicon wafer to eliminate the effect of reflection back from the substrate. Note the improvement in the resist profile by using a resist of higher contrast.*



(a) (b)

Figure 4-8: *Resist sidewall angle for gratings of different densities. (a) Isolated line. (b) Dense lines and spaces. In both cases, the linewidth is about 300 nm. Both patterns were exposed in PFI-88 photoresist using a NA=0.7 zone plate at* $\lambda = 400\,nm$. *Note that the sidewall angle is approximately the same for both cases.*

we can write

$$(1 + f)E(x_0 + \Delta x) = E_0 \tag{4.7}$$

$$(1 + f)(E_0 + \frac{dE}{dx}\Delta x) = E_0 \tag{4.8}$$

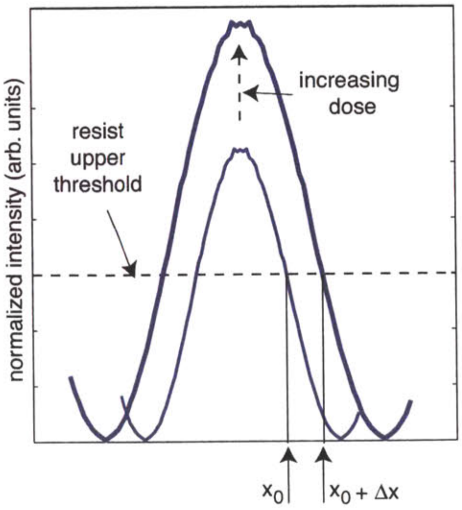


Figure 4-9: *Change of exposure dose leads to change of linewidth. An increase in exposure dose shifts the edge of the resist profile from* x_0 *to* $x_0 + \Delta x$*, increasing the width of the printed feature. Note that all the resist is removed after development for all intensity values above the upper-resist threshold.*

where the derivative is calculated at $x = x_0$.

$$\Delta x = -\frac{f}{1+f} \frac{E_0}{\frac{dE}{dx}} \tag{4.9}$$

Since both the edges move away from each other by Δx, the linewidth changes by

$$\Delta L = 2|\Delta x| = 2\frac{f}{1+f} \frac{1}{(\frac{1}{E}\frac{dE}{dx})} \tag{4.10}$$

For $f << 1$,

$$\Delta L \approx 2f \frac{1}{(\frac{1}{E}\frac{dE}{dx})} \tag{4.11}$$

From this simple model, the exposure latitude (defined as the change in linewidth due to a change in exposure dose) is only dependent on the zone plate through the normalized slope of the image profile. As pointed out earlier, this provides a significant advantage for scanning-spot systems over image-projection systems because of the ability to modulate the exposure dose of each pixel. In real systems, we can expect the resist to play a role in the exposure latitude as well. However, this simple model provides a good first-order understanding of the factors contributing to exposure latitude. We provide experimental measurements of the exposure latitude later in this chapter.

In conclusion, it can be said that resolution in ZPAL is highly dependent on the pattern being written. It depends in a complex manner on the aerial-image contrast, the resist contrast, and the underlying process latitude. For scanning-spot systems, a pattern consisting of dense lines and spaces is the most challenging one to write reliably. Hence, we will use the width of the smallest possible lines and spaces that can be written over the full field as the measure of resolution in ZPAL. In general, in lithography the resolution is expressed using a simple equation

$$W = k_1 \frac{\lambda}{NA} \tag{4.12}$$

where k_1 is a proportionality factor which takes into account the effects of the resist and any post-processing (anything other than the optics). The classical Rayleigh resolution is obtained when $k_1 = 0.61$.

The discussion so far has assumed the plane of best focus for the image formation. In the next section, we investigate the effect of deviations from this plane (defocus).

4.1.2 Depth of Focus

In zone plates, just as in refractive lenses, there exists a plane where the focussed spot has the smallest size. This plane is called the focal plane and it is situated at a distance equal to the focal length from the zone plate. The *depth of focus (DOF)* is the measure of allowable displacement from this plane such that the blurring in the spot does not significantly degrade the aerial image. It can be shown that the DOF is given by

$$DOF = \pm k_2 \frac{\lambda}{NA^2} \tag{4.13}$$

If the criteria for the DOF is such that the peak of the focussed spot is within 20% of its maximum value, then it can be shown that k_2 is then equal to 1 [34].

Figure 4-10 shows how the focussed spot changes at planes displaced from the focal plane. The peak of the focussed spot decreases and energy gets redistributed into the sidelobes from the main lobe, increasing the background. Hence, DOF is another factor which affects the contrast and resolution of the lithography system. These effects are explored in greater detail using simulations later in this chapter. Clearly, defocus also affects the normalized image slope and this, in turn, affects the resist-sidewall angle and the exposure latitude.

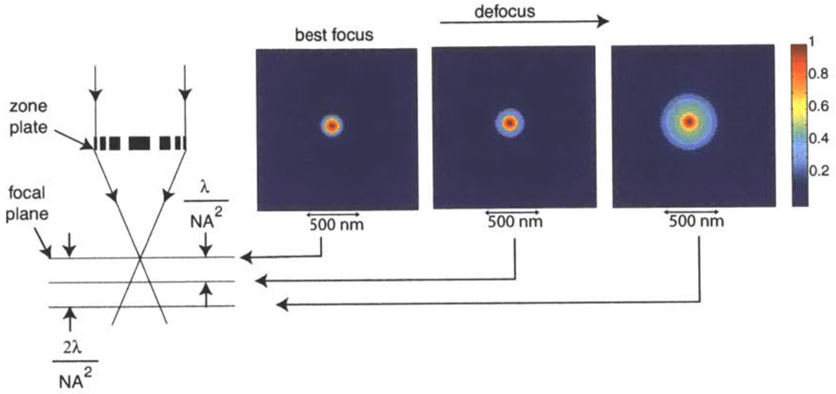


Figure 4-10: *Effect of defocus on the focussed spot. Simulated intensity of the spot at planes displaced from the focal plane are shown. The peak of the spot decreases (not shown since the data was normalized) and the sidelobes increase contributing to background.*

In an incoherent system such as ZPAL, the DOF of the aerial image is dictacted by the DOF of each spot making up that image[2]. This is in contrast to the case of image projection, where the useful DOF depends on the type and size of the features in the image and is typically smaller than the case for the single spot. This is another inherent advantage of scanning-spot systems over image-projection systems.

In order to estimate the linewidth variation due to a defocus of Δz, we can write

$$E(x_0 + \Delta x, z + \Delta z) = E_0 \tag{4.14}$$

where E_0 is the upper threshold (clearing dose) of the resist. Expanding the above equation

---

[2] This assumes that the transverse-scan axes lie in a plane such that every point on that plane is within the DOF of the entire zone-plate array.

in Taylor series and only keeping the quadratic term in Δz, we get

$$\Delta x = -\frac{1}{2}(\Delta z)^2(\frac{1}{E_0}\frac{\partial^2 E}{\partial z^2})\frac{1}{\frac{1}{E_0}\frac{\partial E}{\partial x}} \tag{4.15}$$

where the derivative with respect to x is evaluated at $x = x_0$ and the derivative with respect to z is evaluated at $z = 0$ (at the focal plane). The assumption here is that E is symmetric about the focal plane in z, implying

$$\frac{\partial E}{\partial z}\Big|_{z=0} = 0 \tag{4.16}$$

This assumption is not strictly true for high-NA zone plates. This first-order estimate suggests that DOF is dependent only on the optical system and not on the resist properties. The change in linewidth ($|2\Delta x|$) decreases with increasing normalized-image slope.

For thick resists, the above analysis is modified. The presence of the resist breaks the symmetry of E in z due to refraction at the air-resist interface. This has to be considered when the thickness of the resist is comparable to the DOF of the system. Figure 4-11 shows the rays impinging on a thick resist layer on top of a substrate. At the air-resist interface, the rays undergo refraction and bend towards the normal inside resist because the refractive index of resist, n is typically higher than that for air ($n_{air} = 1$). Since DOF depends on the angle of convergence, the new DOF inside the resist film can be written as

$$DOF' = k_2\frac{\lambda/n}{(\sin(\theta'))^2} \tag{4.17}$$

where θ' is the angle of convergence of the rays inside the resist. Using Snell's law, we know that

$$n\sin(\theta') = \sin(\theta) \tag{4.18}$$

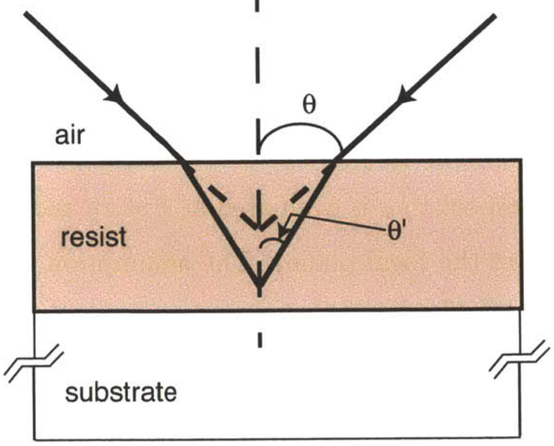


Figure 4-11: *Effect of thick resist on DOF. The rays bend due to refraction at the air-resist interface. This changes the angle of convergence of the rays and hence, modifies the DOF of the system. This is important when the thickness of the resist is comparable to the DOF of the system.*

Hence, the new depth-of-focus is given by

$$DOF' = k_2 \frac{\lambda}{n} \frac{1}{(\frac{\sin(\theta)}{n})^2} = n \times DOF \tag{4.19}$$

where DOF is the depth-of-focus in air given by equation (4.13). Thus, the effect of resist is simply to increase the DOF by a factor equal to its refractive index. In this simple analysis, we assumed that the factor, k_2 does not change with the thickness of the resist. In reality, this factor is process-dependent. However, this analysis gives a good first-order estimate.

In conclusion, the depth-of-focus is a very important engineering-design criterion for any lithography tool. This provides a stringent tolerance for maintaining the gap between the substrate and the zone-plate array in ZPAL. In a later chapter, we describe the technique that is used to set and maintain this gap for the experiments reported in this research. It must be pointed out that from equation (4.13) it is clear that decreasing the wavelength or increasing the NA have the effect of decreasing the DOF. This could lead to serious process latitude issues for future generations of optical lithography.

4.1.3 Throughput

The *throughput* of a lithography system is the measure of the speed with which patterns can be written. This figure-of-merit is important for scanning-spot systems, since writing a single spot at a time to build up the entire pattern is inherently a slow process. However, in ZPAL, this disadvantage is overcome by having a large number of beams writing different parts of a pattern at the same time. Here, we analyze the effect of various factors that contribute to the throughput of a ZPAL system

All patterns in ZPAL are built from single spots. In order to clear the photoresist within a single spot, a certain amount of energy is required. This puts a constraint on the exposure time. If P is the power focussed into a spot by one zone plate, W is the diameter of the spot, and S is the sensitivity of the resist (measured in energy per unit area required to clear the resist), the exposure time, t_{exp} can be calculated from the energy requirement as follows.

$$P \times t_{exp} = S \times \frac{\pi W^2}{4} \tag{4.20}$$

This spot is then scanned across the substrate[3] at a continuous velocity. It will be shown later that in order to obtain smooth edges, neighboring spots are exposed such that they overlap considerably i.e. the scan distance between 2 consecutive exposures is a fraction, f of the spot size, W. Now, if we assume that the frame rate of the multiplexers is $1/T$, then the stage velocity is given by

$$v_{stage} = \frac{fW}{T} \tag{4.21}$$

One underlying constraint in the above analysis is that the exposure time is much shorter than the period of consecutive exposures. In other words,

$$t_{exp} << T \tag{4.22}$$

---

[3] In reality, the substrate is scanned across the spot, but this does not change any of the succeeding analysis.

Otherwise, the spot recorded in resist will be smeared in the direction of the scan. It was found experimentally that the exposure time should be less than about 20% of the exposure period.

It is appropriate at this time to discuss the optimum scan strategy in ZPAL.

Scan Strategy

A significant portion of time is spent at the beginning and the end of each scan for acceleration, and deceleration respectively. Therefore, it is prudent to reduce the number of turns during patterning. With this objective, the most efficient scan is achieved by rotating the zone-plate array (ZPA) such that neighboring zone plates are writing neighboring lines (see figure 4-12). If we assume that the array is close packed, the angle of rotation can be calculated as

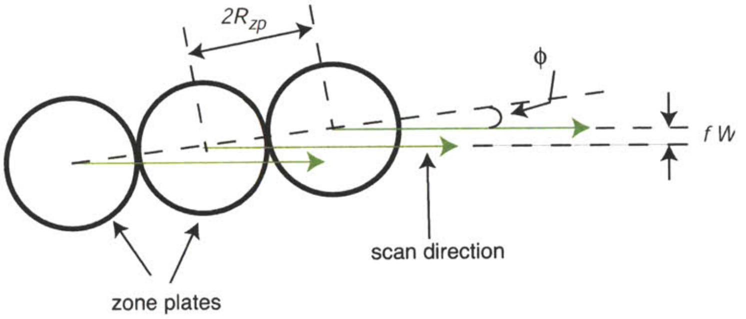


Figure 4-12: *Rotation of zone-plate array for efficient scanning. A close-packed array of zone plates is rotated by an angle, ϕ such that adjacent zone plates write adjacent lines in resist. The scan direction is denoted by the green arrows. This ensures that contiguous areas are written with minimum turns.*

$$\sin(\phi) = \frac{fW}{2R_{zp}} \tag{4.23}$$

where R_{zp} is the radius of one zone plate. As mentioned earlier, the radius of each zone plate is fixed by the wavelength (λ), focal length (f), and the numerical aperture (NA).

The relevant equations are repeated here.

$$N = \{\frac{2f}{\lambda}(\sqrt{1 + tan^2(sin^{-1}(NA))} - 1)\} \tag{4.24}$$

and

$$R_{zp} = R_N = \sqrt{N\lambda f + (N\frac{\lambda}{2})^2} \tag{4.25}$$

where N is the total number of zones in the zone plate.

The largest contiguous area that can be written with the shortest scan length is when the left-most zone plate in the ZPA reaches the starting x-position of the right-most zone plate in the ZPA. This scan strategy is illustrated in figure 4-13. Assuming a square zone-plate array of side L_{zpa}, the scan length is given by

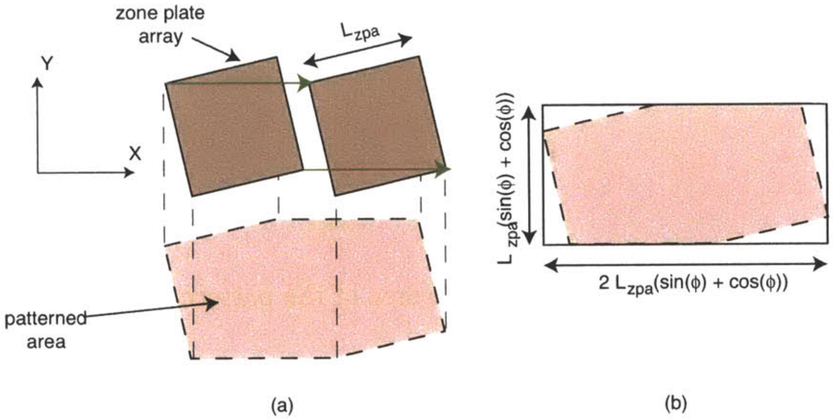


Figure 4-13: *Scan Strategy with rotated zone-plate array. (a) The rotated zone-plate array is scanned in the X-direction such that the left-most zone plate moves until it reaches the starting point of the right-most zone plate. The area written during this one scan is shown below. (b) Written area is a hexagon. The area can be calculated as the area of the rectangle shown minus the four triangles (in white).*

$$L_{scan} = L_{zpa}(\sin(\phi) + \cos(\phi)) \tag{4.26}$$

The time taken for this scan can be calculated from the velocity of the stage. In order to take into account the time required to accelerate, decelerate, and move to a different scan line, an overhead equal to 100% of the time taken to scan a single line is added to the total. Each written hexagon can be tiled to cover the entire substrate as shown in figure 4-14. Thus, the

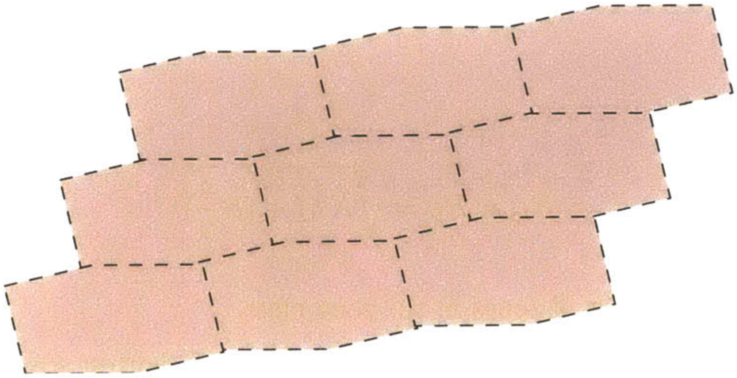


Figure 4-14: *Tiling the hexagons to cover the entire substrate. The hexagonal area written during a single scan can be tiled to cover the entire area of the substrate. The time taken to move from one hexagon to the next is added as an overhead into the calculation.*

time required for one scan on an average is given by

$$t_{scan} = 2 \times \frac{L_{scan}}{v_{stage}} \tag{4.27}$$

Using simple geometry (see figure 4-13), the area of the patterned hexagon is given by

$$\begin{aligned} A &= 2(L_{zpa})^2(\sin\phi + \cos\phi)^2 - 4 \times \frac{1}{2}(L_{zpa})^2\cos(\phi)\sin(\phi) \\ &= 2(L_{zpa})^2(1 + \sin(\phi)\cos(\phi)) \end{aligned} \tag{4.28}$$

Thus, the writing speed of ZPAL is given by

$$WS = \frac{A}{t_{scan}} = L_{zpa}(\frac{1 + \sin(\phi)\cos(\phi)}{\sin(\phi) + \cos(\phi)})(\frac{fW}{T}) \tag{4.29}$$

The writing speed increases by increasing the size of the ZPA (L_{zpa}) or by increasing the refresh rate of the multiplexing element ($1/T$). The writing speed is a weak function of the angle, ϕ, especially for small angles. In figure 4-15, we plot the throughput in terms of number of 12" wafers per hour (obtained by dividing WS in equation (4.29) by the area of one 12" wafer) as a function of the NA of the zone plates in the ZPA. It is interesting to note that the throughput increases with NA. This is mainly due to the increase in the size of each zone plate and hence the increase in the size of the array. The length of one side of the square array is also plotted as a function of NA to make this relationship explicit. The calculation was performed assuming $\lambda = 193$ nm, focal length $= 50\mu$m, the multiplexer refresh rate, $1/T = 10$KHz, sub-pixel stepping equal to half the spot-size, $f = 0.5$, and a zone-plate array consisting of 1000 by 1000 zone plates. Note that for a typical case of NA=0.85, the throughput is about **two** 12" wafers an hour and the size of the ZPA is about 8 cmX8 cm.

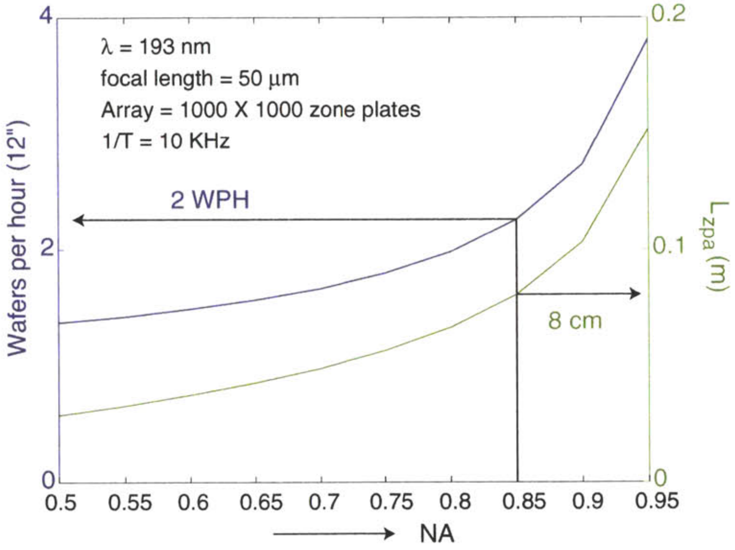


Figure 4-15: *Throughput as a function of NA. The throughput (in 12" wafers per hour) is plotted on the left-Y axis as a function of the NA using equation (4.29). The throughput increases with NA mainly because the size of the ZPA increases with NA. Note that W decreases with NA. The parameters used in this calculation were typical:* $\lambda = 193$ *nm, focal length* $= 50\mu m$, *multiplexer refresh rate = 10KHz, sub-pixel stepping equal to half the spot-size,* $f = 0.5$, *and zone-plate array size of 1000 by 1000 zone plates.*

Source Power Requirements

In order to compute the power requirements of the source to meet throughput requirements, we first assume that 5% of the power from the source makes it down into focussed spots on the substrate. Total source power is then given by

$$P_{src} = \frac{1}{\frac{5}{100}} \times 10^6 \times P \tag{4.30}$$

where P is the power in each focussed spot given by equation (4.20). The factor, 10^6 comes from the fact that the source should be capable of delivering the power, P to all the 10^6 spots focussed by the zone-plate array. Assuming a resist sensitivity of 20 mJ/cm^2, and the exposure time, t_{exp} equal to 20% of the multiplexer-refresh period ($t_{exp} = T/5 = 20\mu$s), and the same parameters used for the plot in figure 4-15, we get $P_{src} = 8.7$ watts for NA=0.85. This power is achievable in pulsed-DUV-laser sources.

Data Bandwidth Requirements

The pattern data is transferred from a control computer to the multiplexing device. This data transfer must be fast enough so that the data on the multiplexer is ready for each refresh cycle. It was shown in some earlier work [35] that 5 bits of grayscaling is sufficient for good pattern control. This corresponds to 32 levels of gray. The data delivery system must be capable of transferring 5 bits of data to every zone plate in the ZPA at a refresh period of T. Assuming a multiplexer refresh rate of 10KHz, the data-transfer rate is

$$\begin{aligned} \text{data rate} &= (1000 \times 1000) \times (5\,\text{bits}) \times 10KHz \\ &= 5 \times 10^{10}\,\text{bits/sec} = 46.57\,\text{Terabits/sec} \end{aligned} \tag{4.31}$$

where 1 Terabit = 2^{30} bits.

This is quite a challenging requirement. However, by using parallel data channels, this bandwidth requirement can be considerably relaxed. In addition, improvements in short-range interconnect technology will make these data rates possible in the near future.

From the above analysis, it is clear that throughput is a metric that depends heavily on the actual design of the system. We have presented an optimum design, which maximizes throughput without consideration to other practical difficulties such as the uniform illumination of a large ZPA (8 cm by 8 cm). Another caveat in the above calculation is that we assumed a 193 nm source which is capable of producing repeatable pulses at a frequency of at least 10KHz. This may be possible in principle but commercial 193 nm sources produce pulses that are not repeatable. To average out this pulse-to-pulse variation, multiple passes through the same feature are required. This will reduce the throughput by a factor equal to the number of multiple passes. Nevertheless, we may conclude that throughput in ZPAL is a highly customizable metric and it is reasonable to claim that with proper design, moderate throughputs are readily achievable.

With these lithographic figures-of-merit, we now proceed to evaluate the performance of ZPAL using both simulation and experiments.

4.2 Lithographic Simulation Process

Lithographic simulations begin with the calculation of the light intensity in the focal plane of one zone plate. The procedure for this calculation was described in the previous chapter. Unless mentioned otherwise, all the point-spread functions (PSFs) used for the lithographic simulations in this chapter were calculated using the body-of-revolution finite-difference time-domain (FDTD) method. A plot of the PSF for a NA=0.7 zone plate at $\lambda = 400\,\text{nm}$ is shown in figure 4-16.

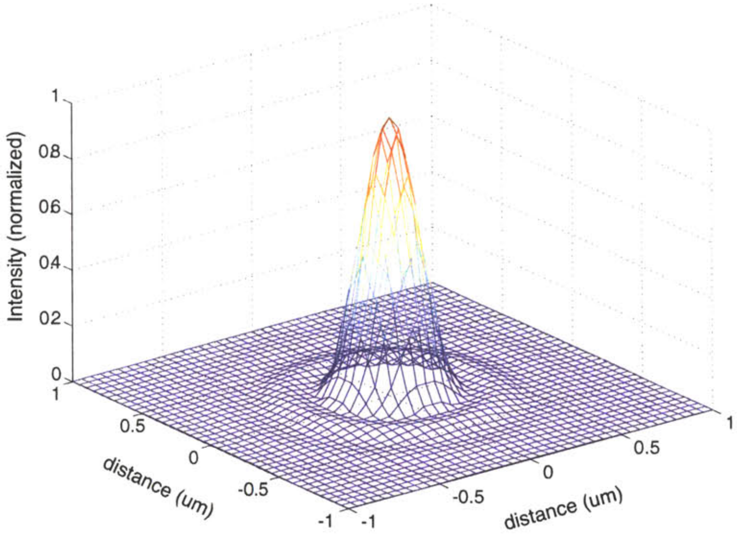


Figure 4-16: *PSF of a NA=0.7 zone plate. The PSF was simulated using the FDTD method with circular polarized incident light at* $\lambda = 400\,nm$.

The exposed pattern is simulated by a discrete convolution of the PSF with the desired *binary* pattern. The binary pattern, which defines the edges of features is typically drawn in a CAD program such as Nanowriter[4]. The convolution is implemented using the FFT algorithm. The simulation also takes into account the "ON-OFF" contrast of the multiplexing device. This simulation process is illustrated in figure 4-17.

---

[4] Nanowriter is a CAD software authored by Dr. Mitchell Meinhold at the Nanostructures Laboratory.

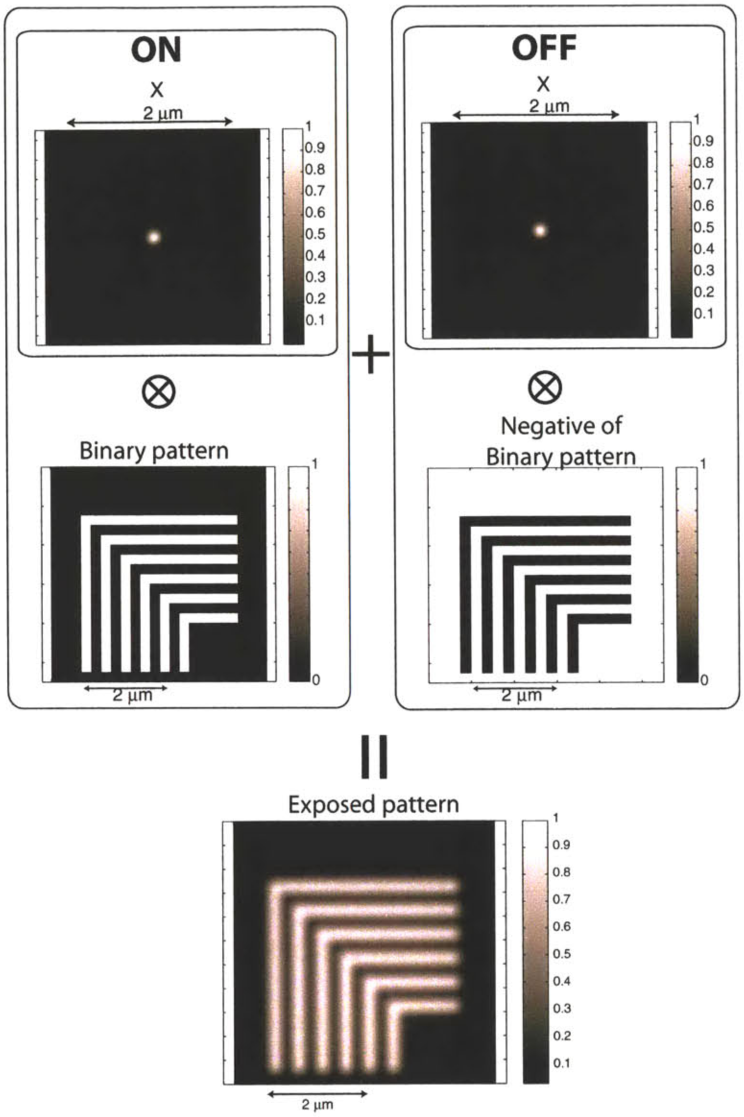


Figure 4-17: *Process for simulating patterns in ZPAL. The PSF is first calculated using FDTD. The exposed pattern is a discrete convolution of the PSF with the binary pattern. The effect of the contrast in the multiplexing device is taken into account by adding the convolution of the "OFF-state" PSF with the negative of the binary pattern. The "OFF-state" PSF represents the light that is focussed by the zone plate even when the multiplexer is in the "OFF" state. The intensity of light in this PSF depends on the "ON-OFF" contrast of the multiplexing device.*

The exposed pattern is simulated as

$$EP = \{(ON \times PSF) \otimes (BP)\} + \{(OFF \times PSF) \otimes (1 - BP)\} \tag{4.32}$$

where EP is the exposed pattern, BP is the binary pattern, and $(1 - BP)$ is the negative of the binary pattern. In order to understand the tolerance of the multiplexer-contrast on image contrast, we calculated the image contrast for a pattern consisting of dense lines and spaces for different values of the multiplexer contrast. The parameters used for the simulation were $\lambda = 400\,\text{nm}$, $f = 40\mu\text{m}$, $NA = 0.7$, and the period of the grating period of 500 nm.

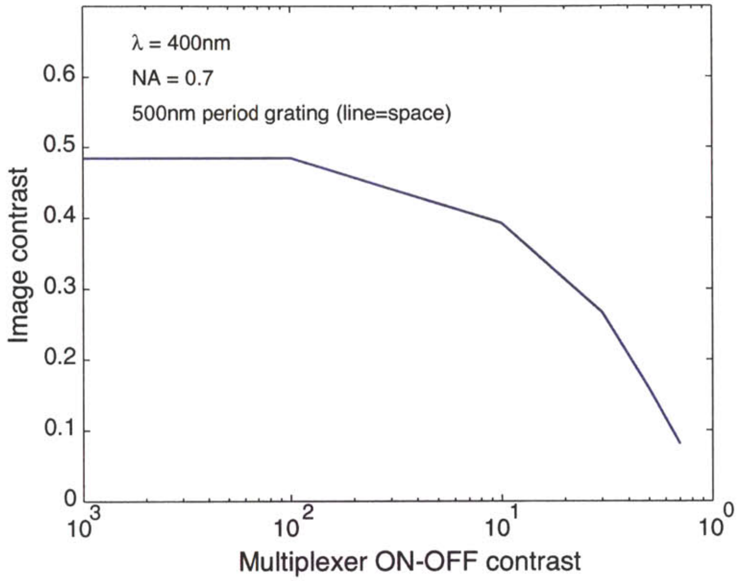


Figure 4-18: *Effect of multiplexer ON-OFF contrast on image contrast. A dense line/space pattern of period 500 nm was simulated using the PSF of a NA=0.7 zone plate at $\lambda = 400$ nm. The multiplexer contrast is plotted in log-scale for clarity. Note that for contrast more than 100 to 1 (ON to OFF), the effect on image contrast is insignificant.*

By using periodic boundary conditions, we can reduce the computational burden for simulating large periodic patterns. The periodic boundary conditions inherent in the discrete FFT algorithm enable us to compute the convolution for a small portion of large periodic patterns and still obtain accurate results. It is important to ensure that the pattern is periodic from one boundary to the next. Otherwise, spurious numerical artifacts such as reflection from the boundaries will distort the results. In figure 4-19, a grating pattern of period 500 nm is simulated using the PSF of a NA=0.7 zone plate at $\lambda = 400\,\text{nm}$. In the top row, the periodic boundary conditions are used, whereas in the bottom row, the boundaries are not periodic. This is evident in the top-down view of the image intensity shown on the left. On the right is the cross-section through the image intensities near the center of the pattern as indicated by the dotted line. Note that there is distortion at the boundaries of the cross-section due to the mismatch in the boundary conditions. This distortion disappears in figure 4-19 (a), where the boundary conditions are matched to the periodicity of the pattern. Therefore, it is important to match the periodicity of the pattern with the periodic boundary conditions for accurate numerical results.

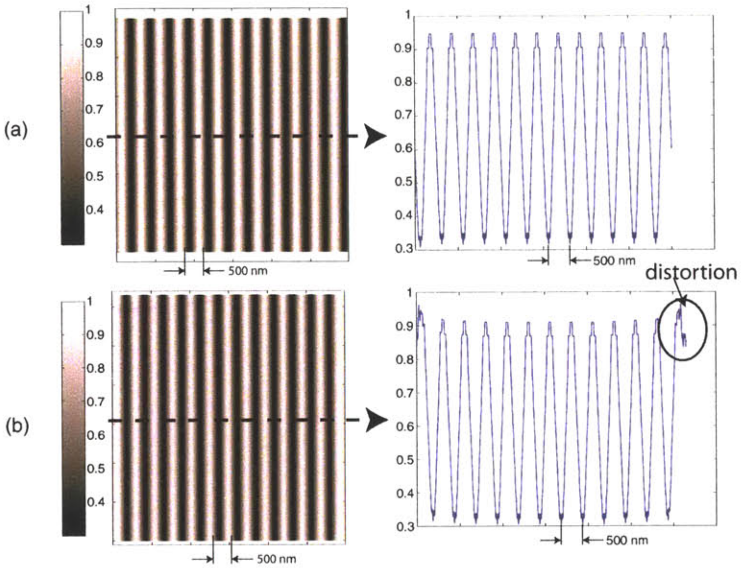


Figure 4-19: *Effect of boundary conditions on image contrast. A grating of period 500 nm is simulated using a zone plate of NA=0.7 and* $\lambda = 400\,\text{nm}$ *in (a) using periodic boundary conditions which match the period of the grating, and in (b) the periodic boundary conditions are not matched to the period of the grating. On the right, are cross-sections through the image intensity which show distortion at the boundaries.*

The PSF falls off rapidly with increasing radial distance. This is evident in figure 4-20 (a), in which the cross-section through a PSF is plotted in log scale as a function of the radial distance. For simulation purposes, the PSF in equation (4.32) must be truncated to a finite distance from its center. It is important to estimate the distance to which the PSF should be computed so that this has little impact on the simulated lithographic patterns. A pattern consisting of dense lines and spaces (of period 500 nm) was simulated with the PSF truncated at different radial distances from the center. The PSF was set to 0 beyond the truncation distance. The image contrast measured in the simulated pattern was plotted as a function of the distance of truncation in figure 4-20 (b). As expected, adding more data to the PSF by increasing the truncation radius, decreases the image contrast. However, the change in image contrast becomes negligible for truncation radii greater than about 3μm. This is due to the fact that the PSF falls off rapidly. At about 3μm away from the center of the spot, the PSF has fallen about 4 orders of magnitude (see figure 4-20 (a)). The exact truncation distance will depend on the type of the simulated pattern. In this study, we are concerned with the worst-case patterns for ZPAL, which are dense gratings. Care was taken during the simulations to ensure that the truncation distance was large enough so as not to affect the results.

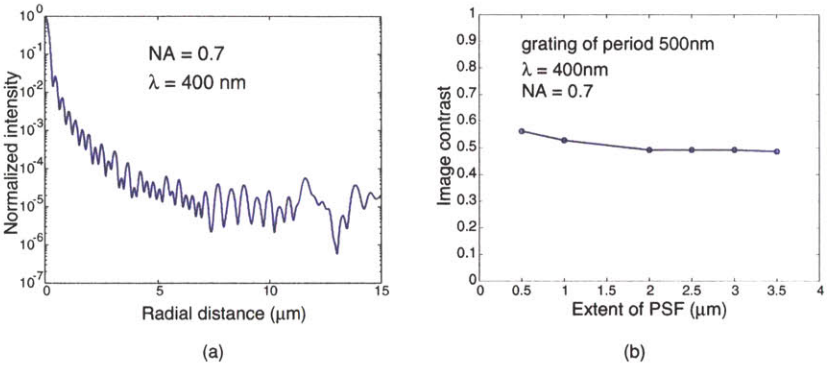


Figure 4-20: *Effect of PSF-truncation radius on image contrast. (a)The PSF of NA=0.7 zone plate at* $\lambda = 400\,nm$ *is plotted in log-scale versus the radial distance (in microns). (b) A dense grating of period 500 nm is simulated using the PSF in (a). The PSF is truncated at different distances and all points beyond this distance are set to 0. The image contrast decreases with increasing truncation distance. The image contrast remains constant beyond a truncation distance of about 3*μ*m. The diameter of the zone plate was about 80*μ*m.*

Armed with these tools to simulate large-area patterns that are exposed using ZPAL, we present a comprehensive study of lithographic patterning in ZPAL.

4.3 Simulation Results

Almost all the simulations presented in this section were performed using the PSF computed by an FDTD method. The results presented also assume phase zone plates. The image contrast is used as the metric in almost all the simulations.

4.3.1 Image Contrast

As described earlier, the image contrast, C is a function of the maximum light intensity in an image (I_{max}) and the minimum light intensity in the same image (I_{min}),

$$C = \frac{I_{max} - I_{min}}{I_{max} + I_{min}}. \tag{4.33}$$

This definition of image contrast is valid irrespective of the type of the pattern in the image. In fact, one would expect the image contrast to degrade for images with patterns of high spatial frequency. Hence, in this section, we consider the worst-case scenario for a scanning-spot lithography system, namely full field exposure of dense lines and spaces (of equal width). If we have acceptable image contrast for this case, we would expect the contrast to be better for all other patterns of practical interest.

In this section, we study the effect of the scan method, NA, and defocus on image contrast of dense gratings.

Sub-pixel stepping

As mentioned earlier, it is necessary to overlap spots for adequate edge-placement and linewidth control. This technique is referred to as *sub-pixel stepping.* In order to quantify the sub-pixel stepping, we define the s-factor as

$$s = \frac{\text{linewidth}}{\text{step size}} \tag{4.34}$$

where the step size is the stepping grid for the stage i.e. the points on the substrate where the focussed spots are put down. It can be seen that by increasing the s-factor, the line-edge roughness of the patterns can be decreased considerably. Going from $s = 1$ to $s = 2$, one can see that the edge-roughness of the grating lines in figure 4-21 improves. The image contrast decreases slightly because of the increase in the number of printed spots.

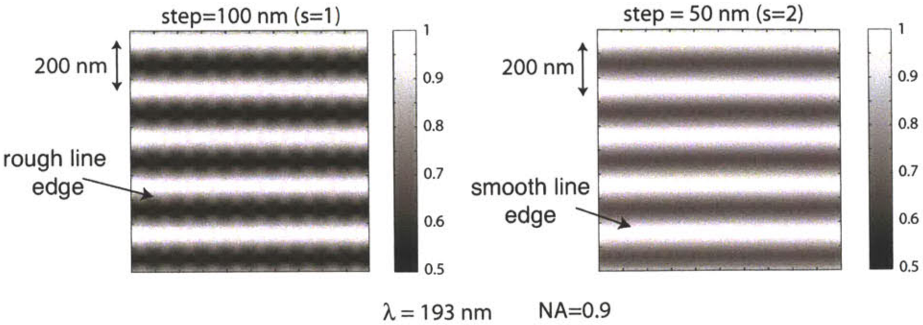


Figure 4-21: *Effect of the s-factor on exposed patterns. A dense grating of period 200 nm was simulated for different s-factors using the PSF of a zone plate of NA=0.9 and* λ*=193 nm. For* $s = 1$*, we can see that the line-edge roughness is worse than when* $s = 2$.

In figure 4-22, the image contrast is plotted as a function of the s-factor for gratings of different periods. The simulations were performed using the PSF of a NA=0.7 zone plate at $\lambda = 400\,\text{nm}$. The contrast increases with increasing grating period. For each grating, the contrast decreases with increase in the s-factor. The s-factor is a measure of the overlap of the exposed spots. As this overlap increases, more spots can be squeezed into the same area, and the background contributions from this increased number of spots degrade the image contrast. Increasing the s-factor beyond 2 does not improve the pattern significantly. Higher s-factors decrease the writing speed. So, s-factor equal to 1 or 2 was used in almost all the experimental and simulation results presented in this thesis.

Numerical Aperture (NA)

Increasing the NA of a zone plate has two main effects on its PSF: the size of the spot decreases and the size of the sidelobes increases. This is because, the area of each additional zone in a zone plate is slightly larger than the area of the previous zone. For high-NA zone plates, this results in a extra emphasis on light diffracted from the outer zones. These zones represent the higher spatial frequencies. Thus, there is an implicit reverse-apodization present in the design of a high-NA zone plate, which causes the sidelobes to be higher. The

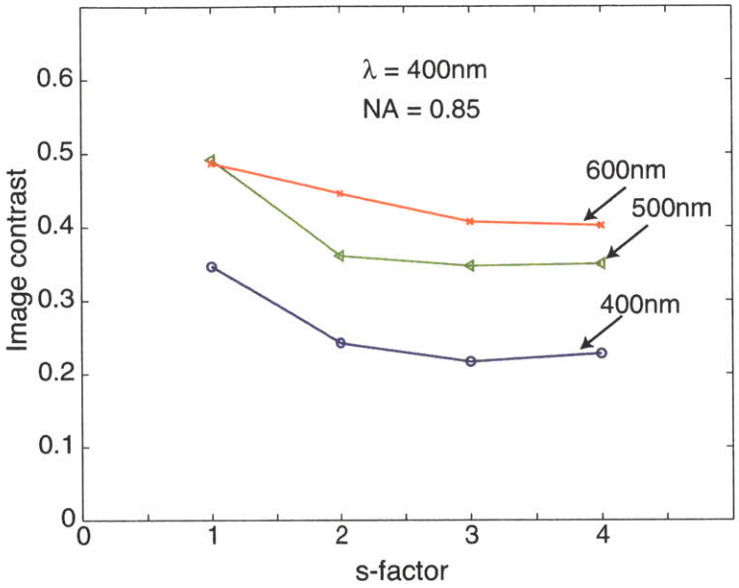


Figure 4-22: *Effect of the s-factor on image contrast. The image-contrasts of gratings of periods, 600 nm, 500 nm, and 400 nm are plotted as a function of the s-factor. As the s-factor increasees, the spots are pushed closer to each other (increased overlap), resulting in increased background which decrease the contrast.*

effects of decreasing spot-size and increasing side-lobes act in such a way as to counter each other with regard to aerial-image contrast. These effects are illustrated in figure 4-23, where cross-sections through the PSFs of two zone plates of NA=0.7 and NA=0.95 are shown.

This would suggest that there is an optimum NA which maximises the image contrast. In figure 4-24, the image contrast is plotted as a function of NA for gratings of different periods. At higher NA, the spot-size is smaller but the sidelobes are more pronounced. The numerical simulations indicate that the optimum NA for $\lambda = 400\,\text{nm}$ is about 0.85, for the smallest period of 400 nm. This result suggests that the normalized slope of the image profile at the resist threshold in also maximum for NA=0.85.

Defocus

The simulations presented until now were based on the spot-intensity in the plane of best focus. In this section, we study the effect of defocus of the spot on image contrast. This is an important metric for process latitude in lithography. In figure 4-25, the image contrast for

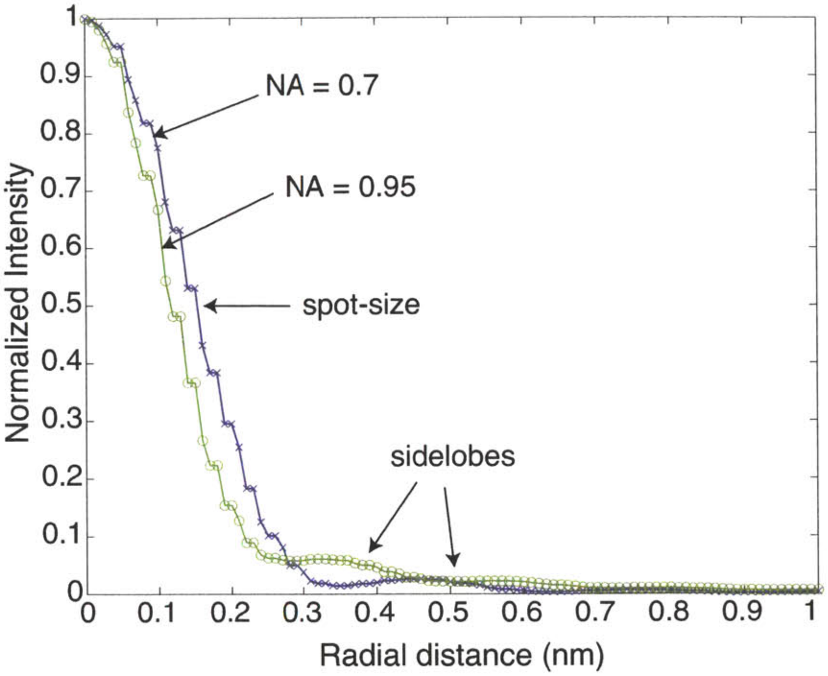


Figure 4-23: *Effect of NA on the PSF. The PSFs for two zone plates at* $\lambda = 400$ *nm and NA=0.7 and NA=0.95 are shown. Note that increasing the NA decreases the spot size but also increases the sidelobes.*

gratings of different periods is plotted as a function of defocus of the spot. These calculations were performed for a zone plate of NA=0.85 and $\lambda = 400\,\text{nm}$. An s-factor of 2 was used as well. The depth-of-focus of this zone plate is given approximately by $DOF = 0.5\frac{\lambda}{NA^2} = 277\,\text{nm}$. The defocus corresponding to the DOF is marked in the figure. For all gratings, the maximum image contrast occurs when the defocus is 0. The contrast falls with defocus because of the blurring of the spot.

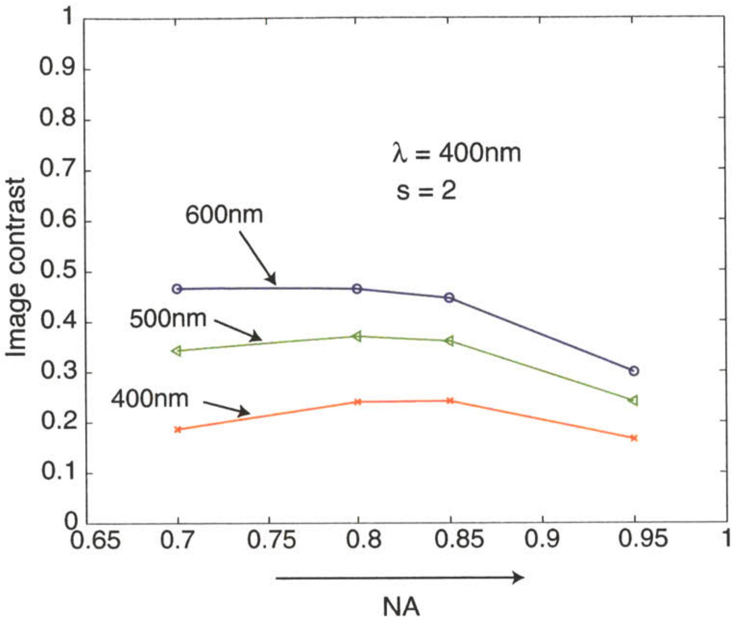


Figure 4-24: *Effect of NA on image contrast for different grating periods. The image contrast is plotted as a function of NA for dense gratings of different periods. Note that at an NA≈0.85, the contrast is maximum for the smallest period of 400 nm.*

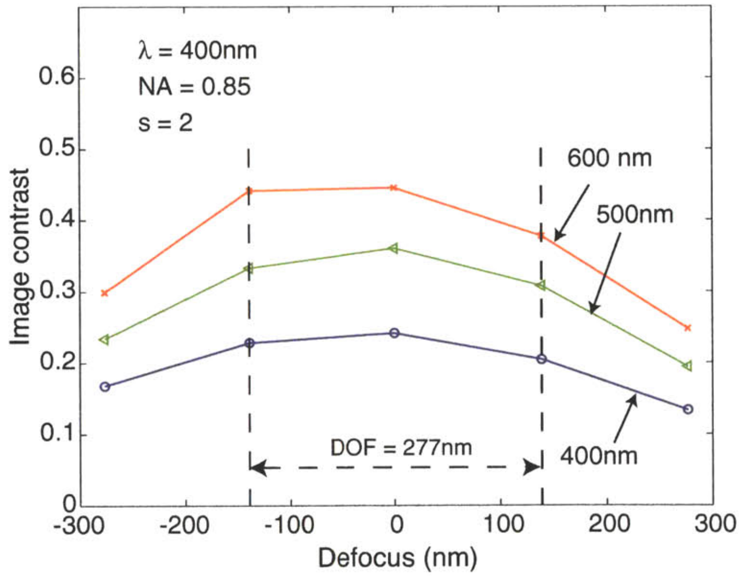


Figure 4-25: *Effect of defocus on image contrast for different grating periods. The image contrast is plotted as a function of defocus for dense gratings of different periods. These simulations were performed for a zone plate of NA=0.85 and* $\lambda = 400$ *nm. The s-factor was set to 2. The image contrast decreases with defocus for all the gratings. This is due to the blurring of the defocussed spot.*

4.3.2 Effect of Fabrication Errors

The results presented until now assumed a perfect zone plate. We now proceed with an extensive analysis of how zone-plate fabrication errors affect contrast in ZPAL. The fabrication of efficient zone plates requires the precise patterning of a large number of sub-wavelength surface-relief structures. For close to ideal performance, three parameters are critical: (1) the phase-shift between alternate zones must be as close to π as possible, (2) the local duty-cycle of each pair of zones must remain close to 50-50, and (3) the period of each pair of zones must be accurately controlled. The following three sections detail these three requirements and the impact of errors on the performance of the zone plates.

Phase Etch Tolerance

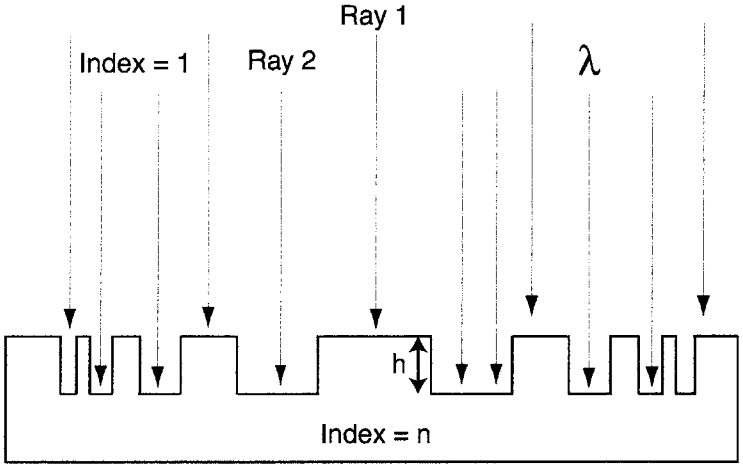


Figure 4-26: *Cross-section of a phase zone plate. Alternate zones are phase-shifted to maximize the efficiency into the first order as well as to cancel the zero order.*

The fabrication of a phase-zone plate entails producing a phase shift in alternate zones of the zone plate. Fabrication procedures can yield errors in the phase shift. These errors can be caused by insufficient control in reactive-ion etchers or other associated processes. Although sophisticated phase shifts can be introduced for high-numerical aperture zone plates resulting in optimized spots, the basic and most widely used phase zone plate requires a π-phase shift between alternate zones. The relation between the height of the etched grooves and the desired π phase shift is given by:

$$h = h_0 = \frac{\lambda}{2(n-1)} \tag{4.35}$$

where λ is the wavelength and n is the refractive index of the zone plate material. This height h_0 reduces the zero-order diffraction efficiency to zero.

A change in the height of this profile results in a proportional change in the phase shift in that zone. This, of course, would result in a reduction in the portion of the incident light that goes into the focused spot. Note that this will not change the focal length since the focal length is only dependent on the radial period of the zones.

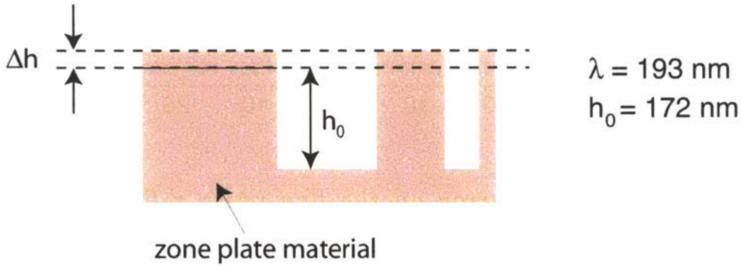


Figure 4-27: *Schematic of phase-shift error (etching error). Δh is the error in the etching height. For a phase zone plate at $\lambda = 193\,nm$ and a material index of 1.5609 (fused silica), the height required for a π phase shift is 172nm.*

We can use the method used in section 3.1.2 to analyze the effect of phase-etch error on the first-order diffraction efficiency of a phase zone plate. This analysis is valid for low-NA zone plates. However, it offers simple (analytic) expressions for an initial understanding of the effect of phase-etch errors. The derivation for this simple expression is given in Appendix B. Here, we present the result for the first-order diffraction efficiency, η_1, as a function of the phase-etch error, Δh,

$$\eta_1 = \frac{2}{\pi^2}\{1 + \cos(\Delta h(n-1)\frac{2\pi}{\lambda})\} \tag{4.36}$$

where n is the refractive index of the zone-plate material and λ is the wavelength. Using the values $\lambda = 193\,\text{nm}$ and $n = 1.5609$, η_1 is plotted as a function of Δh in figure 4-28. As expected, the efficiency is highest (40%) when the phase-etch error is 0. The variation in diffraction efficiency is small when the phase-etch error is small.

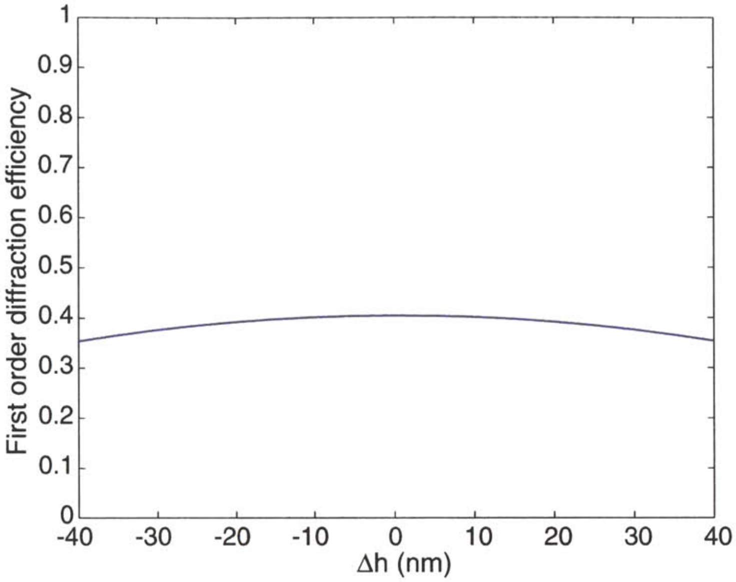


Figure 4-28: *Effect of phase-etch error on the first-order diffraction efficiency. The plot is based on equation (4.36). The parameters used were* $\lambda = 193\,nm$*, and* $n = 1.5609$*. As expected the first-order efficiency is a slowly varying function of the phase-etch error for small values of the error.*

In Figure 4-29, we plot the image contrast as a function of the etching error for a variety of patterns consisting of lines and spaces of various linewidths and densities. All dimensions are normalized to the wavelength. The first thing to notice is that as the linewidth decreases the image contrast also decreases for all cases. This is to be expected. It can also be noted that contrast increases as the density of the grating is reduced. As anticipated, image contrast decreases with increasing etching error. The interesting thing to note is that the image contrast seems to be a slowly varying function of the etching error.

Effect of Duty Cycle Error

Duty-cycle errors are defined as the percent deviation of each pair of zones in a zone plate from their intended linewidths (see figure 4-30). This concept is easier to explain for the case of a regular grating. A grating with period P is said to have a 50-50 duty cycle if all the

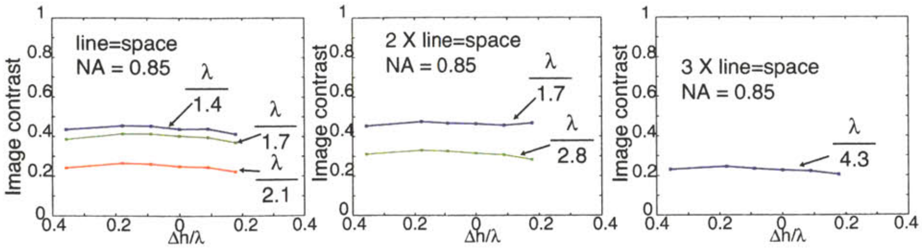


Figure 4-29: *Image contrast versus etching error for lines/space patterns. These simulations were performed using the PSF of a NA=0.85 zone plate with an illuminating wavelength of 193 nm. Image contrast was analyzed for a variety of patterns, including: Left: Dense lines and spaces (1:1) at various linewidths. Center: Semi-Dense lines and spaces (1:2) Right: Semi-Dense lines and spaces (1:3).*

lines and spaces constituting the grating have a width of P/2. If the lines are bigger than the spaces, or vice-versa, the grating's duty cycle is no longer 50-50, and a percent error can be associated with this deviation. A zone plate can be described as a diffractive element containing a large number of small-area gratings, and locally, they should have a duty cycle that is as close to 50-50 as possible. Deviations from such a situation will result in the appearance of even orders in the diffracted radiation, in addition to the always-present odd orders, resulting in a degradation of the focused spot in terms of efficiency. This degradation will have a corresponding impact on the image contrast.

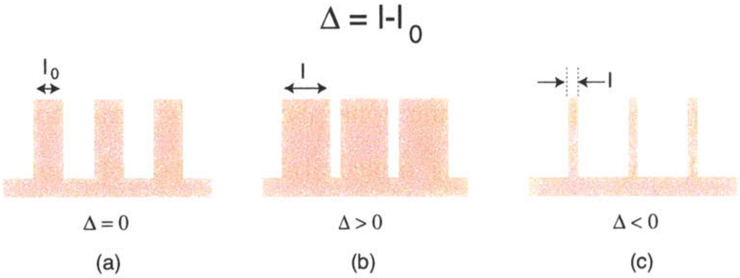


Figure 4-30: *Schematic illustrating duty-cycle errors. (a) Cross-section through a small region of a perfect zone plate. Note the ratio between alternate zone widths is close to 1. This corresponds to a duty cycle of 50%. (b-c) Cross-section of zones exhibiting duty cycle errors. Alternate zones are no longer of equal width.*

Figure 4-31 (left) shows the contrast as a function of the error in duty-cycle for dense gratings of various linewidths.

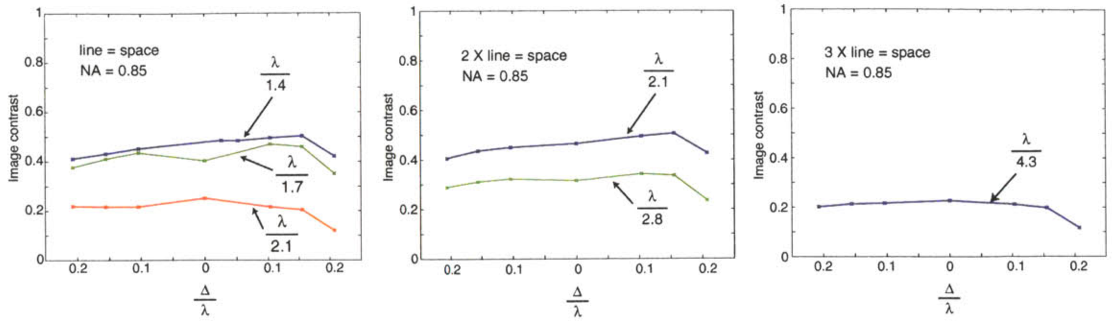


Figure 4-31: *Image contrast versus error in duty cycle for dense gratings of various linewidths and different densities. These simulations were performed with the PSF of a NA=0.85 zone plate with an illuminating wavelength of 193 nm.*

In conclusion, duty-cycle errors can have a significant impact on image contrast, especially for dense patterns. Hence, the zone linewidths should be carefully controlled in the manufacturing process.

Effect of Radial-Period Error

Radial-period control is the determining factor in optimizing how well zone plates are tuned to the wavelength of choice. This, in turn determines our ability to control the focal length of the zone plates. The focal length is critically important, especially for large zone-plate arrays that can span many square centimeters in area. Focal length uniformity is paramount if we want each zone plate to create a diffraction-limited spot at the same focal plane. Deviations from a single focal plane will cause some zone plates in the array to write features with different dimensional control, something unacceptable in state-of-the-art lithography. An appropriate metric to analyze this type of error consists of measuring how contrast degrades, for a fixed focal plane, as radial period errors are introduced.

Calculations indicate that for diffraction-limited performance of the zone plate, all zone boundaries have to be placed within 20% of the smallest zone width[36]. For a zone plate with an outer zone width of 100 nm, this means that all zone boundaries must be placed within 20nm of the intended value. Note that these errors will typically reduce the diffraction efficiency into the first order as well.

The method used to model this radial-period error consists of adding a stochastic vari-

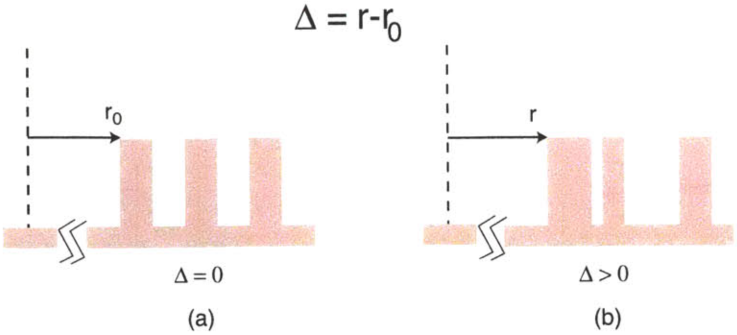


Figure 4-32: *Schematic illustrating possible zone placement errors (radial period errors). (a) Cross-section through a small portion of a zone plate containing no errors. The radius of the first zone is* r_0 *(measured from the center of the zone plate). (b) Same zone plate with errors in zone radii. Specifically, now* r *is the radius of the first zone.* $\Delta = r - r_0$ *is a measure of this placement error. A random distribution of errors across the zones of the zone plate is used to model the placement errors introduced by the actual fabrication process. The standard deviation of these random errors is varied to study the effect on image contrast.*

ation, Δ (see figure 4-32), to each zone radius. These variations represent the possible placement errors introduced in the fabrication process. The errors are modeled as a normal distribution with zero mean and finite standard deviation (the standard deviation is a measure of the placement error). For the analysis presented here, circular symmetry of the resulting zone plate is assumed. Despite this assumption, the model provides an excellent estimate of the effect that pattern-placement fabrication errors produce. The results are presented below.

Figure 4-33 provides a plot of the image contrast as a function of the standard deviation for gratings of various linewidths and densities.

In conclusion, image contrast is a slowly varying function of radial period errors. Again, since no single error should be seen in isolation from other errors that are likely to be incurred in zone plate fabrication, zone placement control in the nanometer range[5] would be beneficial.

Finally, with regard to fabrication errors, zone plates are very robust elements. This is due to the fact that a zone plate is a "Fourier-Transforming" element. Hence, any small

---

[5] This kind of control has been achieved at M.I.T's NanoStructures Laboratory by means of a technique called Spatial-Phase-Locked Electron Beam Lithography [30].

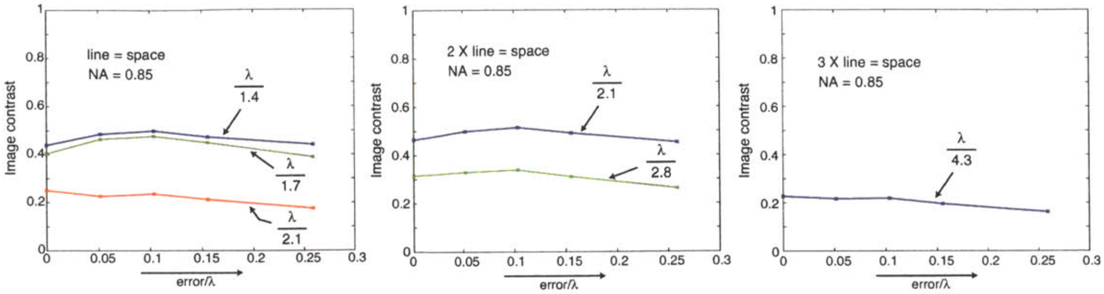


Figure 4-33: *Image contrast versus zone-radii errors for dense gratings at various linewidths and densities. These simulations were performed with the PSF of a NA=0.85 zone plate with an illuminating wavelength of 193 nm.*

local errors in its structure is averaged out in the focussing process.

4.3.3 System-level Tolerances

In this section, we analyze the effect of tolerances at the system-level on the lithographic performance of zone plates. In particular, we study the effect of source bandwidth, and the angle of illumination.

Source Bandwidth

The spectral bandwidth of the source is important in ZPAL since diffractive lenses exhibit strong wavelength dependence (this is known as chromatic aberration). A small spread in spectrum mainly tends to smear the focal spot in the axial direction. A simple criterion for bandwidth is to ensure that this smear is within the depth-of-focus (DOF) of the lens. The DOF [15] and the numerical aperture (NA) of a zone plate are given by the following equations,

$$DOF \approx \frac{\lambda}{NA^2} \tag{4.37}$$

$$NA = \frac{R_{zp}}{\sqrt{R_{zp}^2 + f^2}} \tag{4.38}$$

where R_{zp} is the radius of the zone plate and f is the focal length at the central wavelength λ. The radius of a zone plate containing N zones is given by

$$R_{zp} = \sqrt{N\lambda f + (\frac{N\lambda}{2})^2} \tag{4.39}$$

Finally, the DOF of the zone plate can be expressed as

$$DOF = \lambda\{1 + \frac{f^2}{N\lambda f + (\frac{N\lambda}{2})^2}\} \tag{4.40}$$

If P is the smallest period in the zone plate, then

$$P = \sqrt{N\lambda f + (\frac{N\lambda}{2})^2} - \sqrt{(N-2)\lambda f + (\frac{(N-2)\lambda}{2})^2} \tag{4.41}$$

On inverting the above equation, we get the focal length as

$$f = \frac{(N-1)(P^2-\lambda^2) + P\sqrt{N(N-2)(P^2-\lambda^2)}}{2\lambda} \tag{4.42}$$

From the above equation, it is apparent that the focal length is a function of the wavelength. If the source has a spread in wavelength, this will appear as a blur of the spot along the optical axis. This axial blur is tolerable as long as it is within the DOF of the designed zone plate. We go on to calculate this maximum tolerable bandwidth.

The change in focal length for a change in wavelength is given by

$$\Delta f = \Delta\lambda \frac{df}{d\lambda} \tag{4.43}$$

By differentiating the equation for f with respect to λ, we get

$$\frac{df}{d\lambda} = -(N-1) - \frac{P}{2}\sqrt{\frac{N(N-2)}{P^2-\lambda^2}} - \frac{(P^2-\lambda^2)(N-1) + \frac{P\sqrt{N(N-2)(P^2-\lambda^2)}}{2}}{2\lambda^2} \tag{4.44}$$

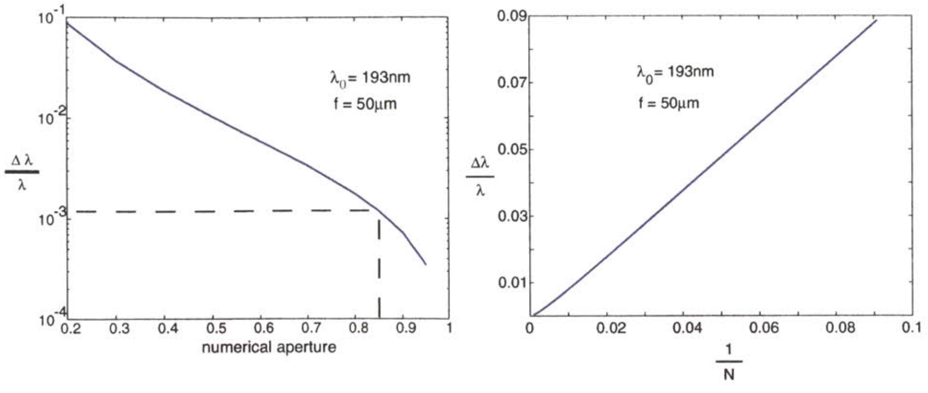
As mentioned earlier, we need to ensure that Δf is less than the DOF of the lens.

$$\Delta f < DOF \tag{4.45}$$

$$|\frac{df}{d\lambda}\Delta\lambda| < DOF \tag{4.46}$$

$$\frac{\Delta\lambda}{\lambda} < \frac{DOF}{\lambda|\frac{df}{d\lambda}|} \tag{4.47}$$

This maximum tolerable bandwidth is plotted as a function of the numerical aperture (see figure 4-34 (a)). This shows that for NA = 0.85, the maximum source bandwidth allowed is about 0.1%. For a wavelength of 193 nm, this means that the wavelength spread must be less than 0.193 nm or 193 pm. For commercially available 193 nm sources, the bandwidth (95% energy integral) is less than 1.5 pm [37]. So the requirement on the source bandwidth is less stringent for ZPAL and easily met with current sources.



(a) (b)

Figure 4-34: *(a) Maximum source bandwidth as a function of zone-plate numerical aperture (NA). Source spectrum tends to blur the PSF of the zone plate. Here, we calculate the maximum source bandwidth such that the blurring along the optical axis is within the DOF of the original spot. As expected, this value decreases with increasing NA. (b) Maximum source bandwidth as a function of the inverse of the number of zones.*

In conclusion, the source-bandwidth requirements for ZPAL are significantly less stringent than for optical projection lithography. A very important consequence of this result is that laser sources for ZPAL can have more relaxed specifications, and hence much lower costs.

Effect of Oblique Incidence

The angle at which the incident beam hits the zone plate is important. With zone plates, as with lenses, the focused spot is laterally displaced when illuminated obliquely. This effect could be unintentional, and could arise from pointing errors present in the multiplexing device (such as is the case for the TI-DMD). As long as this error is small, deterministic, and repeatable, it will only result in a lateral shift of the printed pattern, which can be corrected in pre-processing. This swiveling effect could also be used intentionally, in order to dynamically correct for scanning-stage errors.

The displacement of the spot from the optical axis for an incident angle θ, and for a zone

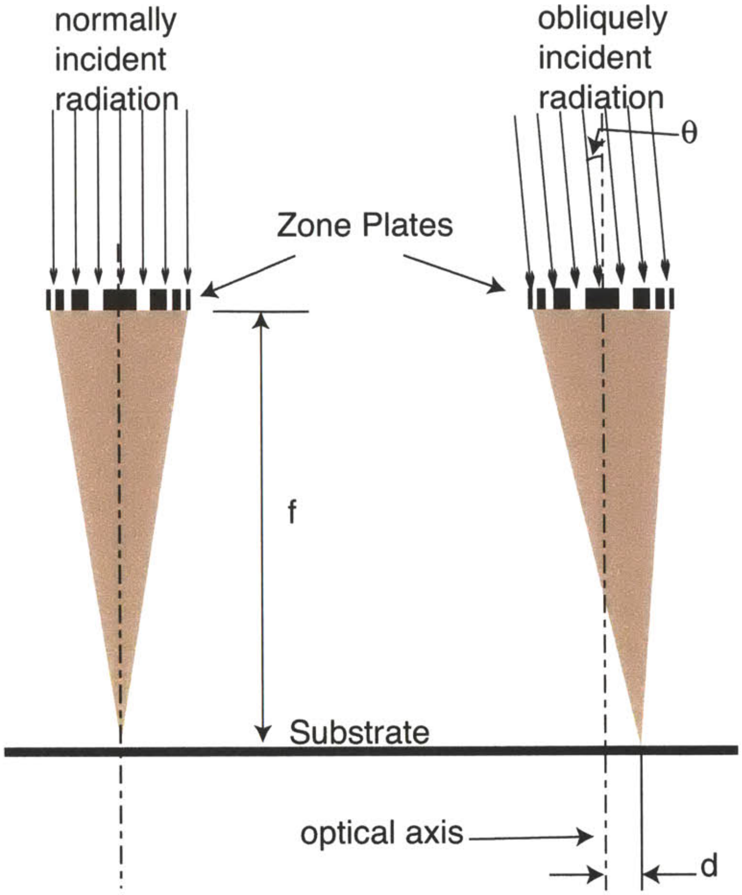


Figure 4-35: *With zone plates, as is the case with lenses, the focused spot can be swiveled by illuminating the zone plates obliquely.*

plate with focal length, f is given by

$$d = f\tan(\theta) \approx \theta \tag{4.48}$$

for small θ. This is illustrated in figure 4-35.

One of the consequences of oblique illumination of a zone plate, is the presence of off-axis aberrations into the focused spot (primarily spherical aberrations and coma). These aberrations will decrease image contrast. Figure 4-36 shows a plot of image contrast for dense lines and spaces as a function of the angle of illumination. The simulation takes into account all the aberrations present.

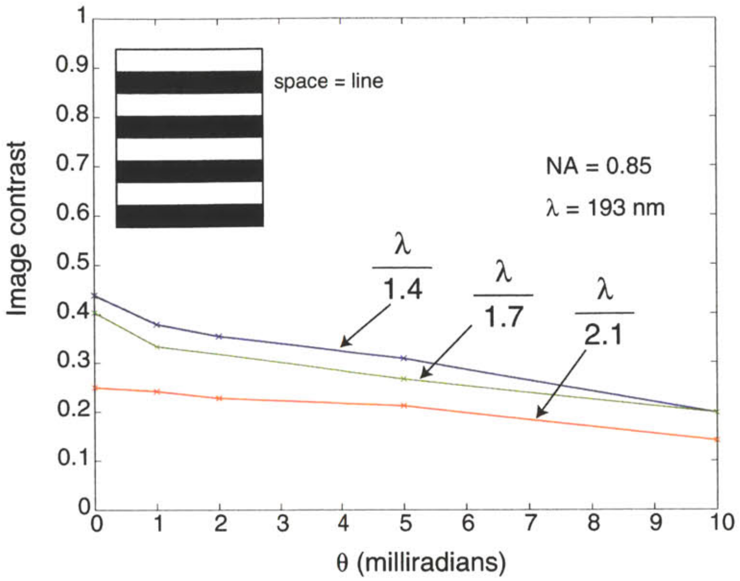


Figure 4-36: *Image contrast as a function of the angle of illumination for dense gratings of various linewidths. These simulations were performed with the PSF of a NA=0.85 zone plate with an illuminating wavelength of 193 nm.*

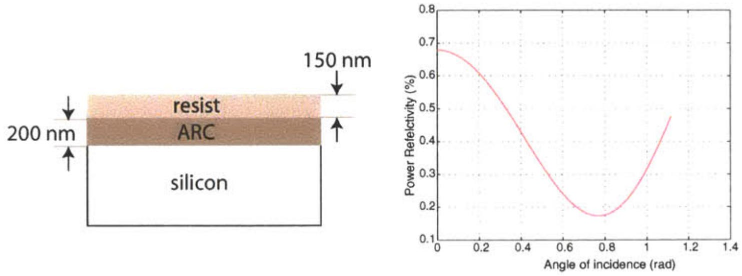
In conclusion, the image contrast decreases with off-axis illumination. In order to print 90 nm lines and spaces at $\lambda = 193\,\text{nm}$, the illumination angle should remain smaller than 5 mrad. For a zone plate with a 50μm focal length, this angle would imply a lateral displacement of the spot by as much as 250 nm. Therefore, this method can be a powerful technique for real-time correction of stage-placement errors.

4.4 Lithographic Performance of Zone Plates: Experiments

The ZPAL system being built at MIT was used to perform extensive characterization experiments in the course of this research. In this section, we begin by describing the experimental procedure and then, present quantitative results, which characterize the lithographic performance of zone plates.

4.4.1 Experimental Process

All the lithography experiments were performed on 3" silicon wafers. The radiation source in the ZPAL system has a wavelength of 400 nm. The photoresist used for the experiments was Shipley 1813, thinned using a P-type thinner in the ratio 1:2. In order to prevent back-reflections from the resist-silicon interface, which produce standing wave patterns in resist, an anti-reflection coating (ARC) was used. The thickness of the stack of resist and ARC was calculated such that the reflection is kept below 1% for the angles of interest. Zone plates of numerical apertures ranging from 0.7 to 0.95 were used in the experiments. Hence, the angles of incidence of the rays vary from 0 to $\sin^{-1}(0.9) = 64°$.



(a) (b)

Figure 4-37: *Resist and ARC stack used in the experiments. (a) Schematic of the stack consisting of 150 nm of S1813 photoresist and 200 nm of BarLi ARC. (b) Plot of power reflectivity as a function of angle for the stack in (a). Note that the ARC ensures that the reflectivity is below 1% for all the angles of interest. This calculation was performed using software authored by Michael Walsh at MIT.*

Figure 4-37 (a) illustrates the stack used in these experiments. A 200 nm thick film of

BarLi ARC was spin-coated on a 3" silicon wafer (spun at 3000 rpm) and then baked on a hot plate at 175° for 90s. A 150 nm thick film of S1813 photoresist was spin-coated on top of the ARC (spun at 5000 rpm) and then baked on a hot plate at 90°C for 60s. In figure 4-37(b), the power reflectivity is plotted as a function of the incident angle for this stack. Note that the power reflected from the ARC into the resist is kept below 1% for all the angles of interest. The calculations were performed using software written by Michael Walsh at the NanoStructures laboratory.

The silicon wafer was mounted on a vacuum pin-chuck. Then, the zone-plate array was brought into close proximity to the substrate. Using a He-Ne laser and a Michelson interferometer setup, the two surfaces were parallelized. A picomotor (described in section 2.1.4) was used to bring the zone-plate array into proper gap with respect to the substrate. The confocal back-signal from satellite zone plates was used to set the zone-plate array in focus. This technique is described in detail in chapter 6. Suffice it to say now that the confocal signal enables us to set the gap between the zone-plate array and the silicon wafer to within the depth-of-focus of the zone plates. After this initial setup, custom software was engaged to control the motion stage. The experiments presented in this thesis were performed without the multiplexing device i.e. all the zone plates in the array were writing the same pattern and they were shuttered in unison using a conpotics shutter. The exposure dose was controlled by changing the dwell time of the spot in a particular location. Since the scan involved a continuous motion of the stage, the dwell time was controlled by controlling the stage speed.

Test patterns were drawn using Nanowriter, which is a pattern-generation software authored by Dr. Mitchell Meinhold at the NanoStructures laboratory in MIT. These patterns were pre-processed to convert the format so as to be recognized by the Labview control software. Custom software in Matlab was written to enable this pattern conversion.

After exposure, the substrate was developed in a diluted 351 developer (1:1 with water) for 45s. A thin layer of gold-palladium[6] was sputtered on the sample to increase the image contrast in the scanning-electron microscope. The coated sample was inspected in a scanning-

---

[6] The sample was kept in the sputterer for about 20s.

electron microscope. Feature-size measurements can be made to a precision within ±10 nm very reliably. Almost all the results presented in this work were top-down views of resist patterns on top of ARC on a silicon wafer. These experimental results are presented next.

4.4.2 Experimental determination of Point-Spread Function (PSF)

The point-spread function (PSF) of the zone plate determines all the lithographic qualities of the written patterns. In the previous chapter, several techniques to simulate the PSF were described. Here, we describe an experimental technique using lithography to determine the PSF. Light exposure to photoresist induces a highly non-linear response. Essentially, one can approximate this reponse as a binary *clipping* of the image on the resist. This clipping level corresponds to the upper resist threshold, E_0 described in section (1.1.1). Single-shot spots were printed at various exposure doses in photoresist. Then, the radii of these spots were measured, and tabulated as a function of the exposure dose. At any exposure dose, the photoresist *samples* the radius of the PSF as illustrated in figure 4-38.

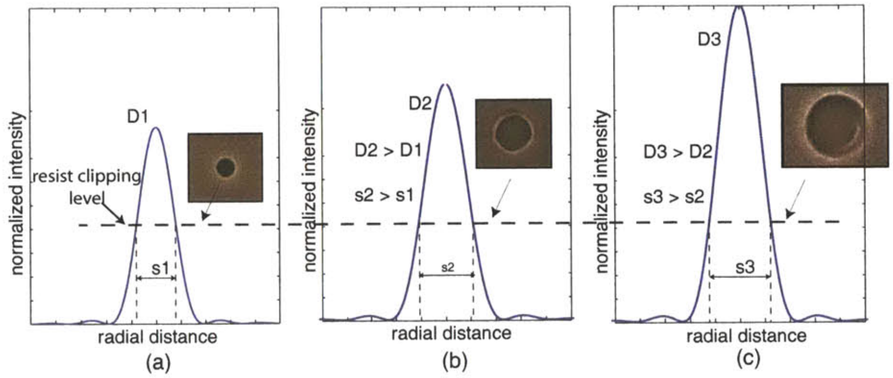


Figure 4-38: *Experimental process for determining the PSF using lithography. Single-spot exposures at different doses were conducted. The exposure dose in (a), D1 is less than that in (b), D2, which is less than that in (c), D3. Assuming a single clipping response of the photoresist, the exposed spots have diameters, s1, s2 and s3 respectively. The insets show scanning electron micrographs illustrating the exposed spots. After development, the radii of these spots are measured and used with the exposure-dose data to determine the PSF.*

In order to sample the PSF accurately and efficiently, the exposure doses must be calculated carefully. An exponential variation of dose such that each dose value was increased by 30% of the previous value was used in the experiments. The underlying assumption was that the change in linewidth due to a 30% change in dose is small enough that the PSF could be accurately reconstructed i.e. inaccuracies due to undersampling are avoided. Twenty five doses ranging from 2 to 1086 ms were used in the experiments. This ensured sufficient dynamic range of the exposure doses with a practical number of exposures. The doses were given by

$$d_n = 2 \times (1.3)^{n-1} \text{ for } n = 1 \text{ to } 25 \tag{4.49}$$

In figure 4-39, the exposure dose values are shown. It is apparent from the plot that the

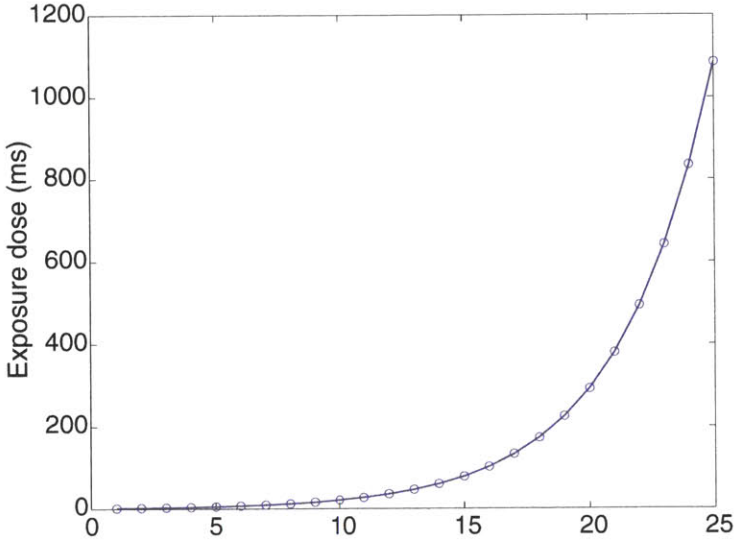


Figure 4-39: *Exposure-dose samples for the experimental determination of PSF. An exponential function of the doses is used for efficient sampling of the PSF.*

sampling doses are closer together at lower doses. This is because all the important structure in the PSF is close to its center.

The twenty-five spots were exposed and their radii measured very carefully. The reciprocal of the dose was plotted as a function of the measured spot-radius. The experimental data was scaled such that one data point matched the corresponding point in the simulated PSF. This experiment was performed for zone plates of different NAs. This PSF is then plotted together with the theoretical PSF for NA=0.85, 0.8 and 0.7 zone plates in figure 4-40. The PSF is plotted in log scale on the right to show the remarkable agreement between the theory and experiment.

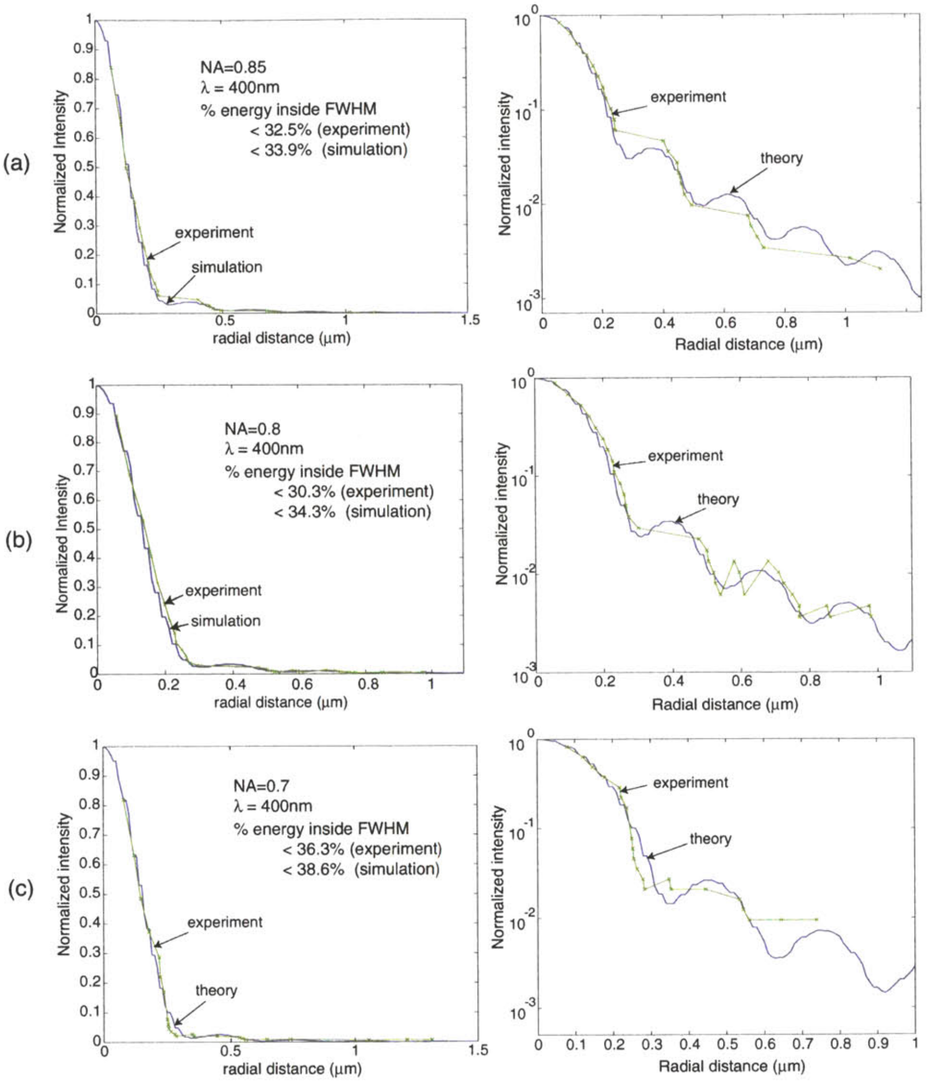


Figure 4-40: *Point-spread functions of NA=0.85, 0.8 and 0.7 zone plates. The experimentally determined PSF is plotted along with the theoretical one for zone plates of NA=0.85 (a), NA=0.8 (b), and NA=0.7 (c). On the right, the same data are plotted in log-scale to emphasize the excellent agreement between experiment and theory.*

The diffraction efficiency into the focussed spot can be calculated as the ratio of the energy inside the spot to the total energy in the focal plane. Using the full-width at half-maximum (FWHM) as the definition of the spot-size, the efficiency can be written as

$$\eta = \frac{\int_{r=0}^{r_{fwhm}} PSF(r) r dr}{\int_{r=0}^{\infty} PSF(r) r dr} \tag{4.50}$$

where r_{fwhm} is the radius of the focussed spot. Only an upper bound for efficiency can be obtained from the experimental data since the data does not extend to infinity. The efficiency values calculated using the experimental data as a function of NA is shown in figure 4-41. The theoretical efficiency is also plotted for comparison. Note that the theory and experiment show excellent agreement. The efficiency decreases with increasing NA. The reasons for this decrease were discussed in detail in chapter 2.

In conclusion, the close agreement between the experimentally determined PSF and the theoretical demonstrates that zone plates that perform very close to theoretical limits have been fabricated. The remarkable agreement also serves as a verification of the theoretical models for calculating the PSF. The ability to accurately model the PSF is extremely important in ZPAL since this enables the accurate prediction of exposed patterns. Knowledge of the PSF is critical for proximity-effect correction, which is necessary for linewidth control and edge-placement. This is the subject of the next chapter.

4.4.3 Experimental Determination of Exposure Latitude

Exposure latitude is an important metric of a lithography system because it directly affects the ability to print features reliably. In order to experimentally measure the exposure latitude, dense line/space patterns were exposed at different doses. After development, the linewidths were measured. From the plot of linewidth as a function of the exposure, the exposure latitude was calculated. This experiment was repeated with gratings of different periods in order to study the effect of feature sizes on exposure latitude. Here, the exposure latitude is defined as the change in exposure dose that can cause a 10% change in the printed linewidth.

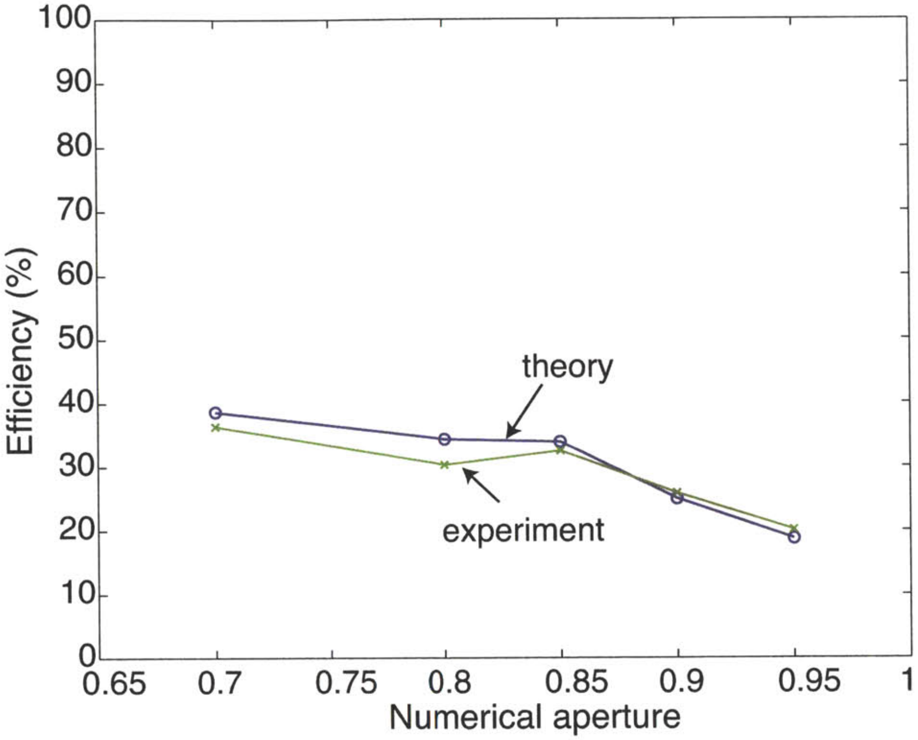


Figure 4-41: *Focussing efficiency as a function of NA. The ratio of the energy inside the FWHM of the focussed spot to the total energy in the focal plane was calculated from the experimentally obtained PSFs. This efficiency was plotted as a function of NA. The corresponding values from the theoretical calculations are shown for comparison. The agreement between theory and experiment is within the range of experimental error.*

The results of the experiment for a zone plate of NA=0.9 are shown in figure 4-43. The linewidths are measured near the center of the grating. In figure 4-43 (a), the linewidth is plotted as a function of the exposure dose for gratings of different periods and s-factor (ratio of exposure-grid size to the spot-size) of 1. In figure 4-43 (b), the s-factor is equal to 2. The exposure latitude was calculated from this data and plotted as a function of the grating period in figure 4-43 (c). The exposure latitude decreases as the grating period decreases. For small periods (340 nm), the exposure latitude is about 6%.

The experiment was repeated for zone plates of NA0.85 and NA=0.8, and the results are shown in figures 4-44, and 4-45 respectively. It is interesting to note that the highest exposure latitude for the smallest period gratings is achieved for a NA=0.85 zone plate. It was shown earlier that the exposure latitude depends on the normalized slope of the image

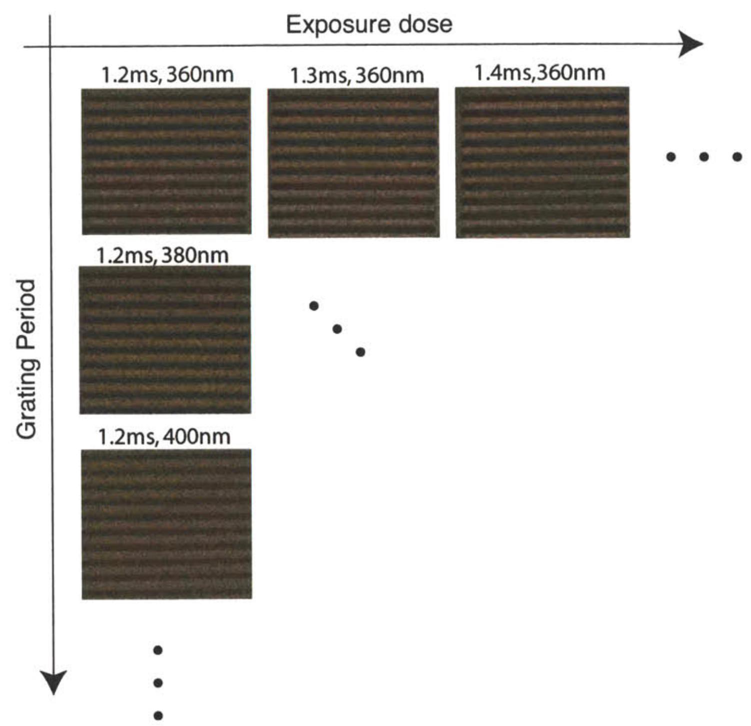
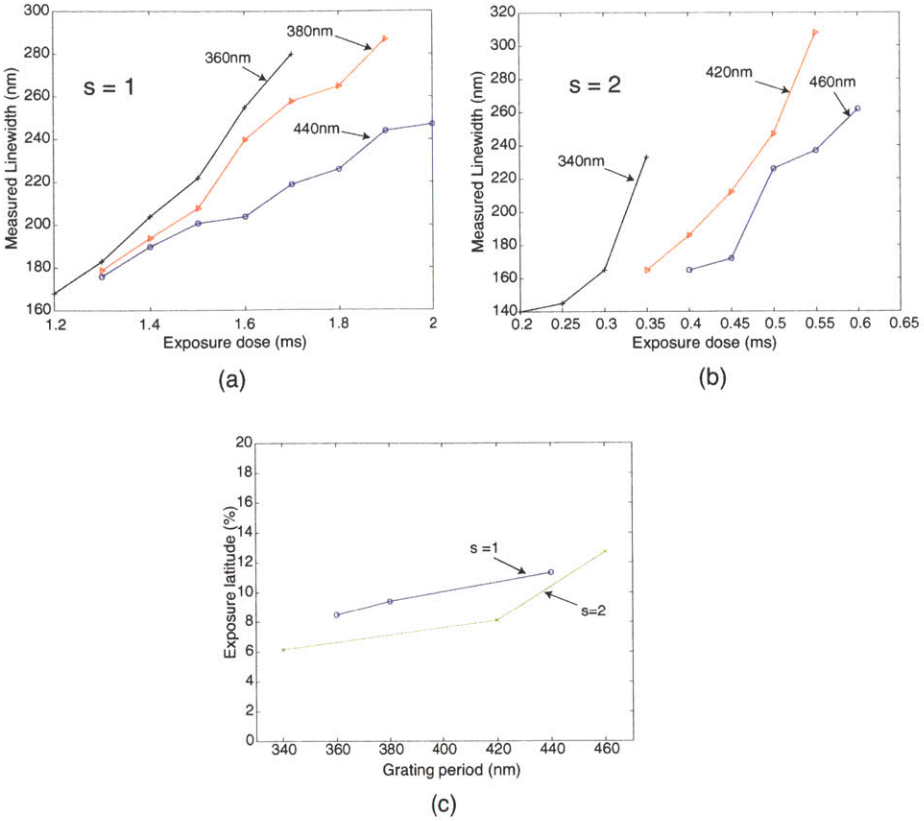


Figure 4-42: *Experiment for determination the exposure latitude. Scanning- electron micrographs of dense gratings are shown. The two variables that were changed during the experiment were the period of the grating and the exposure dose. For each pair of period and dose, the linewidth in the center of the grating was measured. The exposure latitude can be calculated from these measurements. The results shown here are for* $s = 1$. *The experiment was repeated with* $s = 2$ *as well.*

profile at the resist threshold (see section 4.1.1). Contrast simulations indicated that this slope is maximum for $NA \approx 0.85$ (see section 4.3.1). This agrees well with our experimental results which indicate that the exposure latitude is maximum for NA=0.85 (for the smallest lines).

Exposure latitude is a measure of the reliability with which the patterns may be written. In that sense, it is a good measure of the resolution of the system. If we assume that an exposure latitude of at least 10% is required to print features reliably, the resolution (from the experimental results) is about 150 nm with 0.85 NA zone plate. Figure4-46 show scanning-electron micrographs of dense gratings written with a 0.85NA zone plate and $s = 1$. Since the exposure wavelength was 400 nm, the lines and spaces are subwavelength.

It must be pointed out that the smallest lines which can be written with $s = 2$ is slightly



(a)

(b)

(c)

Figure 4-43: *Experimental determination of exposure latitude for a NA=0.9 zone plate. Dense gratings of different periods were exposed at different doses and s-factors. (a) Measured linewidth as a function of exposure dose for s=1. (b) Measured linewidth as a function of exposure dose for s=2. (c) Exposure latitude (in %) as a function of the grating period. The smallest lines that can be printed with NA=0.9 and at least 10% exposure latitude is about 180 nm. These experiments were perfomed with* $\lambda = 400\,nm$ *and a zone plate of NA=0.9.*

bigger than the ones with $s = 1$. This is expected from the earlier analysis which showed that increasing s decreases image contrast. In figure 4-47, gratings of different periods are shown which were written with NA=0.85, $\lambda = 400\,\text{nm}$, and $s = 2$. The smallest lines in this case is about 180 nm.

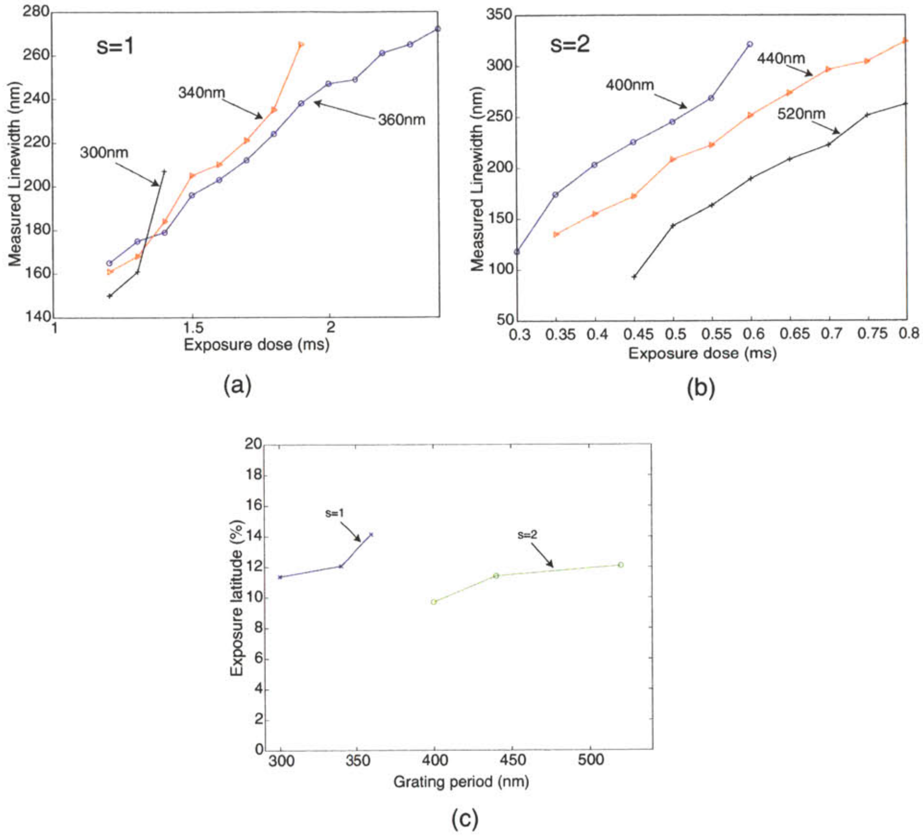
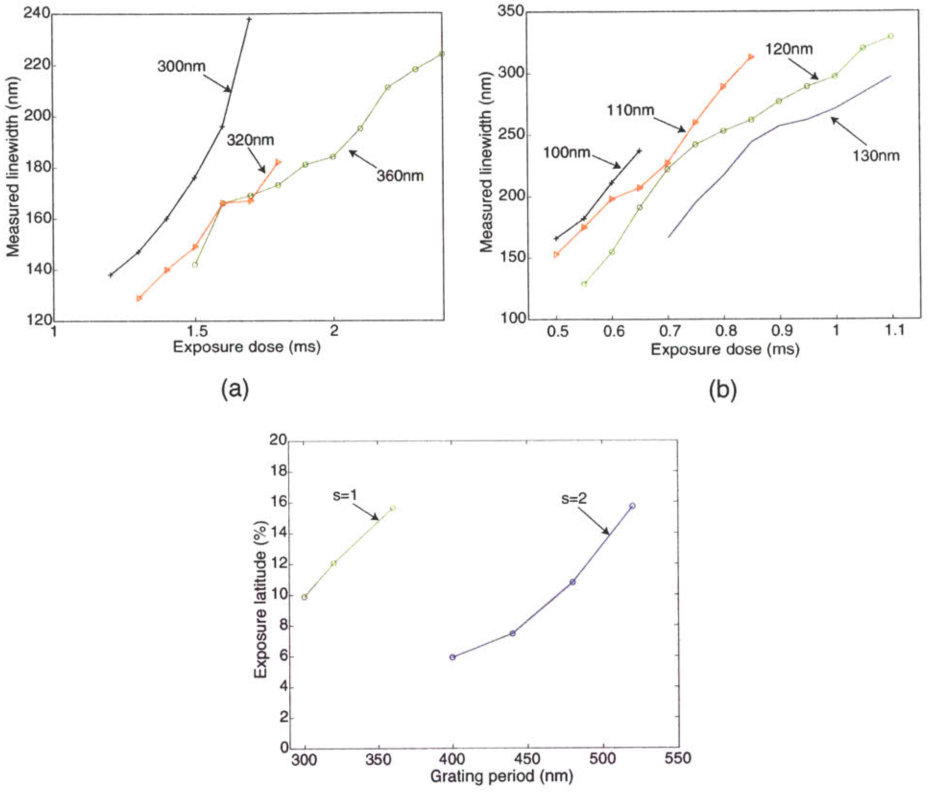


Figure 4-44: *Experimental determination of exposure latitude for a NA=0.85 zone plate. Dense gratings of different periods were exposed at different doses and s-factors. (a) Measured linewidth as a function of exposure dose for s=1. (b) Measured linewidth as a function of exposure dose for s=2. (c) Exposure latitude (in %) as a function of the grating period. The smallest lines that can be printed with NA=0.85 and at least 10% exposure latitude is about 150 nm. These experiments were perfomed with* $\lambda = 400\,nm$ *and a zone plate of NA=0.85.*

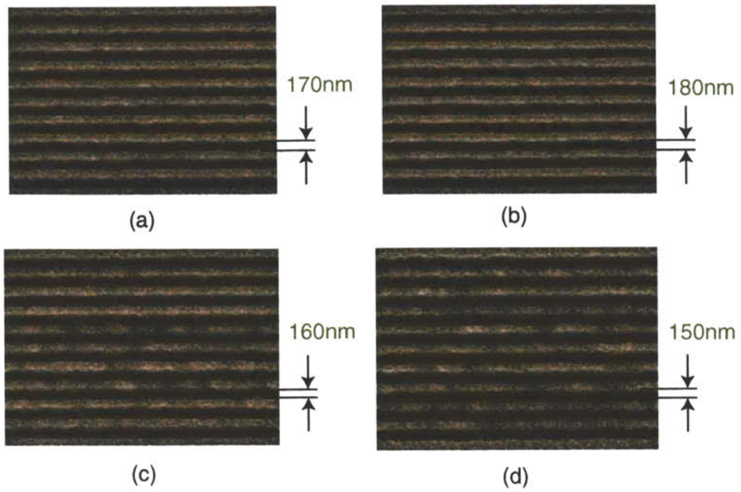


(a)

(b)

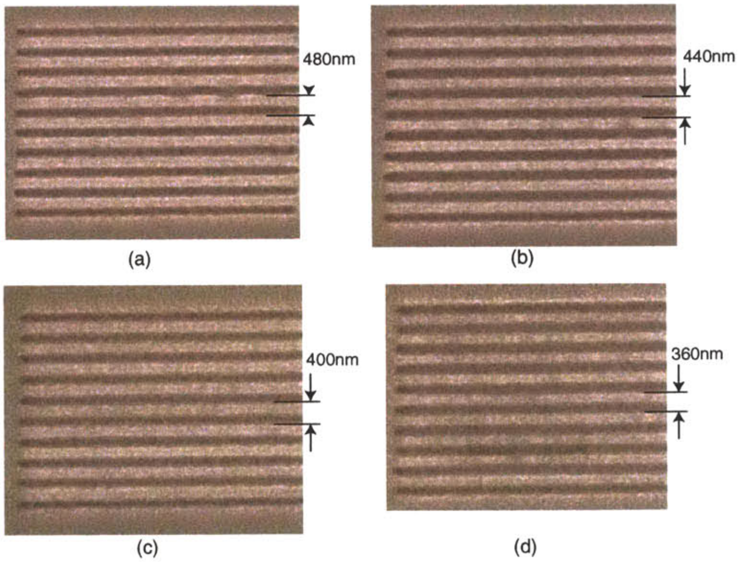
(c)

Figure 4-45: *Experimental determination of exposure latitude for a NA=0.8 zone plate. Dense gratings of different periods were exposed at different doses and s-factors. (a) Measured linewidth as a function of exposure dose for s=1. (b) Measured linewidth as a function of exposure dose for s=2. (c) Exposure latitude (in %) as a function of the grating period. The smallest lines that can be printed with NA=0.8 and at least 10% exposure latitude is about 160 nm. These experiments were perfomed with* $\lambda = 400\,nm$ *and a zone plate of NA=0.8.*



(a) (b) (c) (d)

Figure 4-46: *Smallest period gratings with s=1. Scanning-electron micrographs of gratings exposed using NA=0.85 zone plate,* $\lambda = 400$ *nm, and s=1. The grating periods are (a)360 nm, (b) 340 nm, (c) 320 nm, and (d) 300 nm.*



(a) (b) (c) (d)

Figure 4-47: *Smallest period gratings with s=2. Scanning-electron micrographs of gratings exposed using NA=0.85 zone plate,* $\lambda = 400$ *nm, and s=2. The grating periods are (a)480 nm, (b) 440 nm, (c) 400 nm, and (d) 360 nm.*

4.4.4 Experimental Determination of Image Contrast

Image contrast was discussed earlier and its importance to resolution was emphasized. In this section, we describe a novel method to determine the aerial-image contrast using the same experimental results obtained in the previous section. The experiment involved exposing dense grating lines at different doses. The image contrast can be calculated as follows. Figure 4-48 shows a schematic of the cross-section through the image intensity of a grating pattern. The intensity level, E_2 corresponds to that resist threshold level which results in equal lines and spaces is given by and given by

$$E_2 = \frac{I_{max} + I_{min}}{2} \tag{4.51}$$

Assuming a binary clipping due to the photoresist, the clipping levels are inversely proportional to the exposure doses. At the level corresponding to I_{max}, the pattern would be

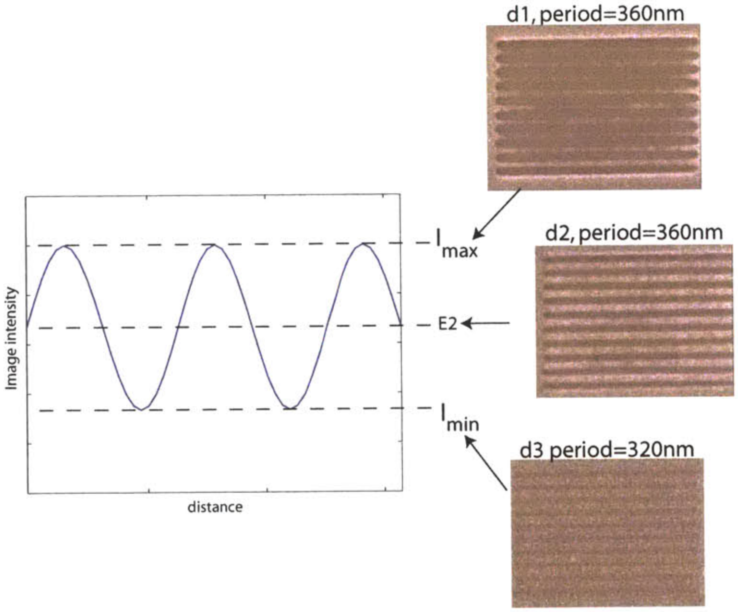


Figure 4-48: *Determining contrast from grating exposures. Cross-section through the intensity profile of the image of a grating is shown. The doses correspond to inversely to the intensity levels. At dose, d_1 (I_{max}), the pattern is almost totally overexposed. At dose, d_2 (E_2), the pattern consists of equal lines and spaces. At dose, d_3 (I_{min}), the pattern is almost totally underexposed. The contrast can be derived from a knowledge of at least two of the three levels shown.*

completely washed out in resist. Let us denote the corresponding dose as d_1. At the level corresponding to E_2, the pattern would consist of equal lines and spaces in resist. The corresponding dose is denoted d_2. At the level corresponding to E_3, the pattern would be barely unexposed. The corresponding dose is d_3. By exposing grating patterns at many different doses, at least two of the doses, d_1, d_2, d_3 can be known[7]. The image contrast can then be calculated as

$$C = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} \tag{4.52}$$

In terms of d_2 and d_1, the contrast is given by

$$C = 1 - \frac{d_2}{d_1} \tag{4.53}$$

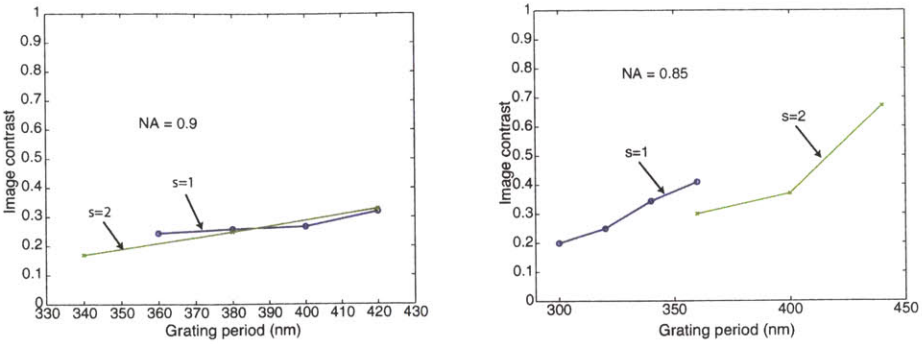
In terms of d_2 and d_3, the contrast is given by

$$C = \frac{d_2}{d_3} - 1 \tag{4.54}$$

By inspecting the scanning-electron micrographs of gratings patterned at different doses, and noting the doses corresponding to total overexposure (d_1), equal lines and spaces (d_2), or total underexposure (d_3), the image contrast can then be calculated. The image contrast is plotted as a function of the grating period in figure 4-49(a) for NA=0.9 and in figure 4-49(b) for NA=0.85. For each case, the contrast was calculated for s=1 and s=2. In both cases, the contrast decreases as the grating period decreases. It is observed that in both cases, the patterns are not well defined for contrast values less than about 0.25. This agrees well with the minimum resolvable image contrast (see section 1.1.1) corresponding to a resist contrast, γ of 2. The resist contrast for Shipley 1813 was experimentally determined and was described earlier in this chapter. In figure 4-49(b), it can be seen that s=1 has higher contrast than s=2 and hence, one can print gratings of smaller periods.

---

[7] All three variables can be known. However, the range of doses used in the experiment only allowed us to obtain either one of d_1 or d_3.



(a) (b)

Figure 4-49: *Image contrast as a function of grating period. The image contrast is plotted as a function of the grating period for (a)NA=0.9 zone plate and (b)NA=0.85 zone plate. The parameters used in the experiment were* $\lambda = 400\,nm$ *and* $s = 1, 2$. *Note that the highest contrast is obtained at NA=0.85 as predicted by the simulations. The minimum resolvable contrast of 0.25 agrees very well with the calculated value based on the measured resist contrast,*γ *of 2.*

In conclusion, the finest resolution of the system is attained when NA=0.85. Based on contrast values, this resolution is about 150 nm. This is the same resolution given by the exposure latitude measurements presented in the previous section.

4.5 Stitching in ZPAL

In ZPAL, a large pattern is built-up by tiling together smaller fields (unit-cells) written by individual zone plates. This process is known as *stitching.* Pattern-placement errors might occur at the boundaries between these fields. These are known as stitching errors.

An example will illustrate the concept of stitching errors. In figure 4-50, the zone plates wrote a pattern consisting of horizontal lines and spaces as shown in the bottom row. If the centers of the zone plates (shown on top) are misaligned for some reason, then the written lines will not match up at the boundary. This gives rise to stitching errors. The misalignment in the centers of the zone plates in a zone-plate array (ZPA) can arise from either one of two causes.

The first one is an actual placement error during lithography while fabricating the ZPA. In the case of scanning-electron beam-lithography, this placement error can be quite small

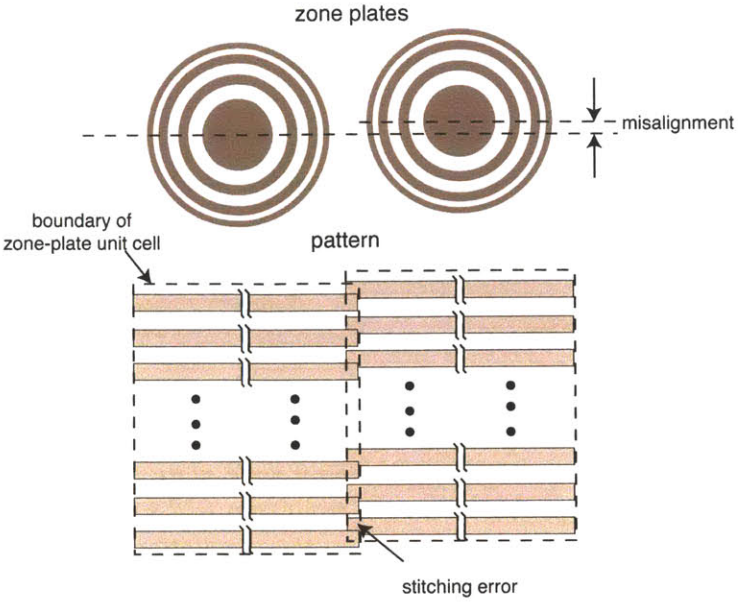


Figure 4-50: *Schematic illustrating stitching errors. Zone plates (on top) wrote a pattern (on bottom) consisting of horizontal lines and spaces spanning the entire unit-cell. If there is a misalignment in the centers of the zone plates, the underlying pattern will not match up at the boundaries of the unit-cell. This causes stitching errors.*

(a few tens of nanometers). Using a novel technique known as spatial-phase-locked e-beam lithography, this placement error can be brought down to about 2 nm or lower [30, 38]. In this case, the error is fixed and deterministic. This error can be measured by patterning gratings and then, measuring the corresponding stitching errors in each unit-cell. Another more efficient technique to measure this error is decribed in chapter 6. Once, this error is know for any ZPA, the pattern may be appropriately distorted prior to writing such that the written pattern shows no stitching errors.

The second cause for the misalignment is the relative rotation of the axes of the ZPA with respect to the axes of the scanning stage. The effect of such a rotation is to give rise to stitching errors as illustrated in figure 4-51. This error can also be measured by patterning gratings over the entire unit-cell. This error is also fixed once the ZPA is fixed. Hence, the pattern can be distorted prior to writing.

In order to study the stitching errors, gratings of period 6μm (linewidth = 2μm) were written over the entire unit-cell using a ZPA containing 1000 by 1000 zone plates, each of 0.7

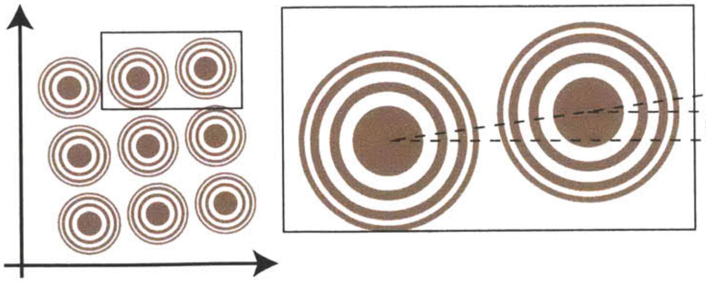


Figure 4-51: *Angular misalignment of ZPA and stage axes.*

NA operating at $\lambda = 400\,\text{nm}$. The stage stepping grid in this case was 250 nm. A scanning-electron micrograph of the result is shown in figure 4-52. The stitching error was found to be about $2\,\mu\text{m}$ in both X and Y directions. The error was measured at several different unit-cell boundaries and was found to be the same. This suggests that the error is mostly due to angular misalignment and not due to placement errors in the centers of the zone plates. The distance between the centers of two adjacent zone plates was equal to $100\,\mu\text{m}$. This implies that the ZPA was rotated by an angle equal to 0.02 radians in the clockwise direction with respect to the axes of the stage.

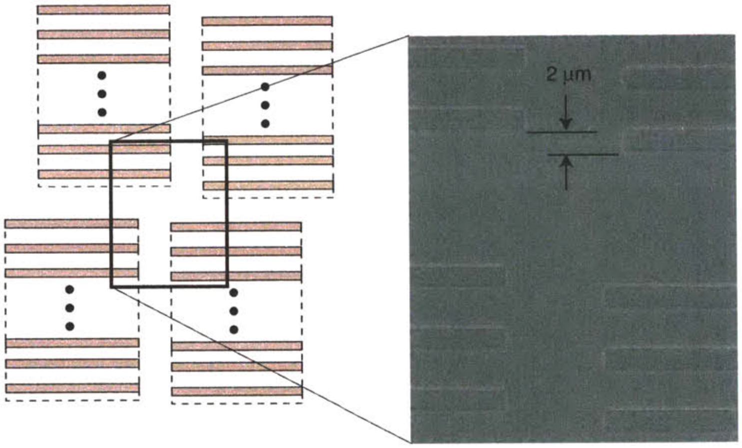
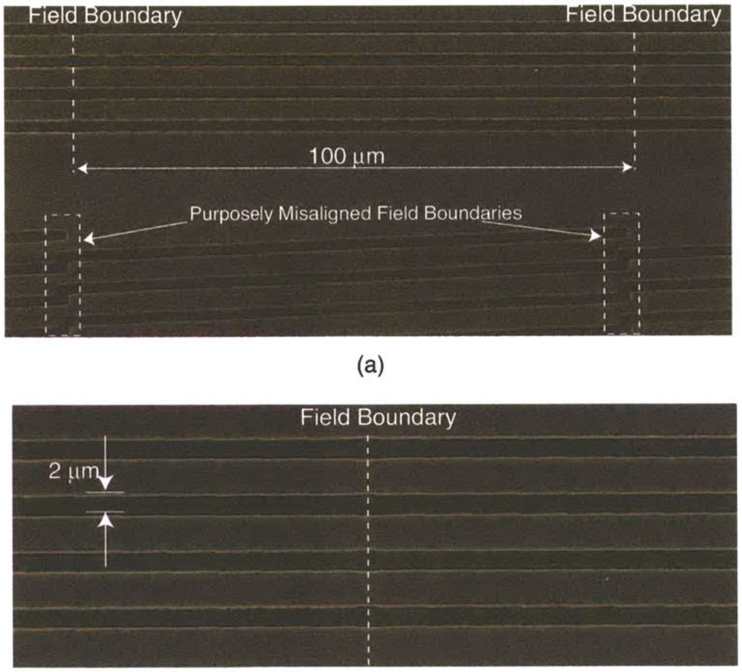


Figure 4-52: *Measuring the angular misalignment between the ZPA and the stage axes. A scanning-electron micrograph of the boundary of 4 unit-cells is shown. The lateral misalignment was measured to be $2\,\mu m$. The exposure was perfomed with a ZPA containing zone plates of NA=0.7 operating at $\lambda = 400\,nm$.*

In order to compensate for this rotation via software, the pattern was rotated in the clockwise direction by an angle equal to 0.02 radians. This ensures that the lines match up at the boundaries of the unit-cell. However, in order to accurately pattern rotated lines, the stage needs to step by a fraction of the spot-size, in this case, 100 nm. This is necessary to ensure that the discretization of the pattern into horizontal and vertical rectangles can accurately represent the slope of the printed line. The results of the experiment are shown in figure 4-53. The scanning-electron micrograph in figure 4-53 (b) shows one unit-cell boundary where the lines are stitched. The slight widening of the lines at the unit-cell boundaries is due to the fact there was a small (1 μm) overlap across the unit-cell boundaries and therefore the net exposure dose in these regions was higher. This can be corrected by dose-modulation methods described in chapter 5. The roughness in the line-edges is due to the discretization mentioned earlier. In practise, this method of correcting for angular



(a)

(b)

Figure 4-53: *Experimental results showing stitching across unit-cell boundaries. Scanning-electron micrographs of a group of four horizontal lines that were stitched across the unit-cell boundaries are shown. The pattern was exposed using a ZPA containing 1000 by 1000 zone plates each of NA=0.7 operating at* $\lambda = 400\,nm$. *Note that the corrected pattern was stitched properly across the boundaries.*

misalignment is very inefficient because it requires a small step-size in order to represent the sloped lines accurately and avoid the line-edge roughness seen in figure 4-53. This will slow down the writing process significantly. Angular misalignment can be corrected by means of alignment marks and careful (one-time) alignment of the ZPA with respect to the stage. Nevertheless, this experiment demonstrates the powerful capability of modifying the pattern in software in order to improve the pattern quality, which is the method of choice for small errors (such as the center to center placement of the zone plates). We will discuss this idea in greater detail regarding proximity-effect correction in chapter 5.

In conclusion, we have shown the feasibility of using software correction for stitching patterns across boundaries of unit cells.

4.6 Quality of patterning in ZPAL

In the concluding section of this chapter, we address the issue of the quality of the patterns produced by ZPAL. To this end, scanning-electron micrographs of a wide variety of patterns are shown. All these patterns were printed very reliably and are repeatable.

Curved Patterns

Most lithographic systems are optimized for *manhattan geometries*, which consist of only straight edges. It is becoming more common, especially in small niche research fields, to pattern curved structures. Since ZPAL is a raster-scan tool, curved patterns are broken down into rectangular boxes. We have written custom software to perform this transformation quite efficiently. If the rectangles are too big, then the approximation to the curve is not good. However, if the rectangles are too small, the number of rectangles become large and this slows down the patterning process. Thus, there is a fundamental trade-off between the accuracy of the curve and the writing speed. Nevertheless, in ZPAL, the presence of a large number of beams writing in parallel can offset the slow writing speed of a single beam.

In figure 4-54, a large variety of patterns with curved features are shown, which were written with ZPAL. Figure 4-54 (a) shows the scanning-electron micrograph of a 3-ring optical resonator device, where the resonators are racetrack-shaped. The associated waveguides were about 100μm long. This pattern was printed with a 0.9 NA zone plate at $\lambda = 400$ nm, and a $s \approx 2$ (step-size =125 nm). Figure 4-54 (b) shows the scanning-electron micrograph of an array of micro-rings, each has an inner diameter of about 1μm, and an outer diamter of 2μm. This pattern was printed with a 0.9 NA zone plate at $\lambda = 400$ nm, and a $s \approx 3$ (step-size=75 nm). Figure 4-54 (c) is a scanning-electron micrograph of a portion of a zone plate, whose outer-zone width is about 1.24μm. Note that curve is extremely smooth. This pattern was printed with a 0.7 NA zone plate at $\lambda = 442$ nm, and a $s \approx 4$ (step-size =75 nm). Figure 4-54 (d) is a scanning-electron micrograph of coupled waveguides. This pattern was printed with a 0.7 NA zone plate at $\lambda = 442$ nm, and a $s \approx 2$ (step-size =150 nm). Figure 4-54 (e) is a scanning-electron micrograph of a zone plate, whose outer-zone width is about 1.37μm. This pattern was printed with a 0.7 NA zone plate at $\lambda = 442$ nm, and a $s \approx 2$ (step-size =150 nm). The curve has rough edges due to the large rectangles used to approximate the curve.

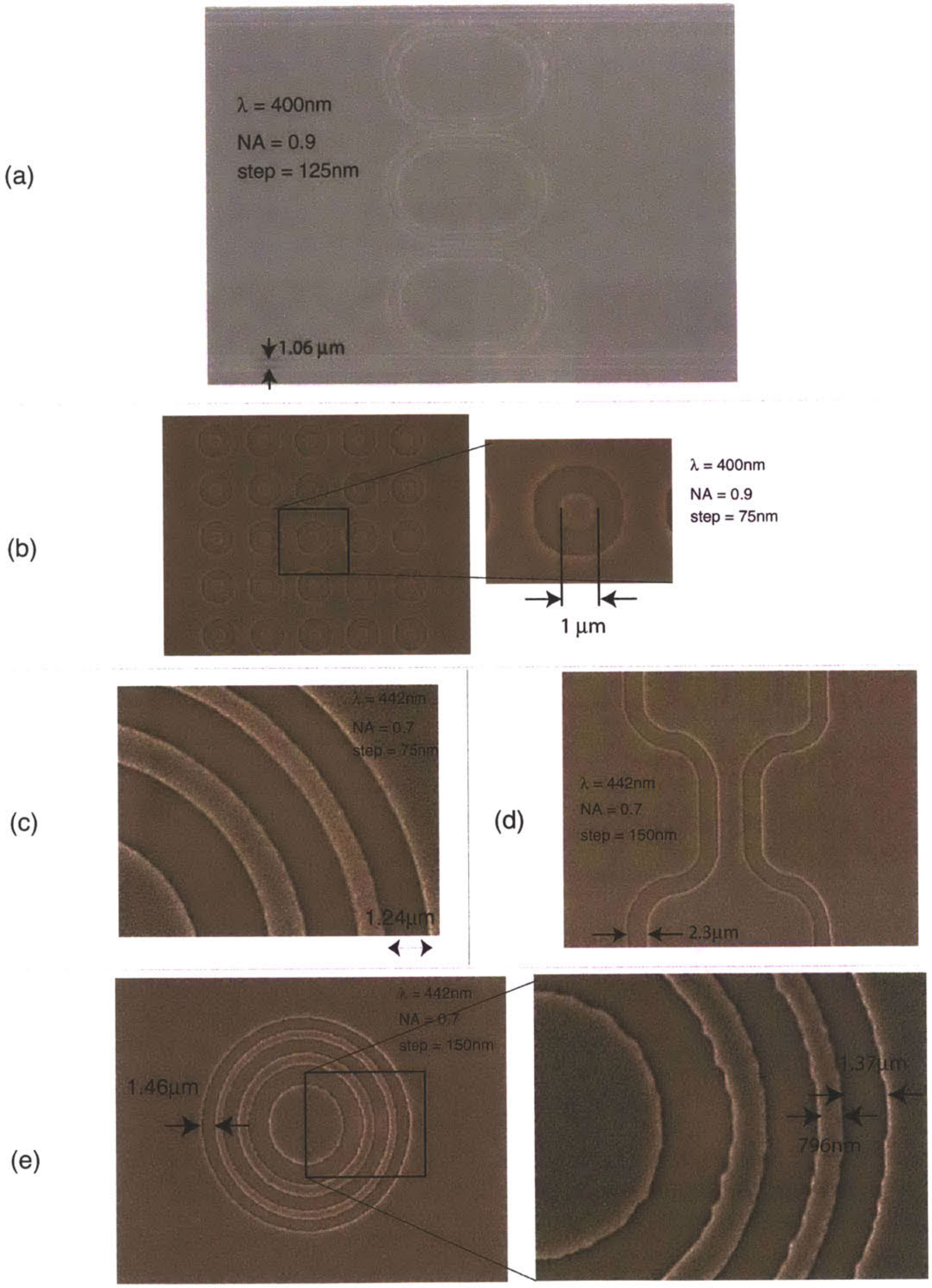
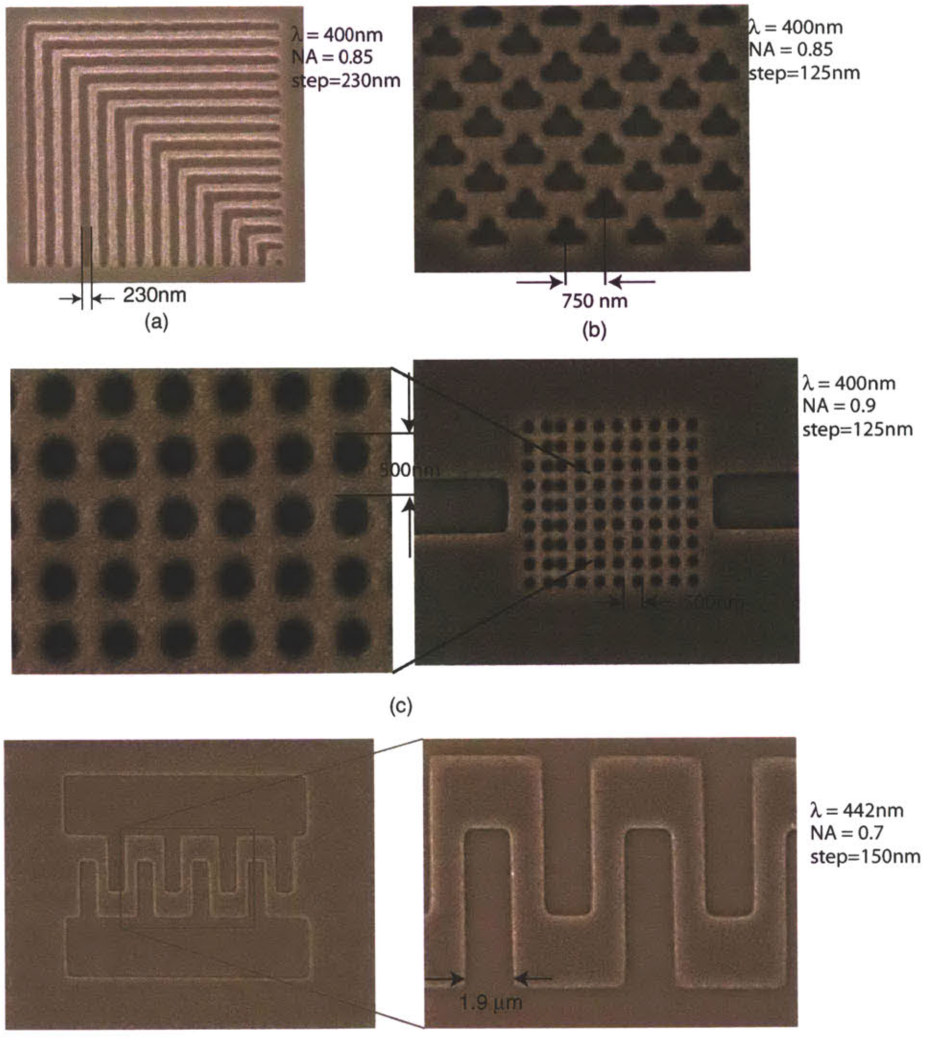


Figure 4-54: *Patterns with curves written with ZPAL.*

Miscallaneous Patterns

ZPAL is capable of patterning arbitrary geometry. This capability is limited only by the pattern-conversion software and the accuracy of the scanning stage. We have developed software that can translate arbitrary patterns into a format that is recognizable by the ZPAL hardware. This capability enables the patterning of a large variety of patterns as shown in figure 4-55.



(a) (b) (c) (d)

Figure 4-55: *Miscallaneous patterns written with ZPAL.*

In figure 4-55 (a), the scanning-electrom micrograph of a pattern consisting of dense nested-L shapes is shown. The period of these lines is about 460 nm. The pattern was exposed using a 0.85 NA zone plate, $\lambda = 400\,\text{nm}$ and a step-size of 230 nm, implying single-pixel lines ($s = 1$). In figure 4-55 (b), the scanning-electrom micrograph of a periodic pattern representing part of a DRAM device is shown. The features lie on a hexagonal lattice. This pattern was printed with a 0.85 NA zone plate at $\lambda = 400\,\text{nm}$, and a $s \approx 2$ (step-size=125 nm). In figure 4-55 (c), the scanning-electrom micrograph of a photonic bandgap structure is shown. A defect-line in the third column of holes from the left was intentionally introduced. This pattern was printed with a 0.9 NA zone plate at $\lambda = 400\,\text{nm}$, and $s \approx 2$ (step-size =125 nm). In figure 4-55 (d), the scanning-electrom micrograph of a microcomb structure is shown. This pattern was printed with a 0.7 NA zone plate at $\lambda = 442\,\text{nm}$, and $s \approx 2$ (step-size =150 nm). These patterns were written very reliably and repeatedly. This myriad of patterns showcases the flexibility of our maskless lithography system.

Finally, ZPAL enables the patterning of truly arbitrary patterns such as the potrait of one's advisor (figure 4-56)!

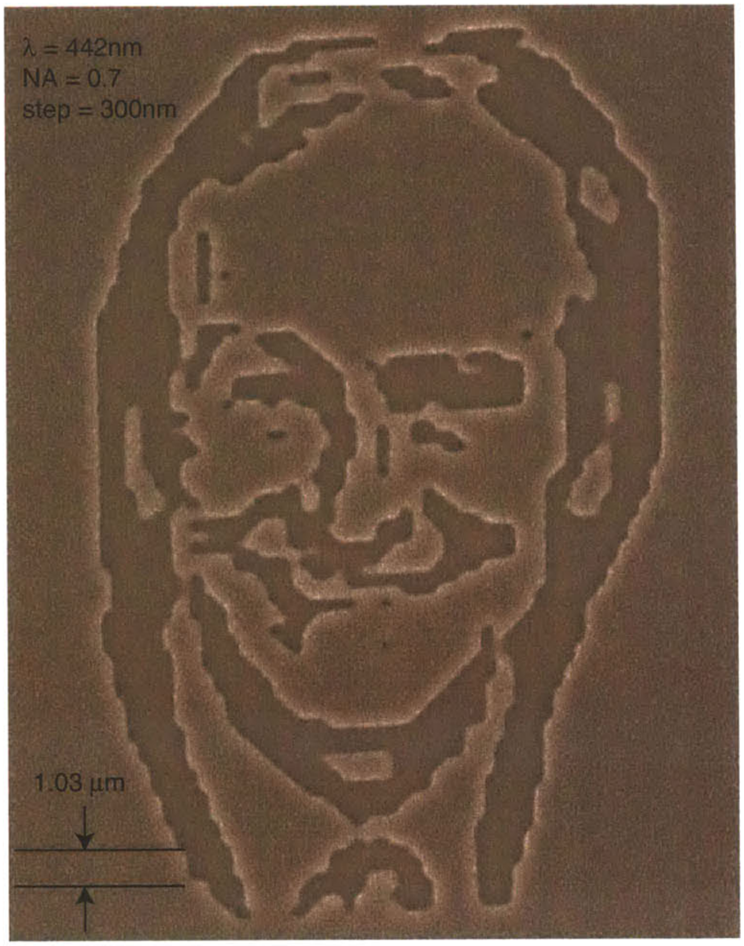
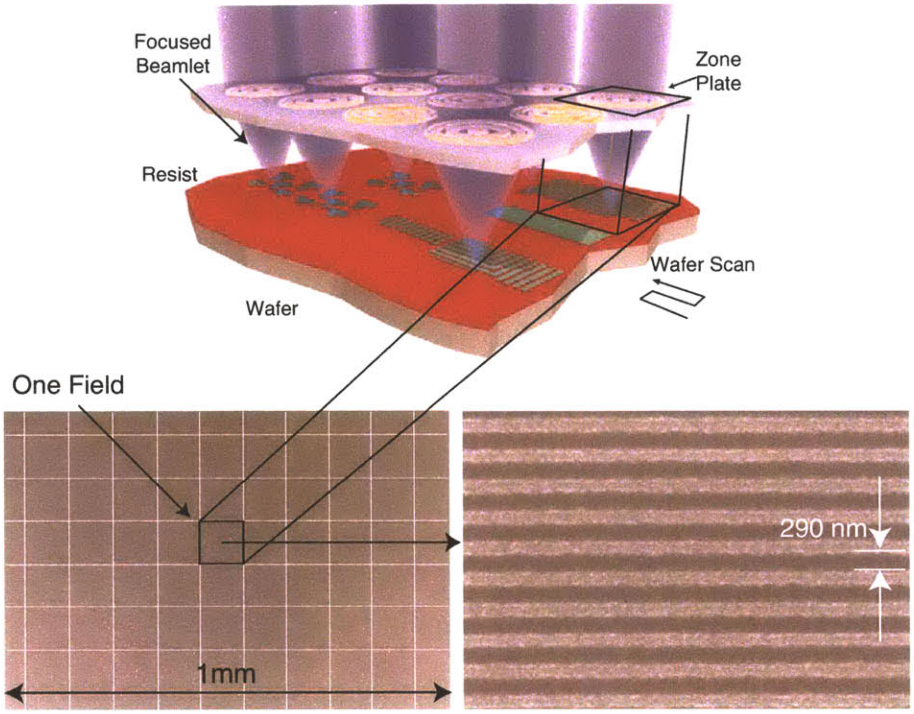


Figure 4-56: *Portrait patterned using ZPAL. The portrait of Professor Henry Smith was patterned using a 0.9 NA zone plate at* $\lambda = 400\,nm$ *and* $s \approx 1$ *(step- size =* $300\,nm$*). This pattern was written in May 2002. This showcases the power of a maskless patterning technique such as ZPAL.*

Large-Area Patterns

In ZPAL, each zone plate writes patterns within its unit cell. By tiling together different unit cells, large patterns are formed. Such patterns were written with an array of 1000 by 1000 zone plates that were fabricated in the NanoStructures Laboratory. Figure 4-57 shows the scanning-electron micrograph (with the ZPAL schematic on top) of a grating pattern (of period 600 nm) spanning an area of about 10 mm^2. These were patterned with a 0.7 NA zone plate at $\lambda = 400\,\text{nm}$ and step-size of 300 nm ($s = 1$).



(λ = 400nm, NA = 0.7, Focal Length = 40 μm)

Figure 4-57: *Full-field patterning with ZPAL. Top: Schematic of a ZPAL system. Each zone plate writes patterns within its field (unit-cell). Tiling together (also known as stitching) several such fields results in a large area pattern. Bottom: Scanning-electron micrograph of a dense grating (of period 600 nm) spanning about 10*mm^2 *area is shown. A magnified micrograph of the center of the grating reveals lines and spaces of good fidelity.*

In conclusion, we have demonstrated the feasibility of the lithographic patterning using spots focussed by a zone-plate array.

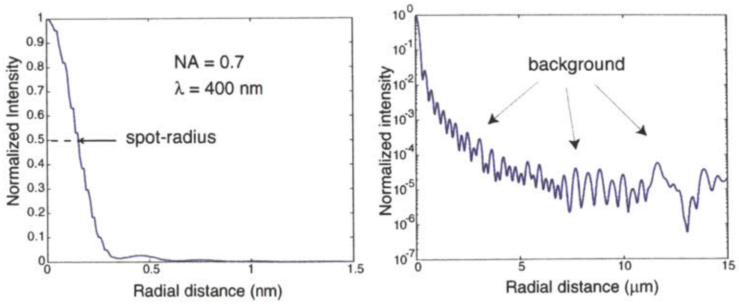
Chapter 5

Proximity Effects and their Correction

All forms of lithography exhibit deterioration in image quality due to reasons ranging from aberrations in the optics to diffraction. In this chapter, we are concerned with the pattern deformation caused by the close proximity of features. We begin by discussing the origin of proximity effects in zone-plate-array lithography (ZPAL). Various methods to compensate for these effects are also described. The methods were evaluated using simulations. Finally, experimental results are presented which showcase the performance of a newly developed proximity-effect correction method.

5.1 Origin of Proximity Effects in ZPAL

In a scanning-spot system such as ZPAL, a focussed spot is scanned to write desired patterns on a substrate. The intensity profile of the spot shows that although the intensity falls rapidly with distance from the center, the distribution extends over a relatively large distance. In this context, the radius of the spot is defined as the radius at which the intensity level is equal to one-half of the peak value, i.e. the full-width at half-maximum (FWHM). The simulated point-spread function (PSF) of a 0.7 NA zone plate at $\lambda = 400$ nm is shown in figure 5-1(a). In figure 5-1 (b), the PSF is plotted in log scale to emphasize the extent of the intensity distribution in. The intensity that is outside the spot-radius (the *tail* of the PSF) acts as background, contributing exposure dose to neighboring regions in resist.



(a)

(b)

Figure 5-1: *Background in the PSF. (a) A cross-section through the center of the simulated-focal spot of a 0.7 NA zone plate at* $\lambda = 400\,nm$. *The spot-size is twice the spot-radius; the radius measured at the half-maximum point. (b) The intensity is plotted in log-scale to emphasize the extent of the tail of the PSF. This tail acts as background to the desired dose (that is within the spot-radius), and delivers exposure dose to neighboring regions in resist.*

The areal image of the exposed pattern is an addition of the doses in the exposed spots comprising the pattern. Therefore, the addition of exposure dose in unintended regions of the photoresist is a cumulative process in ZPAL. When the spots are many and close together (i.e. dense features in the pattern), the cumulative dose tends to distort the exposed pattern (causing removal of unintended regions of resist). This is the *proximity effect* in ZPAL. An example will help illustrate this phenomenon. In figure 5-2(a), a binary pattern consisting of an isolated line and two groups of dense lines of different densities is shown. The desired linewidth is the same for all the lines. Figure 5-2(b) shows the simulated aerial image after exposure. The cross-section through the intensity profile of the image is shown in figure 5-2(c). As mentioned in the previous chapter, the response of the photoresist to this aerial image can be approximated as a simple thresholding function, i.e. all the regions of the resist that receive a total exposure dose higher than the resist threshold will be removed after development. With this model, the resist profile after development is shown in figure 5-2(d)[1]. It is seen that the width of the isolated line on the left, W is larger than the width of the semi-dense lines, W′. As seen from the intensity profile, this is due to the fact that

---

[1] The resist sidewall angle is not depicted in this schematic for the sake of simplicity.

the close proximity of the lines causes exposure dose to spill over into the spaces between the lines, and increases the linewidths. This phenomenon is so severe in the case of the lines of highest density that these lines were totally "washed out" in resist and not resolved at all. This shows that the proximity effects not only modify the intended linewidths of features but also affect the resolution of the lithography system! If there were a technique by which the intensity profile of the most dense lines could be reduced locally, then the lines could be resolved. In the next section, we describe three methods to achieve this goal. The conclusions drawn here are based on the assumption that the resist contrast is infinite. In practise, proximity-effect correction (PEC) provides the capability to achieve the resolution limits sets by the finite resist-contrast.

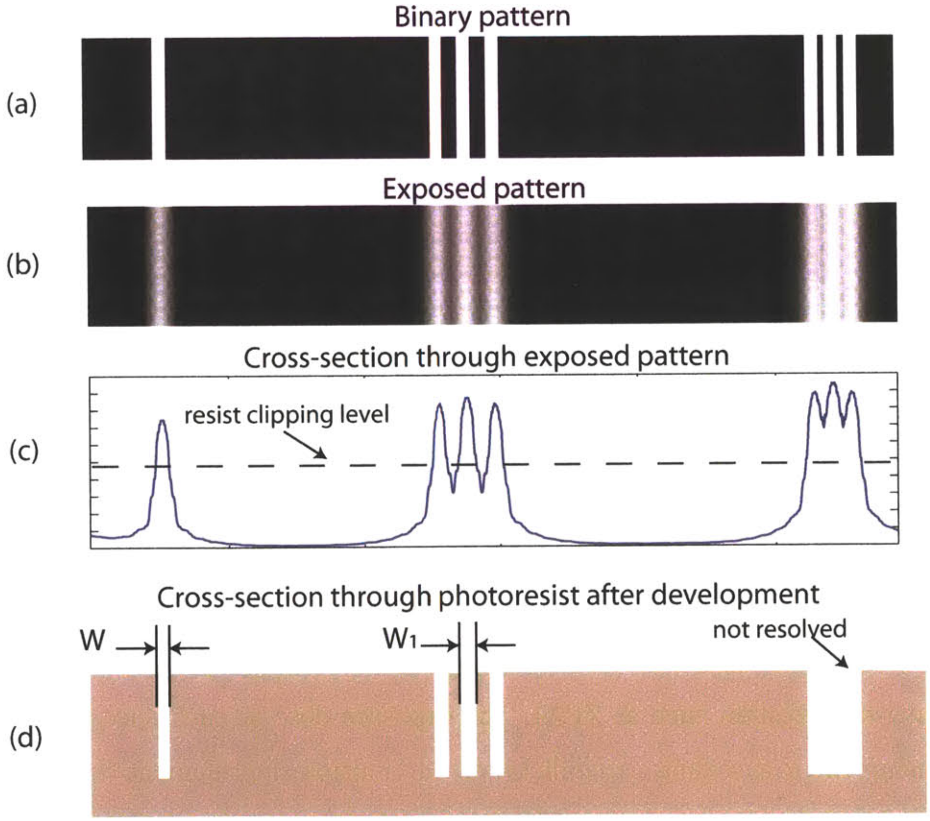


Figure 5-2: *Schematic illustrating proximity effects. (a) Desired binary pattern consists of lines (all of same width) of different densities. (b) Aerial image of the pattern after exposure in a ZPAL system. (c) Cross-section through the image. (d) Pattern in resist after development. Note that the width of the isolated line, W, is different than the width of the semi-dense lines, W1. The most dense lines were not resolved in the resist.*

A dramatic example of proximity effects is shown in figure 5-3. The pattern consists of a single line between two large squares. The scanning-electron micrograph in the figure indicates that the segment of the line between the squares is considerably wider because of the extra exposure dose from the two squares in close proximity to the line. This pattern was exposed with a 0.7 NA zone plate at $\lambda = 442\,\text{nm}$.

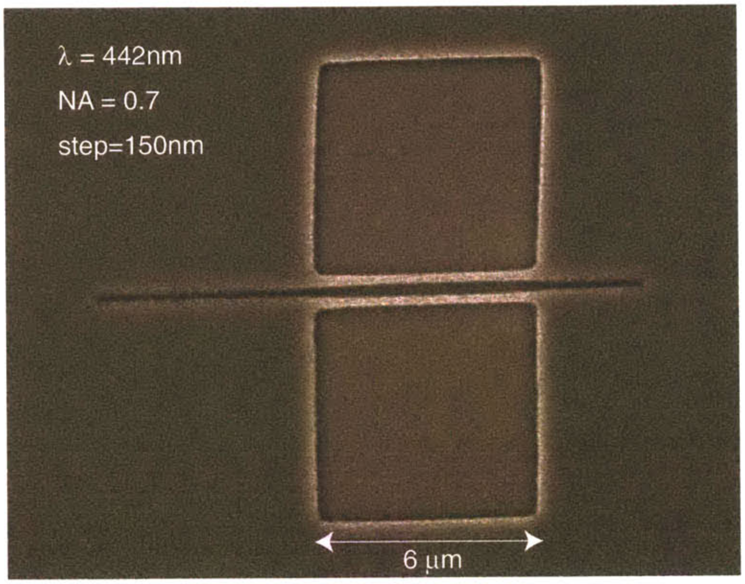


Figure 5-3: *Demonstration of proximity effects. Scanning-electron micrograph of a pattern exposed with a 0.7 NA zone plate at* $\lambda = 442\,nm$ *and a step-size of 150 nm* ($s \approx 2$). *The width of the line-segment between the squares is considerably larger than the rest of the line. This is due to additional dose from the squares in close proximity to the line.*

5.2 Proximity Effect Correction (PEC)

In a scanning-spot system such as ZPAL, the exposure dose at each spot can be controlled either by using the grayscaling capabilities of the multiplexing element or by changing the exposure time at any given position. This *dose modulation* can be used to correct for proximity effects. The technique is illustrated in figure 5-4 using a simple example. In figure 5-4(a), a single spot exposed in resist is shown. Due to the exposure of other spots nearby (not shown), background dose gets added. This causes the size of the spot to increase as shown. The solution to this problem is to underexpose the spot in the first place such that the new dose together with the background produce the spot of the desired size (see figure 5-4(b)).

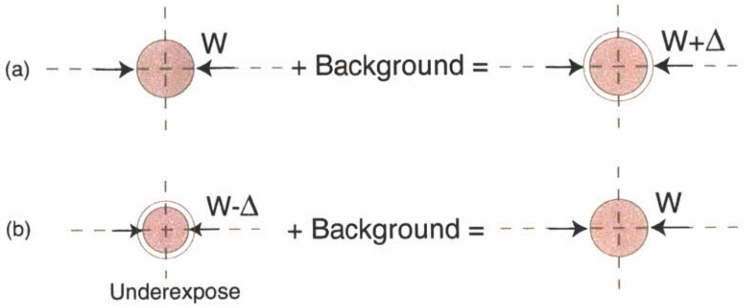


Figure 5-4: *Schematic illustrating the use of dose modulation for PEC. (a) Due to background dose from neighboring spots, the size of a single spot increases, i.e. the spot is slightly overexposed. (b) The correct spot-size can be regained by appropriately underexposing the spot such that the lower dose together with the background produce the correct (nominal) dose.*

Thus, dose modulation provides a powerful means of improving the quality of printed patterns. Along with sub-pixel stepping, dose modulation also offers a means to control the placement of edges of features. In figure 5-5, scanning-electron micrographs of isolated lines patterned using a 0.7 NA zone plate at $\lambda = 400\,\text{nm}$ is shown. The leftmost line is printed using a single colum of spots. By modulating the exposure dose of a second column of spots at a distance equal to a fraction of the spot-size (in this case, 1/3) from the first column, the width of the line can be controlled. In addition, by modulating the dose of the two columns together, the width of the line as well as the placement of both the edges of the line can be independently controlled. In the right-most image, a third column of spots is added to increase the linewidth further.

Thus, ZPAL offers two handles which can be used to optimize the written pattern. The first is the positions of the written spots (defined by the sub-pixel stepping), and the second is the dose in each spot (dose modulation). Proximity effects in ZPAL can be corrected by numerically compensating the exposed pattern prior to actual exposure. This is achieved by modifying the exposure dose in each spot appropriately, such that the final-exposed pattern exhibits no (or little) proximity effects. This requires a calculation of the exposure-doses prior to writing using the simulation tools described in the earlier chapters. Accurate knowledge

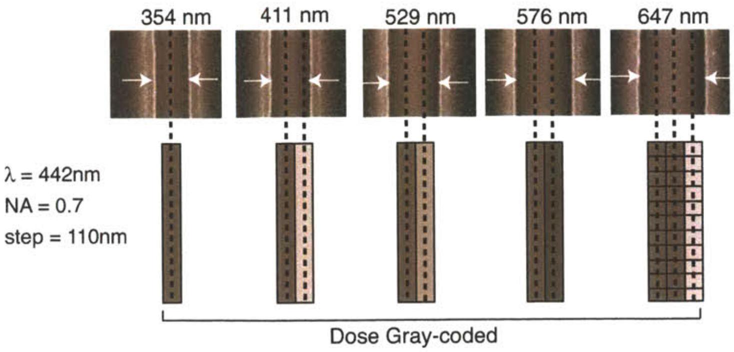


Figure 5-5: *Linewidth and edge-placement control using dose modulation and sub-pixel stepping. Scanning-electron micrographs of lines of different widths are shown. These lines were exposed with a 0.7 NA zone plate at $\lambda = 442nm$. The linewidth was increased by exposing a column of spots at a distance of 110 nm such that the two columns of spots overlapped. This sub-pixel stepping technique together with dose modulation provides sufficient control to place both edges of a line independently.*

of the point-spread function (PSF) of the zone plate is necessary to calculate the proximity effects and hence, to calculate the optimum exposure doses for correcting these effects. First, we present initial results, where the optimum doses were calculated by trial and error, i.e. those doses in regions of high density were reduced relative to those doses in regions of low density and this was carried out in an ad-hoc manner. The scanning-electron micrographs of the same pattern with and without proximity-effect correction are shown in figure 5-6. The pattern consists of dense nested-Ls and a single spot in the bottom-right corner. These patterns were written with a 0.67 NA zone plate at $\lambda = 400\,\text{nm}$. In the uncorrected pattern, all the spots had the same dose, while in the corrected pattern, the doses were modulated to take into account the local-pattern density. The isolated spot is not cleared (underexposed) in the uncorrected pattern, while it clearly shows up in the corrected version. In the uncorrected pattern, the lines within the nested region are wider than the segments which extend outside. In the corrected pattern, the doses in the line-segments outside the nested region are increased such that the lines have a more uniform width throughout the pattern.

Thus, we have shown that dose modulation is an effective means of implementing proximity-effect correction. Next, we discuss the three methods of calculating the optimum doses in a

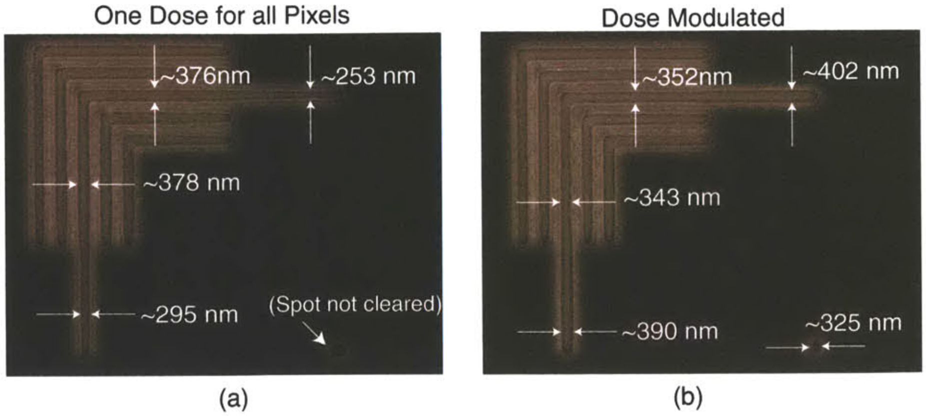


Figure 5-6: *Proximity-effect correction by dose modulation. (a) Scanning-electron micrograph of an uncorrected pattern (all spots have same dose) consisting of dense nested-L shapes and a single isolated spot. Note that the spot is underexposed and hence, not cleared in resist. (b) Scanning-electron micrograph of the corrected pattern. The dose of each spot was modulated to take into account the local density. Note that the isolated spot is cleared. The widths of the lines are more uniform as well. There is some over-correction, i.e. linewidths are wider outside the nested region. This is due to the ad-hoc calculation of the doses. The patterns were exposed with a 0.67 NA zone plate at* $\lambda = 442\,nm$.

rigorous manner. These are: (1) a direct method based on the solution of a system of linear equations, (2) a method based on linear programming, and (3) a method based on simple iterative error-correction. First, we begin by recapitulating the process of simulating the exposed patterns. In figure 5-7 (a) a binary pattern is shown in gray superimposed on a grid of points. This grid represents the discretization of space for the calculation of the exposed pattern i.e. the grid points are the sample points for the aerial image. In figure 5-7 (b), the pattern is made up of spots. In this schematic, the spot steps over a distance equal to half the spot-size for each exposure (s=2). Note that in the schematic, the grid size is one-tenth the size of the spot. This is typical especially for proximity-effect correction calculations, because the resolution of the grid sets the lower limit for the accuracy of edge-placement in the final pattern.

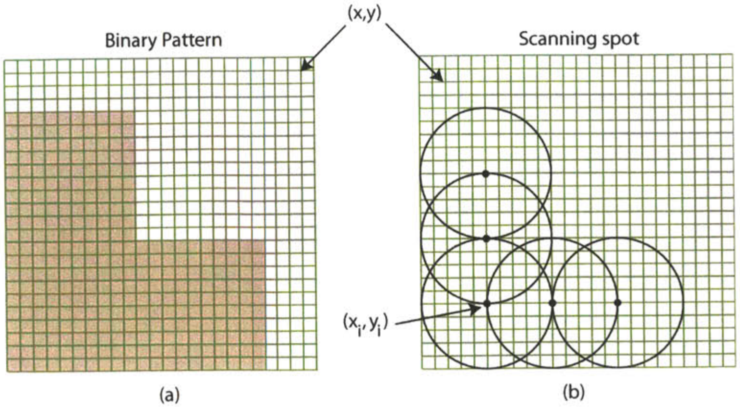


Figure 5-7: *Writing technique in ZPAL. (a) The desired pattern is depicted in gray on top of a grid. The grid represents the points in space, where the aerial-image intensity is sampled for simulation purposes. (b) The focussed spot is scanned to form the pattern. For each exposure, the spot is stepped over by a distance equal to half the spot-size, i.e.* $s = 2$. *The resolution of the grid puts a limit on the accuracy of the placement of feature-edges for PEC.*

The binary pattern $B(x, y)$ is defined as

$$B(x,y) = \begin{cases} 1 & \text{inside pattern boundary} \\ 0 & \text{otherwise} \end{cases} \tag{5.1}$$

The exposed pattern $E(x, y)$ is a discrete convolution of the binary pattern with the PSF and is given by

$$E(x,y) = \sum_{i=1}^{N_{exp}} B(x_i, y_i) \times P(x - x_i, y - y_i) \tag{5.2}$$

where (x_i, y_i) are the coordinates of the poistions of the exposed spots, N_{exp} is the total number of exposed spots, and $P(x, y)$ is the two-dimensional PSF of the zone plate. There are 2 degrees-of-freedom in the pattern-printing process. They are (x_i, y_i), the positions of the exposed spots, and the dose at each exposed spot. As described in the previous chapter, the positions of the exposed spots are chosen so as to achieve smooth features. If the grid on which the exposed spots lie is fixed, then dose modulation is the only means of proximity-effect correction (PEC). If we express the dose-modulated pattern as $D(x, y)$, then the exposed pattern is given by

$$E(x, y) = \sum_{i=1}^{N_{exp}} D(x_i, y_i) \times P(x - x_i, y - y_i) \tag{5.3}$$

PEC involves the calculation of $D(x_i, y_i)$ for a given binary pattern and PSF. We describe three methods of calculating the doses below.

Direct Method

Equation 5.3 is a linear system of equations. It can be rewritten in matrix form as

$$\begin{bmatrix} E(x,y) \\ \vdots \\ \vdots \end{bmatrix} = \begin{bmatrix} \vdots & \vdots & \vdots \\ \vdots & P(x - x_i, y - y_i) & \vdots \\ \vdots & \vdots & \vdots \end{bmatrix} \times \begin{bmatrix} D(x_i, y_i) \\ \vdots \\ \vdots \end{bmatrix} \tag{5.4}$$

If we set the exposed pattern, $E(x, y)$ equal to the binary pattern, $B(x, y)$, and invert this system of equations, the unknown, $D(x_i, y_i)$ can be calculated.

$$\begin{bmatrix} D(x_i,y_i) \\ \vdots \\ \vdots \end{bmatrix} = \begin{bmatrix} \vdots & \vdots & \vdots \\ \vdots & P(x - x_i, y - y_i) & \vdots \\ \vdots & \vdots & \vdots \end{bmatrix}^{-1} \times \begin{bmatrix} B(x, y) \\ \vdots \\ \vdots \end{bmatrix} \tag{5.5}$$

It is necessary to control the edge-placement of features to a small fraction of the feature size. This means that the number of points (x, y) is much larger than the number of exposure spots, (x_i, y_i) and the inverse of the matrix of $P(x-x_i, y-y_i)$ in equation (5.4) may not exist. In this case, the system of equations in (5.4) is overconstrained, i.e. the number of equations is more than the number of unknown variables. This implies that a feasible solution may not exist. Another important disadvantage of this technique is that the computational burden is very high especially when the sizes of the matrices are large. Even when the solutions exist, there is no constraint to ensure that the computed doses are positive (negative exposure is nonphysical). To account for this effect a blanket exposure dose can be applied to all the spots. However, this decreases the image contrast. In practise, this method turns out to be a poor solution, especially for complex and dense patterns.

Linear Programming Method

This method was formulated based upon the the highly non-linear nature of photoresist. In the binary clipping model, the positive photoresist is removed after development only in those regions that receive energy more than an upper threshold. Using this model, the requirement for aerial image, $E(x,y)$ to form the desired binary pattern is

$$E(x,y)\begin{cases} > UT & \text{inside pattern boundary} \\ < LT & \text{otherwise} \end{cases} \tag{5.6}$$

where LT is the lower threshold, and UT is the upper threshold. The mathematical problem now becomes finding a set of values for $D(x_i,y_i)$ that satisfies the equations (5.3) and (5.6). Using equation (5.3), we can write

$$\sum_{i=1}^{N_{exp}} D(x_i,y_i)\times P(x-x_i,y-y_i)\begin{cases} > UT & (x,y) \text{ inside pattern boundary} \\ < LT & \text{otherwise} \end{cases} \tag{5.7}$$

To simplify the notation, we define two more functions,

$$P'(x-x_i,y-y_i)=\begin{cases} +P(x-x_i,y-y_i) & (x,y) \text{ inside pattern boundary} \\ -P(x-x_i,y-y_i) & \text{otherwise} \end{cases} \tag{5.8}$$

and

$$CL(x,y)=\begin{cases} +UT & (x,y) \text{ inside pattern boundary} \\ -LT & \text{otherwise} \end{cases} \tag{5.9}$$

Equation (5.7) can then be written as

$$\sum_{i=1}^{N_{exp}} D(x_i, y_i) \times P'(x - x_i, y - y_i) > CL(x, y) \tag{5.10}$$

We define a *cost function*, C as

$$C = \sum_{i=1}^{N_{exp}} |\epsilon(x, y)| \tag{5.11}$$

where $\epsilon(x, y)$ is the difference between the exposed pattern and the binary pattern.

$$\epsilon(x, y) = E(x, y) - B(x, y) \tag{5.12}$$

Now the goal of the problem has been changed to finding the minimum of the function C, given the constraint in equation (5.10). This is the classic definition of a linear programming (LP) problem [39]. The recasting of the problem into an LP form actually is better than a simple search problem. This is because minimizing the cost function, C in this case, abets the search for the optimum exposure doses.

A test pattern consisting of a single unexposed spot in the center of an exposed square region was used to test the LP algorithm. The results for one case is shown in figure 5-8. The simulations used the PSF of a 0.7 NA zone plate at $\lambda = 400\,\text{nm}$ and $s \approx 1$. The size of the unexposed spot was 200 nm. The binary pattern is shown in figure 5-8 (a). The uncorrected version of the aerial image is shown in figure 5-8 (b). Note that the contrast of the unexposed region is very low. In an actual experiment, this would mean that the spot is not patterned in the resist. This is because of the proximity dose contributions from the regions near the central spot. In figure 5-8 (c), the aerial image of the corrected pattern is shown. The contrast of the unexposed pixel in the center is improved dramatically. However, there is quite a bit of roughness along the edges of the pattern. This is probably due to the fact that step-size=200 nm ($s \approx 1$). Nevertheless, the LP method is a step in the right direction for solving the proximity effect problem.

This method showed moderate success. However, the amount of memory required to

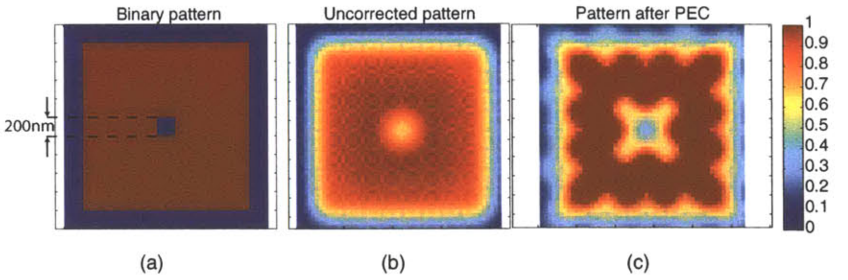


Figure 5-8: *Simulation results of PEC with the LP method. (a) The test pattern consisted of a single unexposed spot in the center of an exposed square. The size of the spot was about 200 nm. (b) Aerial image of the simulated pattern in which the doses of all the spots within the pattern boundary was the same. (c) Aerial image of the simulated pattern where the doses of the spots were modulated. The doses were calculatd based on an LP method. The simulations were based on the PSF of a 0.7 NA zone plate at* $\lambda = 400\,nm$.

solve the LP problem increases dramatically for large patterns. Although this method has mathematical rigor and offers some interesting analysis of the proximity-effect correction problem, its practical application is currently limited to small patterns.

Iterative Error-Correction Method

This method involves modifying the exposed pattern in a step-by-step manner by modulating only those exposure doses that are responsible for causing the deviation from the desired pattern. This modulation is performed in an iterative manner. The steps involved are detailed below:

1. Starting from the binary pattern (all doses inside features equal to 1), the exposed pattern is simulated as described earlier.

2. The exposed pattern is passed through a thresholding function, which determines the reaction of the photoresist on the aerial image. Those regions of the image which have intensity higher than a given threshold is set to 1, and all other regions are set to 0.

3. The difference between the thresholded pattern, E_{th} and the original binary pattern,

BP is denoted ϵ.

$$\epsilon(x, y) = E_{th}(x, y) - BP(x, y) \tag{5.13}$$

4. For each point (x, y), which has non-zero error, a search is performed to determine its nearest-exposed spot, (x_i, y_i)[2]. Depending on the sign of the error at (x, y), the dose at (x_i, y_i) is modified by a dose quantum. The dose quantum is determined by the resolution of dose-control, avaliable in the system.

$$D(x_i, y_i) \Rightarrow D(x_i, y_i) - \text{sign}(\epsilon(x, y)) \times DQ \tag{5.14}$$

where DQ is the dose quantum and sign is the sign-function.

5. Using the dose-modulated pattern, $D(x_i, y_i)$, the exposed pattern is again calculated. Steps 2-4 are repeated until an acceptable level of the error is achieved.

A schematic of this algorithm is shown in figure 5-9. The pattern shown consists of a single unexposed spot in the center of an exposed square area. The size of the spot was intended to be 200 nm. The starting point assumes the same dose (=1) for all spots within the pattern and a dose of zero everywhere else. Note that with the initial dose distribution, the central spot is not resolved as shown by the thresholded pattern. A resist threshold at 0.5 of the peak image intensity was assumed. The simulations shown in the figure were performed using the PSF of a 0.7 NA zone plate at $\lambda = 400\,\text{nm}$. The algorithm ran for about 100 iterations[3] and the final corrected pattern is shown in figure 5-10. The aerial image is shown in figure 5-10(a) and the image after thresholding (assuming a threshold of 0.5) is shown in figure5-10 (b). The edges of the square pattern are much sharper and the single spot in the center is resolved in resist. From this simple simulation, it is evident that this iterative algorithm gives superior results compared to the previous methods. There

---

[2] For sake of clarity, *point* refers to the discretization of space, whereas *spot* refers to a point on the exposure (stage-stepping) grid.

[3] The algorithm converged to a final pattern within about 10 iterations. The extra iterations were performed to ensure that the global optimum solution was found.

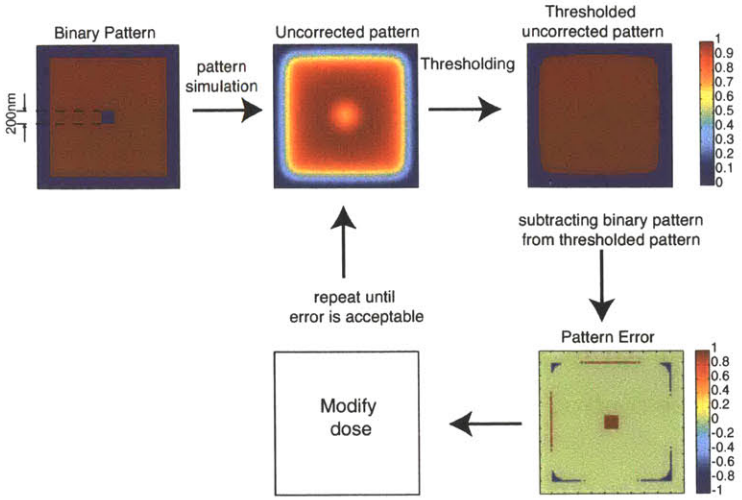


Figure 5-9: *Schematic illustrating the steps involved in the iterative-error-correction method. The binary pattern is the starting point for simulating the exposed pattern. The exposed pattern is thresholded to simulate the effect of the resist. The difference between the resulting pattern and the initial-binary pattern is the pattern error. This is used as the basis for modulating the exposure dose.*

are still some rough edges to the pattern. This is of the order of the size of the space-grid (about 20 nm). In order to investigate the effect of the size of the space-grid, the iterative-error-correction algorithm was used to optimize the same pattern with different grid-sizes. As expected, as the size of the grid decreases, the number of points in the image increases and the computation is slower. Nevertheless, we would expect the calculation to be more accurate since decreasing the grid-size means that the aerial image is sampled at a higher frequency. Subsequently, the corrected image improves slightly. This is shown in figure 5-11.

The previous simulations assumed a s-factor of 2, i.e. adjacent spots were exposed at a spacing equal to 100 nm. Next, we investigate the effect on the optimized pattern of changing the s-factor. The test pattern was simulated and optimized at different s-factors, keeping all other parameters the same. This is shown in figure 5-12. As s increases from 1 to 2, the optimized pattern improves significantly. However, further increase in s from 2 to 3, shows no improvement at all even from the original uncorrected pattern. This was the case even after running the algorithm for a large number of iterations. This is probably

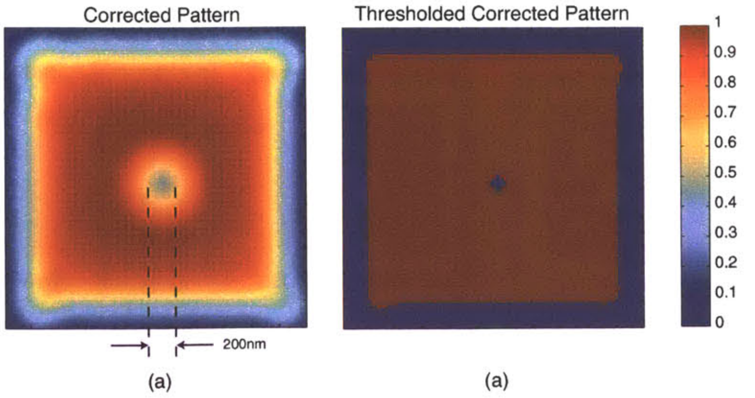


Figure 5-10: *Simulated pattern corrected using the iterative-error-correction method. (a) Aerial image of the corrected pattern. (b) Intensity after passing through a thresholding function (threshold = 0.5). The corrected pattern is better defined compared to the uncorrected pattern in figure 5-9.*

because increasing s means decreasing the spacing between the exposed spots. As the overlap between neighboring spots increases, not only the nearest exposed spot should be considered for error correction but also the next nearest one as well. The simulation was performed using the PSF of a 0.7 NA zone plate at $\lambda = 400\,\text{nm}$, and a space-grid of 20 nm.

The previous simulations used 5 bits of grayscaling (32 dose levels). One of the advantages of the iterative error-correction algorithm is that the dose quantization is automatically taken into consideration during the calculation. In order to study the effect of this quantization on the quality of the corrected patterns, we simulated the same pattern with different levels of dose quantization. The results are shown in figure 5-13. As the number of bits increases, the control over the dose increases and hence, the quality of the corrected patterns increases as well. However, beyond about 5 bits, the improvement is not very significant.

In the iterative method, the number of iterations required to optimize a given pattern is an important computational metric. The pattern consisting of a single unexposed spot in the center of a square is again used for the study. The pattern was simulated and the mean of the error in the pattern is plotted as a function of the iteration number. The mean error

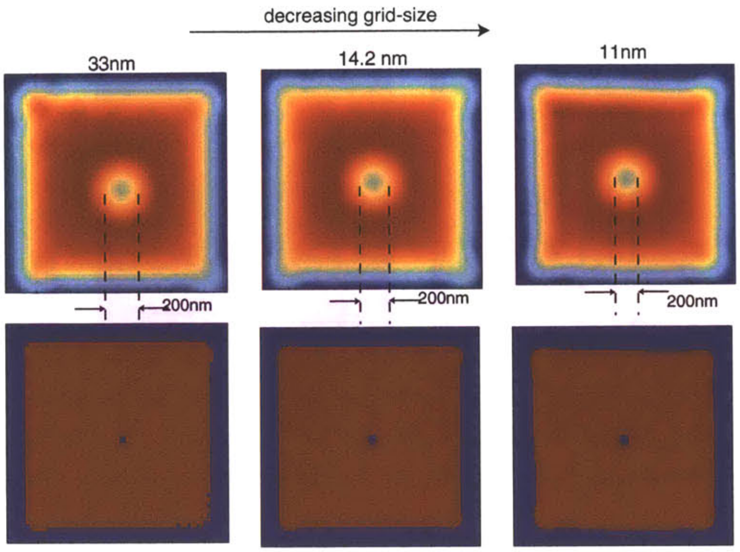


Figure 5-11: *Effect of space-grid size on PEC. The size of the space grid is decreased from left to right. Top row shows the aerial intensity pattern Bottom row shows the corresponding pattern after resist thresholding (assuming a threshold of 0.5). As the grid decreases, the pattern after thresholding improves slightly. The simulations were performed using a 0.7 NA zone plate at* $\lambda = 400\,nm$ *and* $s = 2$ *(step-size=100 nm).*

is defined as

$$< \epsilon >= \frac{\sum_{x,y} \epsilon(x,y)}{N_x N_y} \tag{5.15}$$

where N_x and N_y are the number of sample points of the aerial image in the X and Y directions respectively. The results are shown in figure 5-14. The error converges to a stable value in very few iterations (15). This is seen to be true for almost all feasible patterns. The fact that the error stabilizes at a non-zero value even after a large number of iterations attests to the fact that there is no exact solution to the system of equations. The solution is not feasible. The fast convergence implies that the iterative error-correction algorithm is very efficient in terms of finding the minimum error.

The iterative error-correction algorithm was implemented using Matlab. The pattern conversion software was modified to accomodate a different dose in each pixel. This conversion software was also written in Matlab.

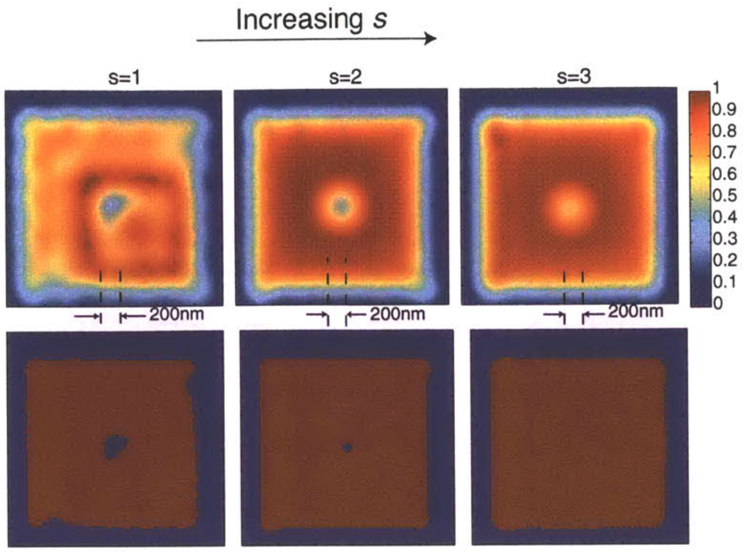


Figure 5-12: *Effect of spacing of exposed spots (s-factor) on PEC. The spacing between the exposed spots decreases from left to right (s-factor increases). The corrected-aerial image is shown in the top row. The corresponding thresholded image is shown in the bottom row. As s increases from 1 to 2, the corrected pattern improves significantly. However, there is no improvement even from the original uncorrected pattern for the case of* $s = 3$. *The simulations were performed using the PSF of a 0.7 NA zone plate at* $\lambda = 400\,nm$ *and a space-grid of* $20\,nm$.

Finally, we report the experiments performed to test the iterative error-correction algorithm. First a pattern was exposed with all the spots having the same dose. Then the iterative error-correction method was used to compensate the pattern for proximity effects by dose modulation. The dose modulation was achieved in practise by changing the exposure time of each spot. Scanning-electron micrographs of the corrected (right) as well as the original pattern (left) are depicted in figure 5-15. The single unexposed spot in the center of the exposed square (on the bottom right) is only resolved after PEC. This is an example of a case where PEC could improve the resolution of an isolated feature. The edges of the square are much sharper in the corrected version. In the nested-L pattern (on the bottom left), the smallest nested-L is much better defined with sharp edges in the corrected pattern. The rounding at the edges of the lines (on the top left) is corrected by PEC to form square edges. On the top right, one array of small circles and an array of squares were printed. In the uncorrected pattern, the edges of the squares are very rounded such that the large squares

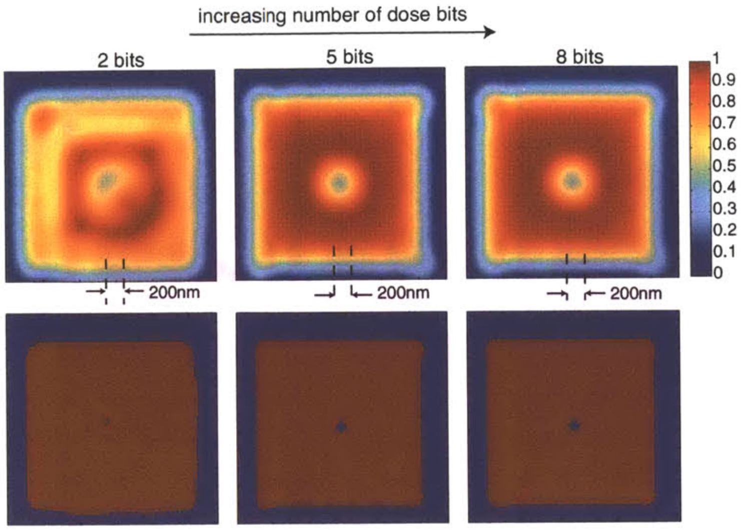


Figure 5-13: *Effect of quantization of dose on PEC. The same pattern was simulated using the PSF of a 0.7 NA zone plate at* $\lambda = 400\,nm$ *and* $s = 2$. *As the number of dose-levels increases, the increased control over the dose enables better corrected patterns. However, the improvement is negligible beyond 5bits (32 levels of dose).*

look like circles. The small circles are barely cleared in resist. In the corrected pattern, on the other hand, the edges of the squares are well-defined and the square shapes are clearly visible. The small circles are well-cleared in resist as well. Magnified scanning-electron micrographs of these features are shown in figure 5-16. These patterns were exposed using a 0.7 NA zone plate at $\lambda = 400\,\text{nm}$ and step=150 nm ($s = 2$).

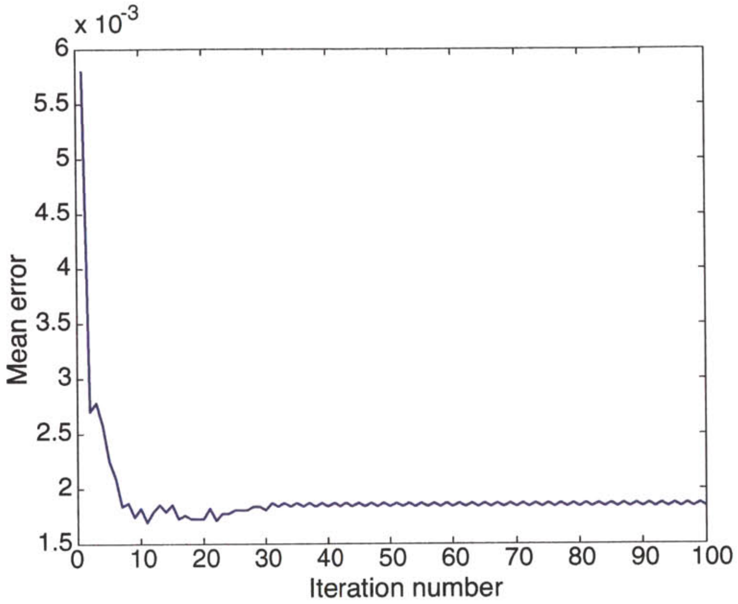
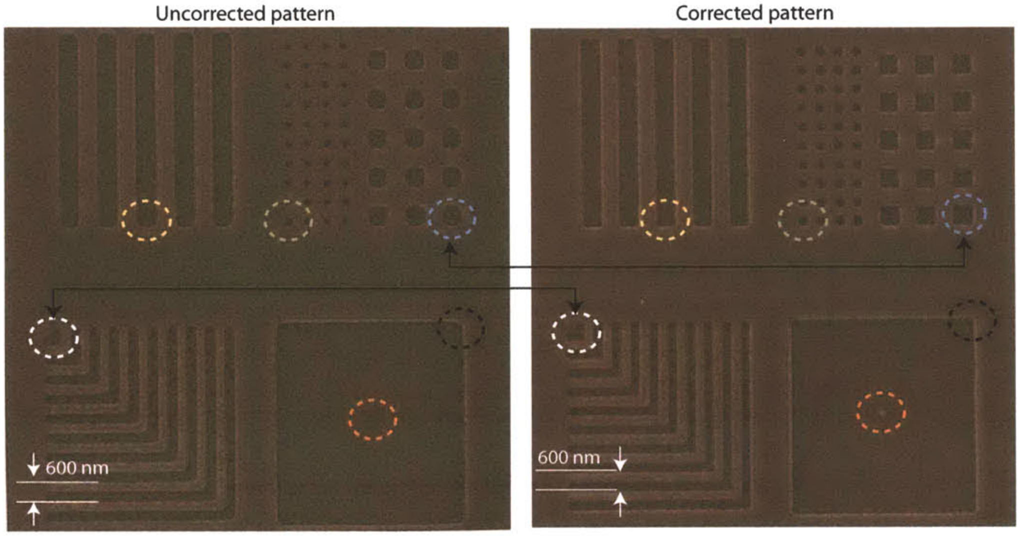


Figure 5-14: *Mean error as a function of the number of iterations. The mean error (sum of the error in the entire pattern divided by the number of points in the pattern) is plotted versus the iteration number. The data is from the optimization of a pattern consisting of a single unexposed spot in the center of an exposed square. The simulations were performed using the PSF of a 0.7 NA zone plate at* $\lambda = 400\,nm$ *and* $s = 2$ *(step-size=100 nm).*


Uncorrected pattern

Corrected pattern

(λ = 400nm, NA = 0.7, Focal Length = 40 µm)

Figure 5-15: *Experimental results of PEC using the iterative algorithm. Left: Original (uncorrected) pattern. All the spots have the same dose. Right: Corrected pattern. The doses of individual spots were pre-calculated to compensate for proximity effects using the iterative error-correction algorithm. These experiments were performed using a 0.7 NA zone plate at* $\lambda = 400\,nm$ *and step=150 nm.*

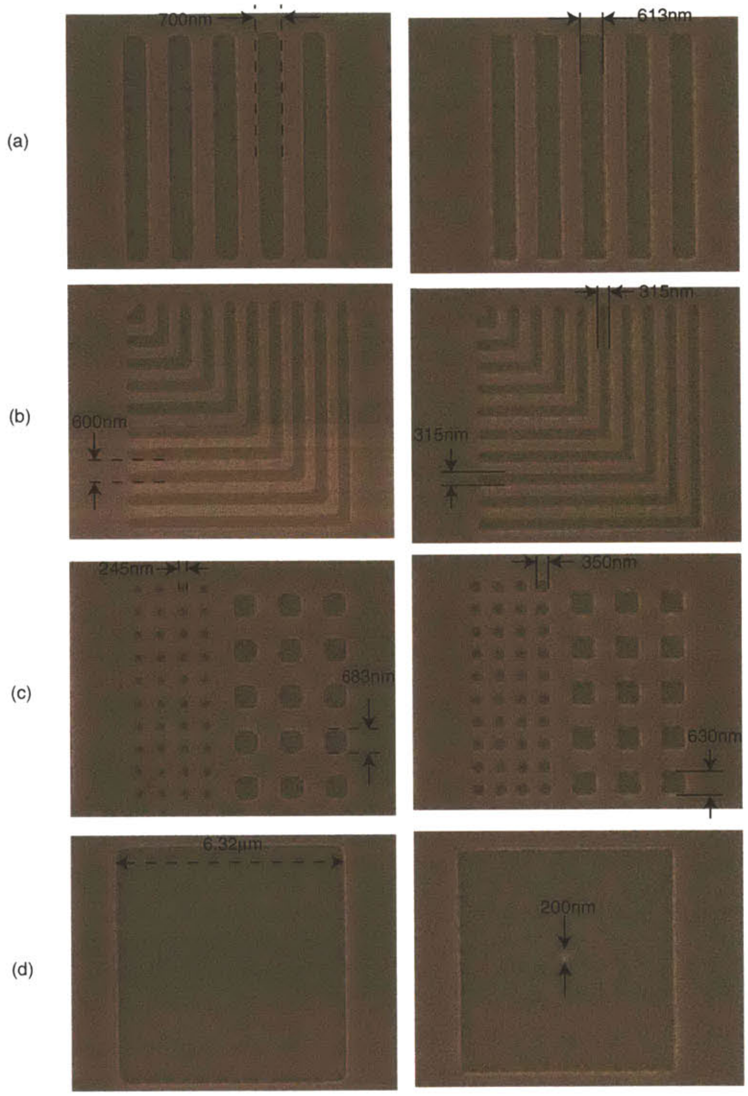


Figure 5-16: *Magnified images of the experimental results of iterative PEC.*

5.3 How does PEC in ZPAL compare to other lithography techniques ?

In this section, we want to understand how the proximity effects and their methods of correction in ZPAL differ from those in other lithography methods, namely electron-beam lithography and optical-projection lithography.

5.3.1 Electron-Beam Lithography

In electron-beam lithography, a focussed beam of electrons is scanned over a substrate to define the pattern. In this case, proximity effects arise from the forward scattering of the electrons in photoresist as well the backscattering of electrons from the substrate. Due to the complex behavior of electrons in the resist, the PSF in an electron-beam lithography system must take into account the effect of scattering from the resist and the substrate. Therefore, this is a complex numerical problem usually solved by Monte-Carlo methods. An approximate model can be built using empirically determined parameters.

$$PSF_{ebeam}(r) = ke^{-\frac{r^2}{\beta_f^2}} + \eta_E e^{-\frac{r^2}{\beta_b^2}} \tag{5.16}$$

where r is the radial distance from the center, β_f is a factor defined primarily by the forward scattering of electrons, η_E is the ratio of the total integrated dose due to backscattered electrons to that due to the incident electrons, and β_b is a factor determined primarily by the backscattered electrons [40]. Using typical values, $\beta_f = 0.1\mu\text{m}$, $\beta_b = 3\mu\text{m}$, and η_E=0.5, we can plot the normalized "intensity"[4] in the focussed electron-beam as a function of the radial distance. This plot is shown in figure 5-17. The PSF of a 0.7 NA zone plate at $\lambda = 400\,\text{nm}$ is shown for comparison. The size of the incident electron-beam spot is smaller. Photons undergo insignificant scattering compared to electrons. From the plots, it seems that even though the size of the spot is smaller for electrons, the background seems to be about the same as for the zone plate.

---

[4] Intensity for electron-beam lithography is the energy deposited per unit volume.

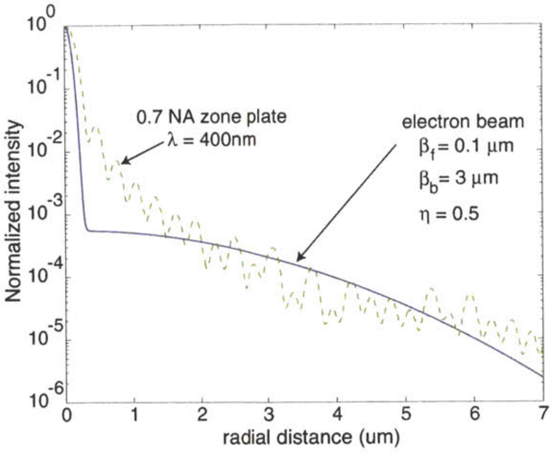


Figure 5-17: *Focussed spot in an electron-beam lithography system. The electron-beam focussed spot is plotted with the PSF of a zone plate of NA=0.7 and* $\lambda = 400\,nm$. *Photons undergo very little scattering compared to electrons. Therefore, even though the size of the electron spot is smaller, the background seems to be about the same.*

Just as in ZPAL, PEC in electron-beam lithography can be achieved by dose modulation. Methods similar to the ones discussed in this chapter may be used to calculate the modulated doses. Several different methods for computing the modulated doses have been proposed in the past. These range from using neural networks [41] to rule-based [42], and simplex methods [43]. Another technique that is often used in electron-beam lithography is *biasing*. In this approach, the dense features are modified by reducing their linewidths slightly. This method is computationally similar to dose modulation. However, pattern biasing is effective only when the feature sizes are large compared to the spot-size. An alternative method is termed *GHOST*. This method involves exposing the negative tone of the pattern using a defocussed electron beam, which simulates the distribution of the backscattered electrons. Once the actual pattern is written, the resist has an overall uniform background exposure which averages out proximity effects. This method requires no computation. However, there is considerably higher writing time. [PUT REFERENCE TO PECKERAR].

5.3.2 Optical-Projection Lithography

In optical-projection lithography, the pattern on a photomask is imaged onto the substrate using a complicated optical system. The optical system is complicated due to the need to maintain high resolution of the printed features over a large area (equal to the field-size), the need to maintain uniform illumination and minimal distortion, and the need to correct for aberrations which exist in refractive lenses. Proximity effects, in this cases, arise from a complex interaction of light as it diffracts through dense features on a photomask. As is the case in ZPAL, proximity effects appear as a broadening of linewidths of dense gratings, for example. The fundamental difference between optical-projection lithography and ZPAL is that image formation in ZPAL is an incoherent process and hence, calculation of the exposed image is relatively easy[5]. In the case of optical projection, the image-formation process is partially coherent, which complicates the simulation of the printed pattern, and hence, the calculation of the influence of proximity effects. Therefore, proximity-effect correction in optical-projection lithography (commonly known as optical-proximity correction or OPC) is much more complicated. In summary, OPC depends on not only on the pattern density, and the imaging system (as was the case in ZPAL) but also on illumination paramters such as partial coherence and the use of off-axis illumination.

Proximity-effect correction in optical-projection lithography is primarily carried out by modification of the features on the photomask such that the printed pattern on the substrate is closer to the desired pattern. For example, in the case of dense gratings, the linewidths on the mask may be biased (made smaller) to account for proximity effects. In practise, any modifications to the photomask is described under the umbrella term, OPC. For example, the *corner-rounding* which make squares print as circles or *line-edge shortening*, where the length of the printed pattern is smaller than the length of the corresponding feature on the mask (after taking into account the optical demagnification). These examples are illustrated in figure 5-18.

In conclusion, proximity-effect correction in ZPAL is a simpler process in terms of the

---

[5] As shown earlier, for incoherent imaging like ZPAL, the image is the convolution of the PSF with the binary pattern.

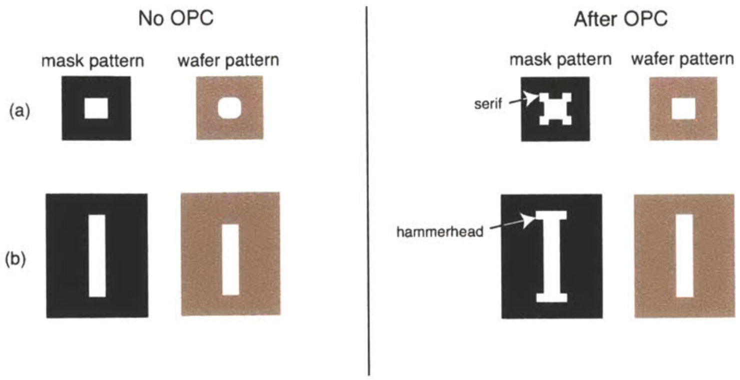


Figure 5-18: *Examples of optical-proximity correction. (a) Corner-rounding causes squares to print as circles. OPC adds sub-resolution serifs to the corners such that the sharp edges are printed as shown on the right. (b) Line-edge shortening causes the length of lines on the substrate to be shorter than that on the mask (after taking into account the demagnification of the optics). OPC adds a hammerhead to the ends of the line. These are examples of intra-feature proximity effects.*

calculation compared to optical-projection lithography. In comparison to electron-beam lithography, the complexity of the calculations are of the same order. However, the effectiveness of the simple iterative error-correction algorithm seems to suggest that in practise, PEC might be much more effective in ZPAL.

Chapter 6

Zone-Plate-Array Microscopy

In this chapter, we describe an imaging technology which is a natural extension from the lithography technology. Experimental results verifying this imaging tool are also presented. Finally, promising applications for this technology are described.

6.1 Introduction to Zone-Plate-Array Microscopy

In ZPAL, an array of zone plates focus incident light into an array of spots on the substrate. If the light scattered from the substrate is collected and redirected through a lens and a pinhole into a detector, we can extract information about the substrate. An example will help illustrate the concept. Imagine a plane mirror in the focal plane of a single zone plate as shown in figure 6-1. When the mirror is in the focal plane of the zone plate, the reflected light that is collected by the zone plate forms a collimated beam which is then directed through a pinhole, using a lens, onto a detector. When the mirror is displaced from the focal plane, the reflect light that is collected by the zone plate is not collimated by the zone plate. This light is now diffused in the plane of the pinhole and hence, a very weak signal reaches the detector. The signal on the detector can be used as a measure of the displacement of the mirror from the focal plane. This is the basic idea of confocal microscopy [44]. By scanning the focussed spot in the transverse plane, topographic information of a general substrate can be obtained. By performing these scans at different transverse planes, topographic sections

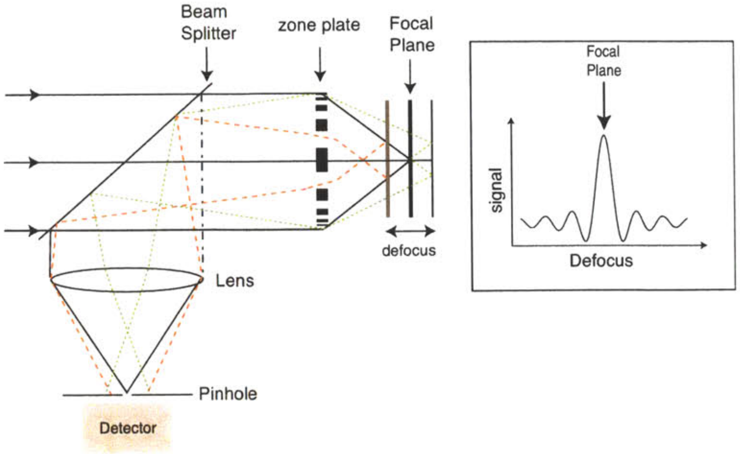


Figure 6-1: *Schematic of confocal microscopy. Light focussed by the zone plate is reflected from the substrate (in our example, a mirror). When the mirror is in the focal plane of the zone plate, the reflected light is collimated by the zone plate and then focussed into a pinhole in front of a detector. When the mirror is displaced from the focal plane, the reflected light is not collimated and blurred in the plane of the pinhole and the signal on the detector is decreased. A schematic of the signal as a function of defocus is shown on the right.*

through the sample can also be obtained. Thus, the confocal setup offers depth selectivity.

In ZPAL, there is an array of focussed spots scanning different regions of the sample at the same time. The signals from these spots can be separated in the image plane of the zone-plate array (ZPA), and detected on a CCD camera. We call this parallel-confocal imaging technology, Zone-Plate-Array Microscopy (ZPAM). The main difference at the system level between ZPAL and ZPAM is the absence of the multiplexing device in ZPAM. A schematic of ZPAM is shown in figure 6-2. The advantages of ZPAM over traditional confocal microscopy are several. Since each zone plate images a small portion of the sample. A small scan area (shorter scan time) yields the image from a large area of the sample by tiling together the sub-images taken by each zone plate in the array. This massive parallelism makes the imaging system much faster [45]. In addition, planar fabrication techniques provide the means to make large-element arrays of high-numerical-aperture zone plates. The main disadvantage of this method is that the light collection efficieny is low because the light has to pass through the zone plate twice.

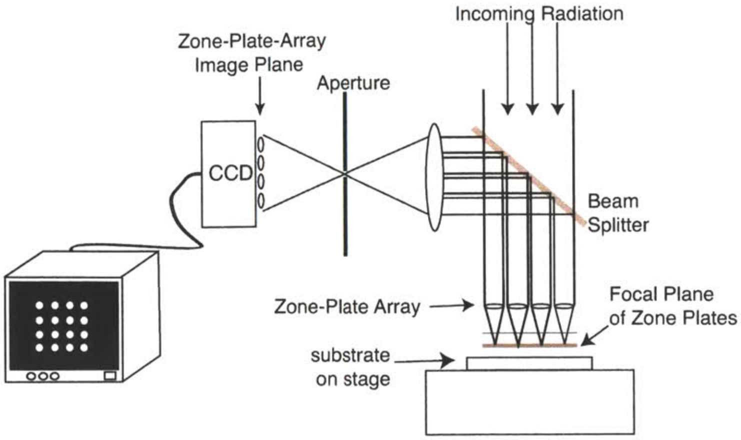


Figure 6-2: *Schematic of zone-plate-array microscopy. The confocal signals from the zone plates are separated by collecting them in the image plane of the zone-plate array. The detection device can be a CCD camera.*

6.2 Theory

In this section, we develop simple models to understand the resolution of a ZPAM system. In an ideal confocal microscope, the sample is illuminated by a diffraction-limited spot and the signal is detected through an infinitesimal pinhole (or a point detector). A schematic of an ideal scanning-confocal microscope (SCM) is shown in figure 6-3(a). Figure 6-3(b) shows the unfolded schematic of the same. The field after passing through the object is given by

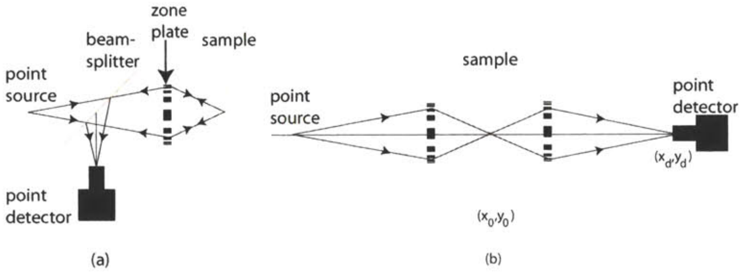


Figure 6-3: *The zone plate focuses light onto a sample, which reflects the light back onto the zone plate. The zone plate focuses the reflected light onto a point detector as the sample is scanned. The detected signal is combined with the scan coordinates to reconstruct an image of the sample. The schematic is optically unfolded to analyze the rays in (b).*

$$U(x_0, y_0; x_s, y_s) = h(x_0, y_0)t(x_0 - x_s, y_0 - y_s), \tag{6.1}$$

where (x_0, y_0) are the coordinates in the object plane, (x_s, y_s) are the scan coordinates, h is the amplitude point-spread function of the zone plate, and t is the object.

The field in the detector plane is given by

$$U(x_d, y_d; x_s, y_s) = \iint_{-\infty}^{\infty} U(x_0, y_0; x_s, y_s) h(x_d - x_0, y_d - y_0) dx_0 dy_0. \tag{6.2}$$

This can be re-written as

$$U(x_d, y_d; x_s, y_s) = U(x_2, y_2) \otimes h(x_d, y_d), \tag{6.3}$$

where $\otimes$ denotes convolution.

Assuming a point detector, $x_d = 0, y_d = 0$ and noting that h is an even function, we get

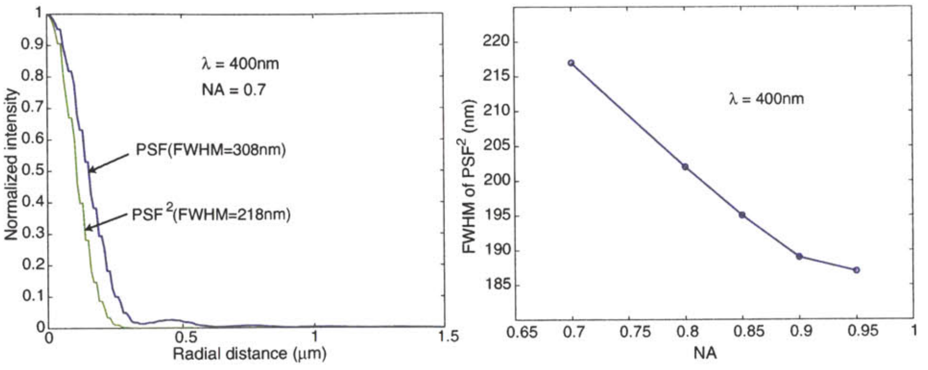
$$U(x_s, y_s) = h^2 \otimes t. \tag{6.4}$$

Thus, the point-spread function (PSF) of the confocal-imaging system is the square of the PSF of the lens [46] (the zone plate in our case).

The case of a finite pinhole is a little more involved and has been studied earlier [47, 48, 49]. The main effect of the finite pinhole (or finite detector area) is to increase the size of the PSF. The lateral resolution increases moderately whereas the depth-resolution increases significantly. In the subsequent analysis, we assume the ideal case.

Lateral Resolution

We define the lateral resolution of the confocal imaging system as the full-width at half-maximum (FWHM) of its PSF. The PSF of a confocal system with a zone plate of NA=0.7 and $\lambda = 400\,\text{nm}$ is shown in figure 6-4 (a). The PSF of the corresponding zone plate is also shown for comparison. The FWHM of the confocal system is smaller than that of the PSF of the zone plate. The sidelobes are suppressed as well. In figure 6-4 (b), the FWHM of the PSF of the confocal system is plotted as a function of the NA of the zone plate. As expected, the resolution decreases with increasing NA. The theoretical PSF indicates that the lateral resolution can be as small as 190 nm for a zone plate of NA=0.95 operating at $\lambda = 400$ nm.



(a)

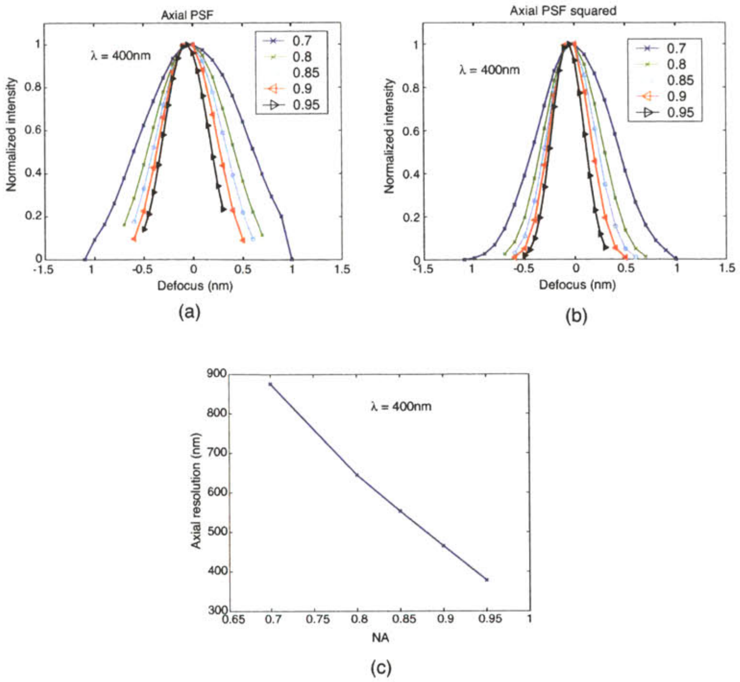
(b)

Figure 6-4: *Lateral resolution of a ZPAM system. (a) PSF of the confocal system and the PSF of the corresponding zone plate are shown. The simulations were performed for a zone plate of NA=0.7 and* $\lambda = 400$ *nm. Since the PSF of the zone plate is squared in the case of the confocal system, not only is the spot-size smaller but the sidelobes are suppressed as well. (b) Spot-size measured as the FWHM of the square of the PSF of the corresponding zone plate is plotted as a function of the NA of the zone plate.*

Axial Resolution

The axial (or depth) resolution of the confocal-imaging system can be defined as the full-width at half-maximum in the axial direction of the PSF of the system. It is interesting to note the requirements of high resolution in the axial direction is at odds with the requirement of large depth-of-focus in lithography. The axial PSF is defined as the plot of the intensity on the optical axis a function of displacement from the focal plane. The axial point-spread functions of zone plates of various numerical apertures is shown in figure 6-5 (a). The depth-of-focus of the system is approximately equal to the defocus that produces a 20% decrease from the peak value at focus. The squares of the axial PSFs are shown in figure 6-5 (b). These are the axial PSFs of the corresponding confocal system. The axial resolution is then the FHWM of the plots in figure 6-5 (b). The axial resolution is plotted as a function of NA in figure 6-5 (c). AS expected the resolution is smaller for higher NA. These simulations were performed at $\lambda = 400$ nm.

Thus, confocal microscopy offers very high resolving power in both the lateral and axial



(a) (b) (c)

Figure 6-5: *Axial resolution of a ZPAM system. (a) PSF of the confocal system and the PSF of the corresponding zone plate are shown. The simulations were performed for a zone plate of NA=0.7 and* $\lambda = 400\,nm$. *Since the PSF of the zone plate is squared in the case of the confocal system, not only is the spot-size smaller but the sidelobes are suppressed as well. (b) Spot-size measured as the FWHM of the square of the PSF of the corresponding zone plate is plotted as a function of the NA of the zone plate.*

directions.

The ZPAL system operating at $\lambda = 400\,\mathrm{nm}$ was modified into a ZPAM system. Experiments were performed on the system mainly to characterize the performance of the zone plates. The experimental results are described next.

6.3 Knife-Edge Scans

By scanning the focused spot along a sharp edge and collecting the reflected signal in a confocal manner, the "knife-edge scan" data can be obtained. It will be shown in this section that further processing of this data can yield PSF of the zone plate. We use this technique to analyze the focussing performance of the zone plates.

The knife-edge sample consisted of a cleaved silicon wafer, which was bonded on top of another silicon wafer. The support silicon wafer was necessary because of the vacuum-pin chuck used to hold the substrates in the system. The sample was mounted on a scanning stage. The zone-plate array was parallelized with respect to the sample. The gap between the ZPA and the sample was set to the focal length using the confocal signal. This technique is described later in this chapter. Then, custom software written in LabView was engaged to control the stage motion. Intensity data from the CCD camera was collected asynchronously and the position of the stage was queried in with each frame of data collected. Thus, a bank of data on the detector signal for each zone plate as a function of the lateral scan coordinates was built up. The knife-edge scan data was averaged over several linescans to reduce the effect of vibrations and other noise.

The detector signal as a function of the scan coordinate is shown in figure 6-6. The different scans were performed with zone plates of different NA. As expected, the slope of the knife-edge scans increases with increasing NA. The stage-scan step was 50 nm. This ensures very low noise in the data, which is necessary for the reconstruction of the focussed spot. This is described in the next section. Note that the knife-edge scans have steeper transitions for zone plates of higher NA.

The knife-edge scan contains information regarding the PSF of the zone plate. Here, we work out a technique to extract the PSF in a rigorous manner.

Reconstruction of the PSF

Assuming an ideal confocal microscope, the signal on the detector is described by equation (6.4), which is repeated here.

$$U(x,y) = h^2(x,y) \otimes t(x,y) \tag{6.5}$$

where h is the PSF of the zone plate and t is the object (in this case, the knife-edge).

The knife-edge is modeled as a step function.

$$t(x) = \begin{cases} 1 & \text{x¡0} \\ 0 & \text{else} \end{cases} \tag{6.6}$$

Evaluating the convolution in equation (6.5), we get

$$U = -\int_{-\infty}^{\infty} dy' \int_{-\infty}^{x'} dx' h^2(x', y') \tag{6.7}$$

Differentiating the above equation, we get the line-spread function.

$$\frac{dU}{dx'} = -\int_{-\infty}^{\infty} dy' h^2(x', y') \tag{6.8}$$

The line-spread function is the projection of the point-spread function along the scan direction. This is nothing but the Radon transform of h^2, where the transform integral is taken along the X-axis. A rigorous proof for this statement is given in Appendix C. Strictly, we also make the assumption that the PSF is radially symmetric. Borrowing the idea from tomography, the PSF can be reconstructed by taking the *inverse Radon transform* of the line-spread function.

$$h^2(x', y') = \Re^{-1}\{-\frac{dU}{dx'}\} \tag{6.9}$$

where $\Re^{-1}$ denotes the inverse Radon transform. Note that this assumes that the focussed spot is radially symmetric. This is not strictly true when the incident light is not circularly polarized. However, in the experiments, the incident light was close to being circularly polarized and hence, this analysis should give accurate results. The recontructed PSF for a zone plate of NA=0.7 and $\lambda = 400\,\text{nm}$ is shown in figure 6-7.

A cross-section through the center of the PSF is plotted in figure 6-8. The theoretical PSF is shown for comparison. The reconstructed PSF shows good agreement to the theoretical one close to the center, such that the measured FWHM is close to the theoretical one. However, the reconstructed PSF shows significantly more background further from the center. This may be due to background and noise in the detector during the experiment.

This operation is extremely susceptible to noise. Hence, a slow scan and small sampling step is very important. Care was also taken to keep the environment dark to ensure no stray light reached the detector.

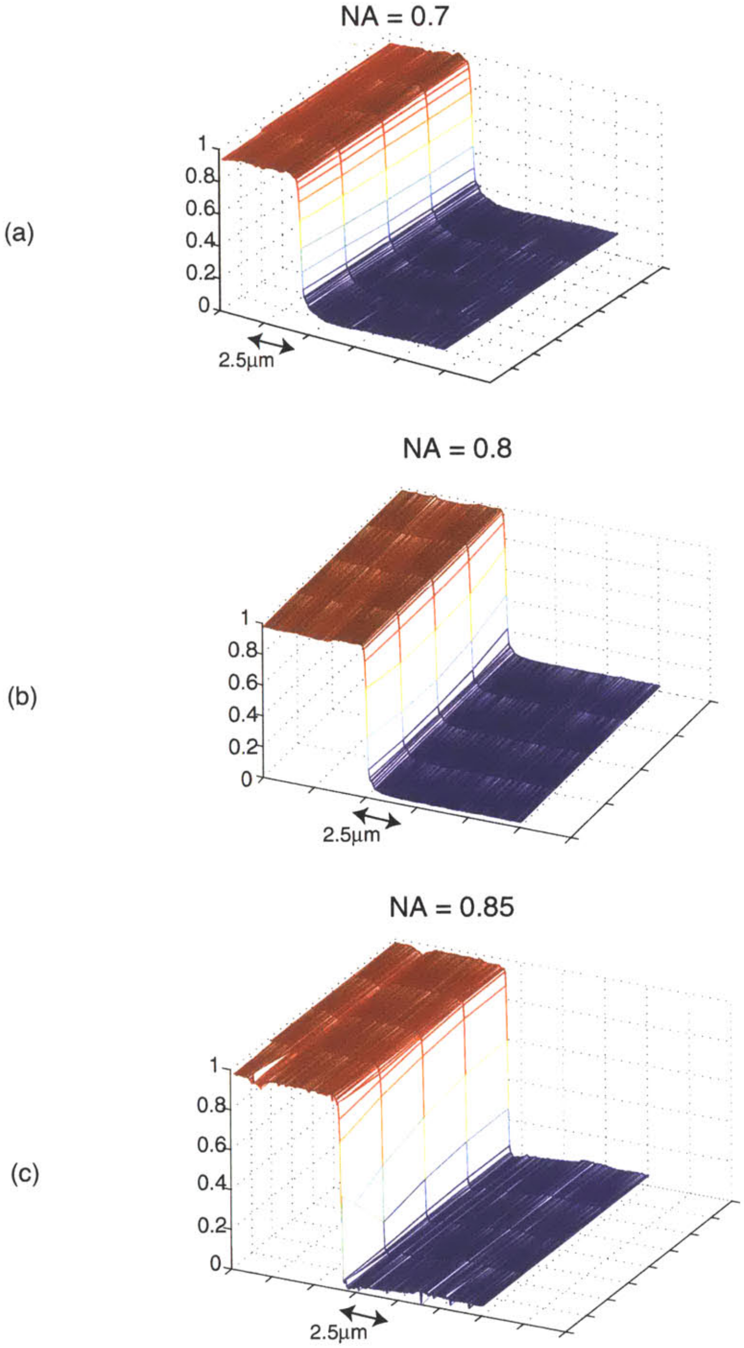


Figure 6-6: *Knife-edge scan results. A knife-edge was scanned across the focussed spot of zone plates of (a) NA=0.7, (b) NA=0.8, and (c) NA=0.85. The confocal signal was plotted as a function of the stage position. Several lines were scanned for averaging in order to reduce the effects of noise during further processing. These experiments were performed at* $\lambda = 400\,nm$.

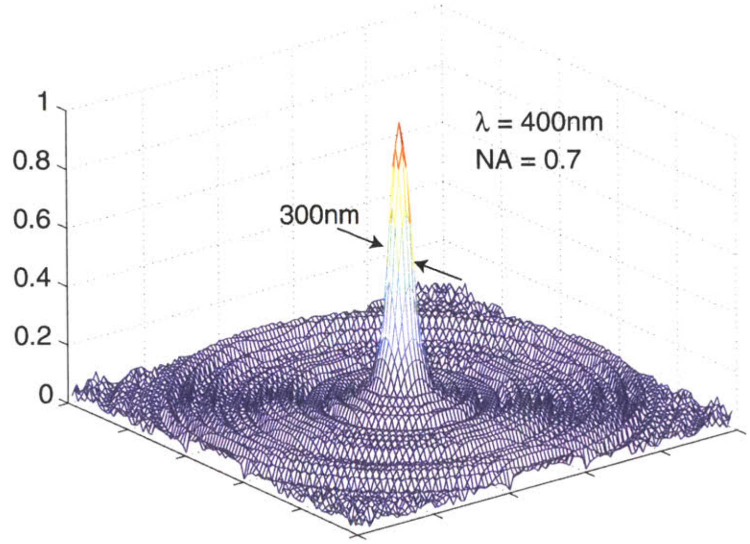


Figure 6-7: *Reconstructed PSF of NA=0.7 zone plate at* $\lambda = 400\,nm$. *The FWHM of the PSF is measured to be about 300 nm*

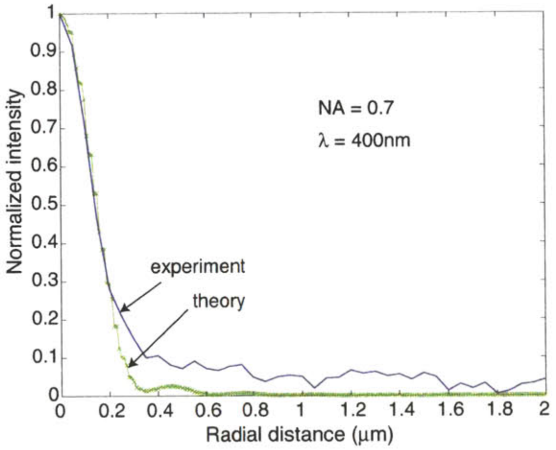


Figure 6-8: *Cross-section through the reconstructed PSF. The cross-section through the PSF shown in figure 6-7 is shown. The theoretical PSF is also shown for comparison. The agreement is good close to the center of the PSF. The data represents the PSF of a zone plate of NA=0.7 and* $\lambda = 400\,nm$.

6.4 Imaging with ZPAM

ZPAM was used as an imaging tool to test its capabilities. A resolution-test sample comprising of groups of concentric-circular gratings of periods ranging from about 1.9μm to 300 nm was fabricated using scanning-electron-beam lithography. A schematic of the test sample is shown in figure 6-9. A scanning-electron micrograph of the largest period gratings are shown on the left and that of the smallest gratings is shown on the right.

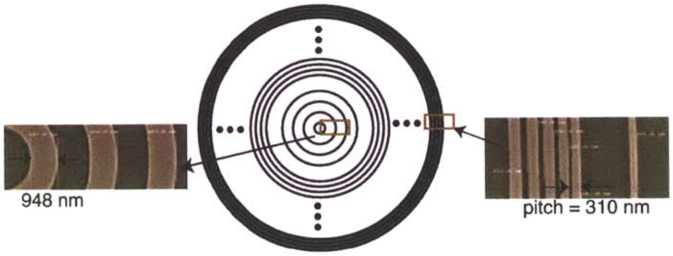


Figure 6-9: *Schematic of resolution-test sample. The sample consists of sets of concentric-circular gratings; each set having a different period. The periods range from about 1.9μ m to about 300 nm. The sample was fabricated by means of scanning-electron-beam lithography.*

This sample was mounted on the stage in the ZPAM system and scanned using zone-plate arrays of different numerical apertures. The intensity measured on the detector (CCD camera) for each zone plate in an array was tabulated as a function of the stage position. The signal from a single zone plate is then plotted as a function of position as shown in figure 6-10 (a). This experiment was repeated with zone plates of different numerical apertures. The measured images are shown in figures 6-10 (a)-(e) for NA=0.7, 0.8m 0.85, 0.9 and 0.95 respectively. As the NA of a zone plate increases, the lateral resolution of the system decreases as shown in figure 6-4. Hence, zone plate of higher NA will be able to resolve smaller grating lines compared to zone plates of lower NA. This is borne out by the images of the gratings of the smallest period (on the right) which get progressively better resolved from NA=0.7 to a NA=0.95 zone plate. This experiment is also a qualitative indication that zone plates of NA as high as 0.95 perform as expected.

A silicon wafer containing various microstructures was used as the sample for the image shown in figure 6-11. The images taken from nine zone plates was stitched together to form an image of nine times the area of the image obtained by one zone plate. Since the nine

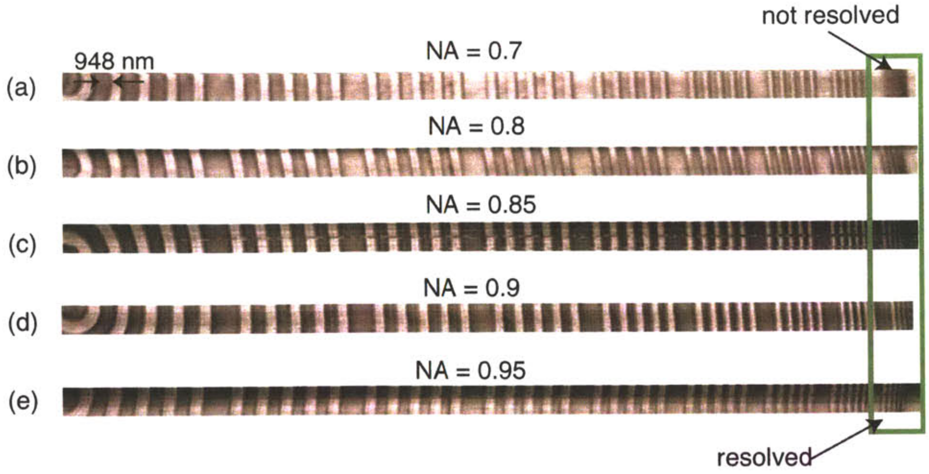


Figure 6-10: *ZPAM images of resolution-test sample. The confocal signal measured by the CCD camera for a single zone plate is plotted as a function of the stage position to form the image. (a)-(e) are ZPAM images taken with zone plates of NA=0.7, 0.8, 0.85, 0.9 and 0.95. As the NA increases the zone plates can resolve smaller lines. Hence, the NA=0.95 zone plate resolves the smallest lines, whereas the NA=0.7 one does not. The experiment was performed at* $\lambda = 400\,nm$.

images are captured simultaneously, this experiment showcases the throughput capabilities of ZPAM. A zoomed in image of a scratch on the surface of the sample is shown on the right. An image of the same scratch taken with an optical microscope is shown beside it for comparison. Due to the small field-of-view (FOV) at high resolution in the optical microscope, it was extremely difficult to find the scratch. In the case of ZPAM, the FOV is decoupled from resolution since it only depends on the number of zone plates in the array. This is an important advantage of scanning-spot systems such as ZPAM. The experiment was perfomed with a 0.7 NA zone plate operating at $\lambda = 442\,\text{nm}$, and a step-size of 350 nm.

In order to show the quality of the images, we performed imaging experiments on a silicon sample which contained various microstructures. Figure 6-12 shows three sets of images taken using ZPAM with a zone plate of NA=0.7 and $\lambda = 442\,\text{nm}$. Each set consisted of 3 by 3 images that were taken using a 3 by 3 element zone-plate array simultaneously. The step-size was increased from 2μm in (a) to 350 nm in (b) and (c). For larger step-sizes, the image can be acquired faster than for smaller step-sizes. However, the resolution of the

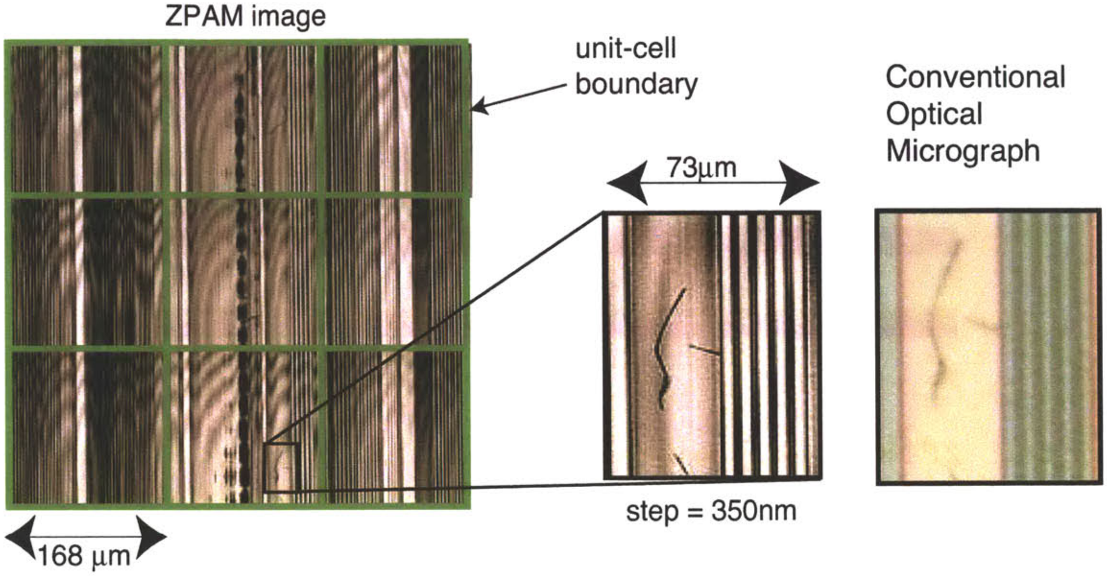


Figure 6-11: *Stitched images from adjacent zone plates to increase the field-of-view. The sample was a silicon wafer which had various microstructures previously fabricated. An array of 3 by 3 zone plates was used to image this sample. The nine images so obtained are tiled (stitched) together to form the large field-of-view. A zoomed-in image of a scratch is shown along with an optical micrograph of the same for comparison. The experiment was perfomed with a 0.7 NA zone plate operating at* $\lambda = 442\,nm$*, and a step-size of 350 nm.*

image is improved by stepping at smaller sizes. This is seen in figures 6-12 (a) and (b). This suggests that quick but coarse images can be obtained before doing a slower higher resolution scan of a region of interest. These experiments were perfomed with zone plates of NA=0.7 and $\lambda = 442\,\text{nm}$. The fringe pattern seen in some of the images in figure 6-12 is an artifact of the ZPAM system. These fringes arise from the interference between the light reflected from the back-surface of the zone-plate array and the light reflected from the sample and collimated by the zone plate.

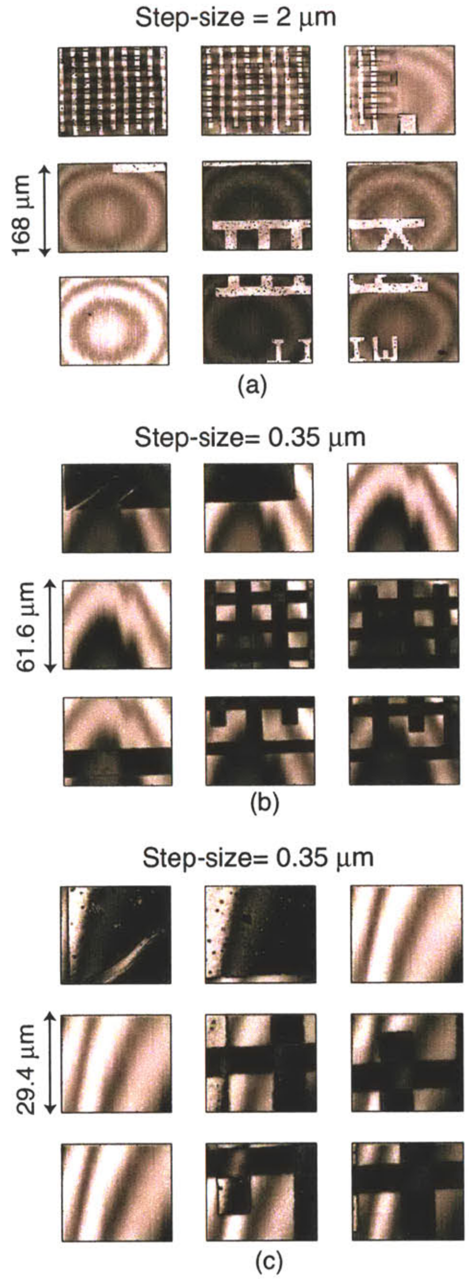


Figure 6-12: *High-resolution ZPAM images. An array of 3 by 3 zone plates was used to image a silicon wafer that contained various microstructures. The imaging was performed at 2 different step-sizes. (a) shows the image of a quick scan at a step-size=2μ m but the entire field (168* μ*m by 168* μ*m. (b) and (c) show images of 2 different portions of the sample imaged at a step-size equal to 350* nm*. Decreasing the step size provides images with higher resolution. These experiments were perfomed with zone plates of NA=0.7 and* $\lambda = 442$ *nm.*

6.5 Applications of ZPAM

In this section, we describe the different applications of ZPAM. ZPAM was devised as a technology to solve the problem of setting the gap between the ZPA and the substrate to the focal length of the zone plate. In this section, we describe this original application of ZPAM. A few other proposals for ZPAM to be used as a complimentary technology to ZPAL are put forward. In addition, ZPAM as a stand-alone technology also offers much promise. Hence, we discuss potential applications there as well.

Setting the working distance in ZPAL

As mentioned earlier, the confocal signal in ZPAM is maximum when the substrate (or sample) is in the focal plane of the ZPA. This signal falls rapidly with any displacement from the focal plane (defocus). The axial resolution of the system was discussed earlier. In the case of ZPAL, this axial resolution provides the precision to which the working distance can be set. In our ZPAL system, this is accomplished by fabricating an extra set of zone plates operating at a different wavelength (in our case, 633 nm) so that the photoresist on the substrate is not affected while imaging. These *satellite zone plates* were placed in the periphery of the zone-plate array that is used for lithography.

Figure 6-13 shows micrographs of two different substrates containing the satellite zone plates. Figure 6-13 (a) shows a scanning-electron micrograph of two sets of zone plates at NA=0.9 and 0.95; each set containing a linear array of 4 zone plates. Around each set of zone plates are a satellite zone plates that operate at $\lambda = 633\,\text{nm}$. Figure 6-13 (b-left) shows an optical micrograph of a large-area ZPA that was back-illuminated with white light. The satellite zone plates on the sides appear in a different color since they were designed for $\lambda = 633\,\text{nm}$. Figure 6-13 (b-right) shows a scanning-electron micrograph of a portion of the ZPA. The zone plates were designed to operate at a wavelength of 400 nm and were placed on a pitch of about $100\,\mu\text{m}$. The confocal signal from three satellite zone plates can also be used to parallelize the ZPA and the substrate.

Calibrating the ZPAL system

ZPAM can be used to calibrate for inter-zone-plate differences within a ZPA. For a large ZPA, the power focussed into a spot can be different in different portions of the ZPA. This could be due to small differences in the fabricated zone plates themselves or due to non-uniform illumination of the ZPA. In either case, the efficiency difference can be compensated by modulating the nominal dose incident on each zone plates by means of grayscaling. In order to do this, the relative efficiency of all the zone plates in the ZPA must be mapped out. This can be done in a parallel fashion using ZPAM. This is accomplished by placing a highly-reflective substrate in the focal plane of the ZPA, and measuring the relative confocal signal from all the zone plates. In this manner, the relative efficiencies of a large number of zone plates can be measured very quickly. The relative efficiencies can be tabulated and used as a look-up table for compensating the dose prior to writing.

The stitching of patterns across the zone-plate unit-cells depend on the placement of the centers of the zone plates and the angular misalignment between the ZPA-axes and the stage-axes. Both these errors can be measured by scanning a large-area grating [1] over the entire zone-plate unit-cell and measuring the misalignment at the boundaries of the unit-cell. An experiment which illustrates this was discussed in Chapter 4. The errors for large ZPAs can be measured quickly because of the parallel nature of ZPAM. These errors can be tabulated and used as a look-up table to compensate the pattern by pre-distorting it prior to actual exposure.

Imaging applications

Since ZPAM is a high-throughput, high-resolution confocal microscope, it has several potential applications in microscopy. In the semiconductor industry, ZPAM may be used for mask and wafer inspection. The defect-resolution capability of ZPAM can be further improved by means of super-resolution algorithms and pattern-matching software. Flourescence imaging and confocal microscopy for biological samples could be important imaging applications for ZPAM as well.

---

[1] Gratings that are coherent over a large area can be patterned by means of interference lithography [50].

Optical Tweezers

An optical tweezer is formed by focussing light to a very small spot using a high numerical aperture lens, and using the electromagnetic gradient forces to manipulate microscopic particles. This tool is used extensively to study the properties of macromolecular objects (colloids, polymers, membranes) and other biological systems [51]. ZPAM would enable an ideal extension of this concept which includes a large number of focussed spots manipulating a large number of microscopic particles at the same time. By using the multiplexers, the angle of illumination to the zone plates may be controlled independently, which enables the focussed spot to be swivelled around the optical axis. This was described as a technique to correct for small stage-placement errors in ZPAL in chapter 4. This technique can be used for optical-tweezer arrays to stretch DNA molecules for example, since the two ends of the molecule could be moved in different directions using the focussed spots of two zone plates (see figure 6-14).

As shown in the course of this research, zone-plate arrays with zone plates of very high numerical aperture can be fabricated reliably. High-NA is useful as it provides large trapping forces. The force is proportional to the gradient of the electric field. Moreover, since diffractive optics can be designed to provide focal spots of various shapes [52], it is possible to optimize the spot for the purposes of particle trapping.

Parallel micro-spectroscopy

Since zone plates are diffractive elements, they exhibit strong wavelength-dependent focussing. This property may be used to measure the spectrum of microscopic samples. This idea is best explained using an example. Consider a sample which is an array of chemical cells, each cell having a different reaction which emits fluorescence at a characteristic wavelength when excited. ZPAM can be used to obtain images of the sample at a given wavelength. By moving the confocal pinhole and repeating the scan, the image of the sample at a different wavelength can be obtained. This can be understood as follows. The radii of the zones of a low-NA zone plate is defined as

$$r_n = \sqrt{n\lambda f} \tag{6.10}$$

where n is the zone number, λ is the wavelength, and f is the focal length. For a different wavelength, λ', the same zone plate will focus at a different focal length, f', such that

$$r_n = \sqrt{n\lambda' f'} \tag{6.11}$$

This implies that in ZPAM, the light emanating from the sample will have different divergence angles depending on the wavelength, after being collected by the zone plate. The signals from the different wavelengths will then be focussed at different distances along the optical axis as shown in figure 6-15. By simply translating the pinhole along the optical axis, the signal as a function of wavelength for every point on the sample can be obtained. This can be done fairly quickly as signals from a large number of zone plaes can be obtained simultaneously.

In conclusion, ZPAM is a technology that has pontential applications in several different areas. The power of diffractive optics coupled with the massive parallelization due to the array provides significant advantages in many of these applications.

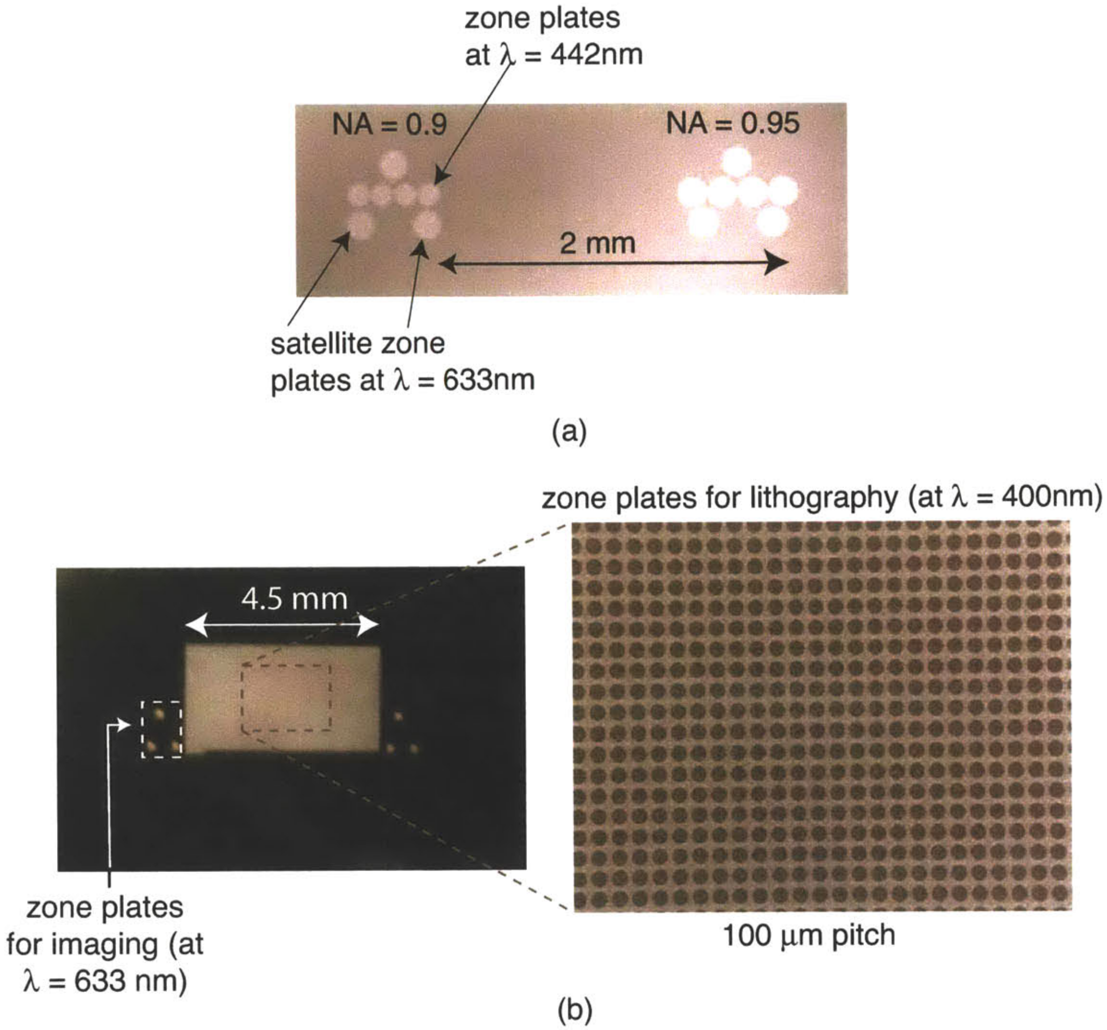


Figure 6-13: *Satellite zone plates for gapping. Satellite zone plates operating at* $\lambda = 633\,nm$ *are used to set the gap between the ZPA and the substrate. (a) Scanning-electron micrograph of zone plates of NA=0.9 and 0.95 operating at* $\lambda = 442\,nm$*. (b-left) Optical micrograph of a ZPA consisting of 1000 by 1000 zone plates. The satellite zone plates are visible at a different color. (b-right) Scanning-electron micrograph of a portion of the ZPA. The zone plates of NA=0.7 and operating at* $\lambda = 400\,nm$ *are on a grid of pitch 100* μm*.*

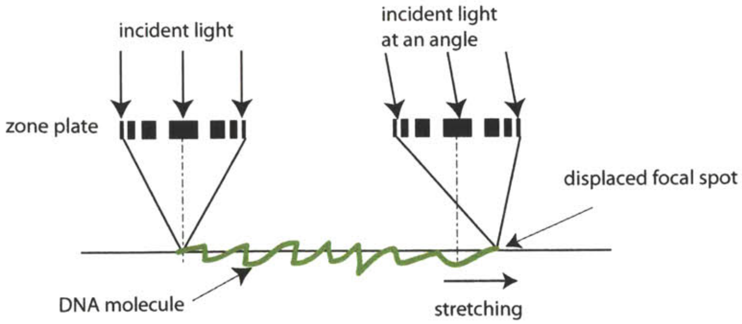


Figure 6-14: *Schematic of parllalel-optical tweezers. The focussed spot of a zone plate can be used to trap microscopic particles. By changing the angle of incidence on a zone plate, the focussed spot may be swivelled around the optical axis. Thus, particles can be manipulated in parallel by a large number of zone plates in a ZPA at the same time.*

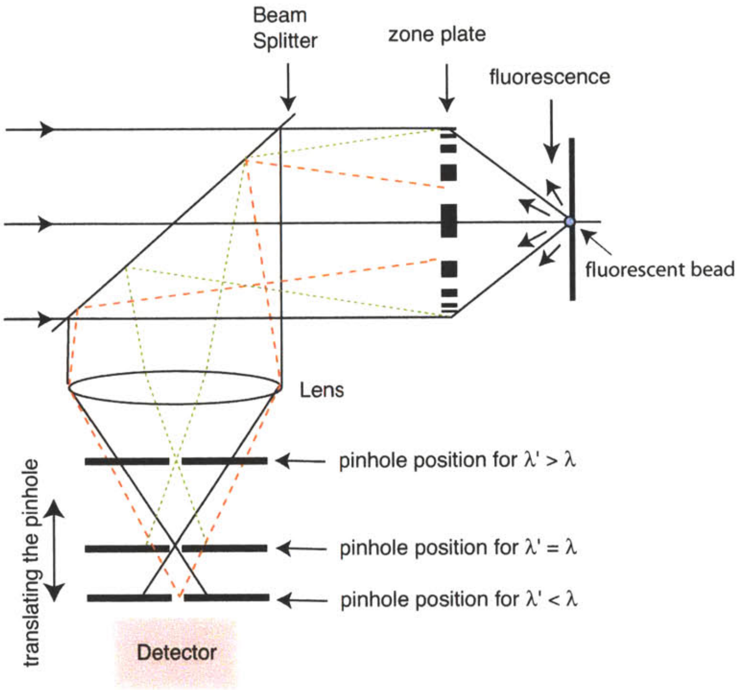


Figure 6-15: *Schematic of a parllalel spectroscope. Being a diffractive element, the zone plate is highly dependent on the wavelength. For low-NA zone plates, different wavelength simply implies a shift in the focallength. This can be used to separate the signals at different wavelengths by translating the confocal pinhole to the planes at which the signals from the different wavelengths are focussed. With a large number of zone plates scanning the sample at the same time, ar large-area spectral map of the sample can be obtained.*

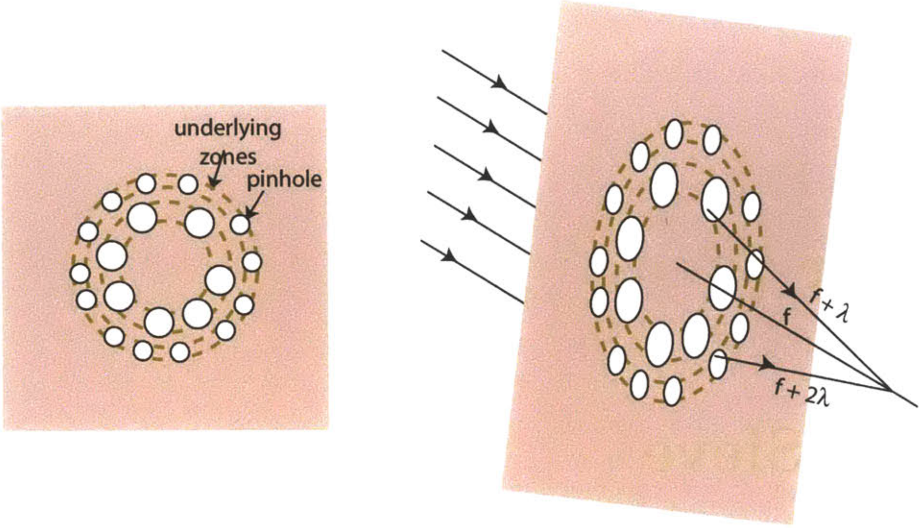
Chapter 7

Photon Sieve

Diffractive elements are very useful because they can be customized to meet certain requirements. One such element that was proposed to improve the focussing properties of zone plates is the *photon sieve* [53]. In this chapter, we begin by describing the photon sieve and its principle of operation. We analyze the performance of the photon sieve using simulations. We also describe methods of fabricating these photon sieves and present the first experimental results showing focussing and lithography using high-numerical-aperture photon sieves.

7.1 Introduction

The geometry of the photon sieve is based on the geometry of an underlying zone plate. Instead of concentric zones, the photon sieve is comprised of pinholes arranged within the boundaries of the zones of the underlying zone plate. This arrangement of the holes ensures that a portion of the light that is diffracted from the pinholes interferes constructively at the focus of the underlying zone plate. A schematic of a photon sieve is shown in figure 7-1. We begin by describing the design procedure for a photon sieve.



(a) (b)

Figure 7-1: *Schematic of a photon sieve. (a) Top view of an amplitude photon sieve. The holes in white are transmissive whereas the gray areas are opaque. The zones of the underlying zone plate are marked with dashed circles. (b) Plane wave incident on a photon sieve is focussed to a point on the optical axis. The path lengths of the holes in the first two zones are indicated.*

7.2 Photon Sieve Design

A uniform plane wave incident on the photon sieve is brought to a focus just as in the case of the underlying zone plate. The pinholes are arranged such that all the pinholes within one zone have the same radius. In addition, all the pinholes are centered on the corresponding zone. The coordinates of the center of a pinhole are therefore given by

$$\begin{aligned} x_{np} &= \rho_n \cos\theta & (7.1) \\ y_{np} &= \rho_n \sin\theta & (7.2) \end{aligned}$$

where the subscript np refers to the p^{th} pinhole within the n^{th} zone, and ρ_n is given by

$$\begin{aligned} \rho_n &= \frac{r_n + r_{n+1}}{2} & (7.3) \\ &= \frac{\sqrt{n\lambda f + (\frac{n\lambda}{2})^2} + \sqrt{(n+1)\lambda f + (\frac{(n+1)\lambda}{2})^2}}{2} & (7.4) \end{aligned}$$

where the equations of the zone radii from chapter 2 are used. The angular coordinate, θ is chosen in a random manner. This is to ensure that there is no azimuthal symmetry in the photon sieve. It will be shown shortly that such a random distribution of pinholes gets rid of higher diffraction orders.

Next, the radius of each pinhole is chosen such that its field contribution at the focus is maximum. The field computation at the focus is a complicated problem as was seen in the case of zone plates. It can be shown that the field contribution is maximized when the ratio of the diameter of the pinhole to the width of the underlying zone is equal to about 1.5, 3.5, 5.5, 7.5, ... This is shown using an analytical expression under far-field approximation as well as using numerical simulations under the Fresnel-Kirchoff regime in appendix D. Using the Fresnel-Kirchoff diffraction theory, we can calculate the field at the focus due to a single pinhole centered on the nth zone of the underlying zone plate. Figure 7-2 shows plots of the field contribution at the focus as a function of the ratio of the pinhole-diameter, $2a_n$ to the underlying zone-width. The underlying zone plate in the simulation had 50 zones and a focal length, f equal to 100 times the wavelength, λ. Figure 7-2 (a) shows the field contribution for a pinhole centered on the 40^{th} zone, while figure 7-2 (b) shows the contribution for a pinhole centered on the 45^{th} zone. Note that the field (both real and imaginary parts) differ in phase by π from the 40^{th} to the 45^{th} zone. This is expected since adjacent zones are chosen such that the path lengths differ by half a wavelength. In both cases, the fields show extrema (maxima and minima) when the ratio of diameter of pinhole to zone-width is equal to about 1.5, 3.5, 5.5, 7.5, etc, with alternating maxima and minima. This is true for pinholes in all the zones. In order to ensure that the fields from all the pinholes interfere constructively, pinholes centered on even zones must have a normalized diameter equal to 3.5, 7.5, etc, and pinholes centered on odd zones must have a normalized diameter equal to 1.5, 5.5, etc or vice-versa, where the normalized diameter, K is defined as

$$K = \frac{2a_n}{w_n} \tag{7.5}$$

This behavior of the field can be explained using a simple model shown in figure 7-3. The figure shows a small portion of a photon-sieve and its underlying zone plate. For simplicity,

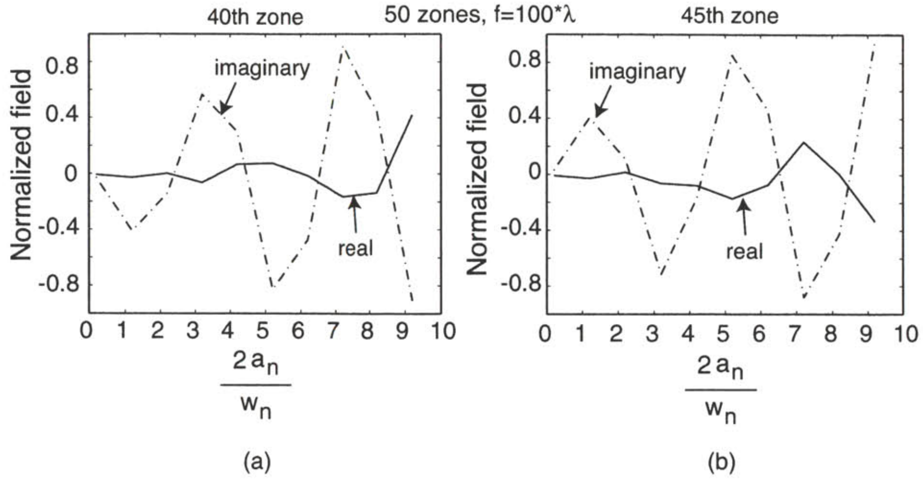


Figure 7-2: *Field contributions at the focus due to a single pinhole. The real and imaginary parts of the field at the focus due to a single pinhole is plotted for increasing pinhole radius for a pinhole centered on the 40th zone (a) and centered on the 45th zone (b). The zone plate was assumed to have 50 zones and its focal length was normalized such that* $f = 100\lambda$.

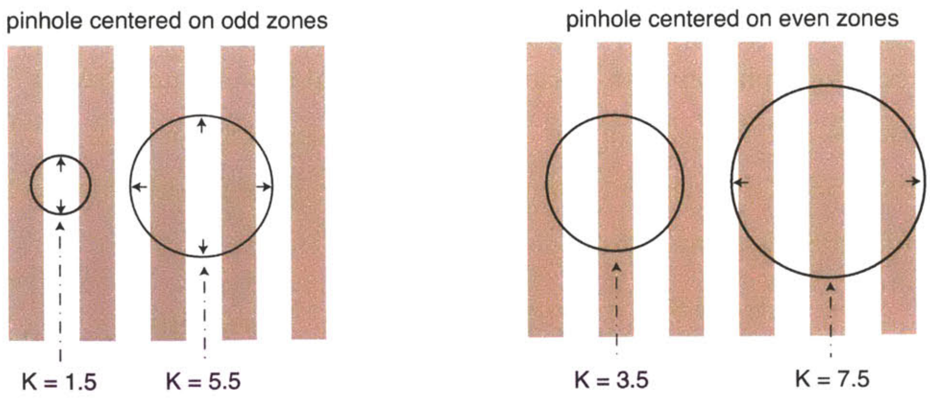
the odd zones are shown in white and the even zones are in gray. Light from the white and gray regions have a path-length difference equal to λ at the focus, and hence, they interfere destructively. For a pinhole centered on an odd zone, the field increases as K increases until it reaches about 1.5. This is because the area of the white region increases until it is countered by the area of the gray region. Increasing K further increases the field again when the "white area" becomes greater than the "gray area" within the pinhole. The opposite is true for pinholes centered about the gray region. The zeros of the field occur when the areas of the white and gray regions are equal.

In order to ensure equal contributions of light from pinholes in all the zones, the total area of the pinholes in a given zone must be equal to that in any other zone. This is the same requirement as in a zone plate, where the areas of the zones have to be equal[1]. If there are P_n pinholes within a zone, n, then the total transmitting area within that zone is equal to

$$A_n = P_n \times \pi a_n^2 \tag{7.6}$$

---

[1] This is not strictly true for high-NA zone plates.



(a) (b)

Figure 7-3: *Schematic explaining the focal field behavior. A small portion of a photon-sieve and its underlying zone plate is shown. The white regions denote the odd zones while the gray regions are the even zones. At the focus, light from these two regions interfere destructively. Hence, the field at the focus due to a pinhole depends on the relative area of the gray and white regions within the pinhole. For example, in (a), the field contribution from the pinhole is maximised when K=1.5 because beyond K=1.5, the gray area starts to counter the effect of the white areas. (a) and (b) show examples of pinholes centered on the odd and even zones respectively.*

The design procedure for a photon sieve begins with the design of the underlying zone plate. This follows the same steps as described in chapter 2. The total number of pinholes is divided among the underlying zones such that the transmitted areas are about the same based on equation (7.6). For each zone, the angular position of the centers of the pinholes are chosen randomly. The pinholes have to be placed such that there is no overlap. Geometrical algorithms were devloped to perform the check for overlap very efficiently. Finally, the design of the photon sieve is output into a file whose format can be read by the electron-beam lithography tool.

In a low-NA zone plate, the zones are periodic in r^2, where r is the radial distance from the center. This gives rise to higher diffraction orders, which contribute to background during exposure. The random distribution of the pinholes azimuthally ensures that there is no symmetry in the transmission function of the photon sieve along any radius. This suppresses diffraction into higher orders. In our designs, the angles were randomly chosen using a random number generator available in the standard C-language library.

This concludes the design procedure for a photon sieve. Figure 7-4 (a) shows the transmission function of a photon sieve designed to operate at $\lambda = 400\,\text{nm}$, NA=0.75, $K = 1.5$ in the even zones and $K = 3.5$ in the odd zones, and containing a total of 738 pinholes. The simulated PSF of this sieve is shown in figure 7-4 (b). The full-width at half-maximum (FWHM) of the PSF is about 326 nm.

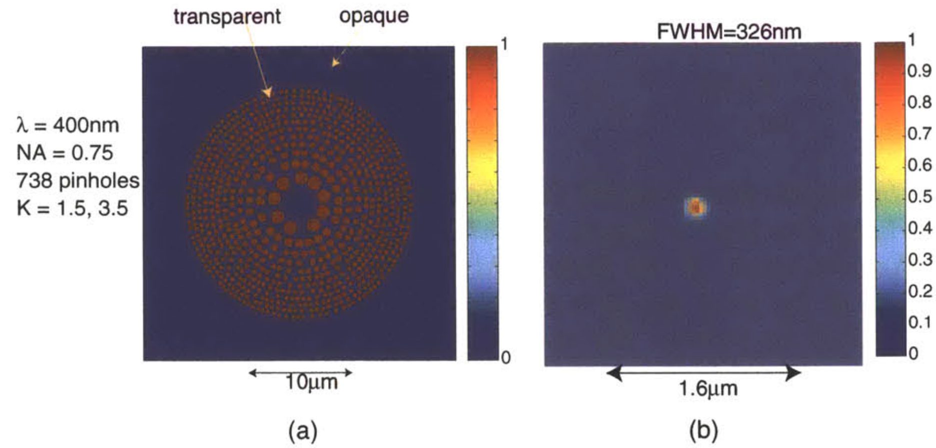


Figure 7-4: *Amplitude photon sieve. (a) Tranmission function of a photon sieve comprised of transparent pinholes in an opaque aperture. This element is designed to operate at* $\lambda = 400$ *nm with NA=0.75,* $K = 1.5$ *for even zones and* $K = 3.5$ *for odd zones, and contains 738 pinholes. (b) Simulated PSF of the sieve. The FWHM is about 326 nm.*

Figure 7-5 shows the scanning-electron micrograph of a photon sieve pattern that was written using scanning-electron-beam lithography. This pattern was exposed in PMMA. The parameters of this sieve are $\lambda = 400\,\text{nm}$, NA=0.7, f=40 μm. Note that the size of the smallest pinholes is about 480 nm.

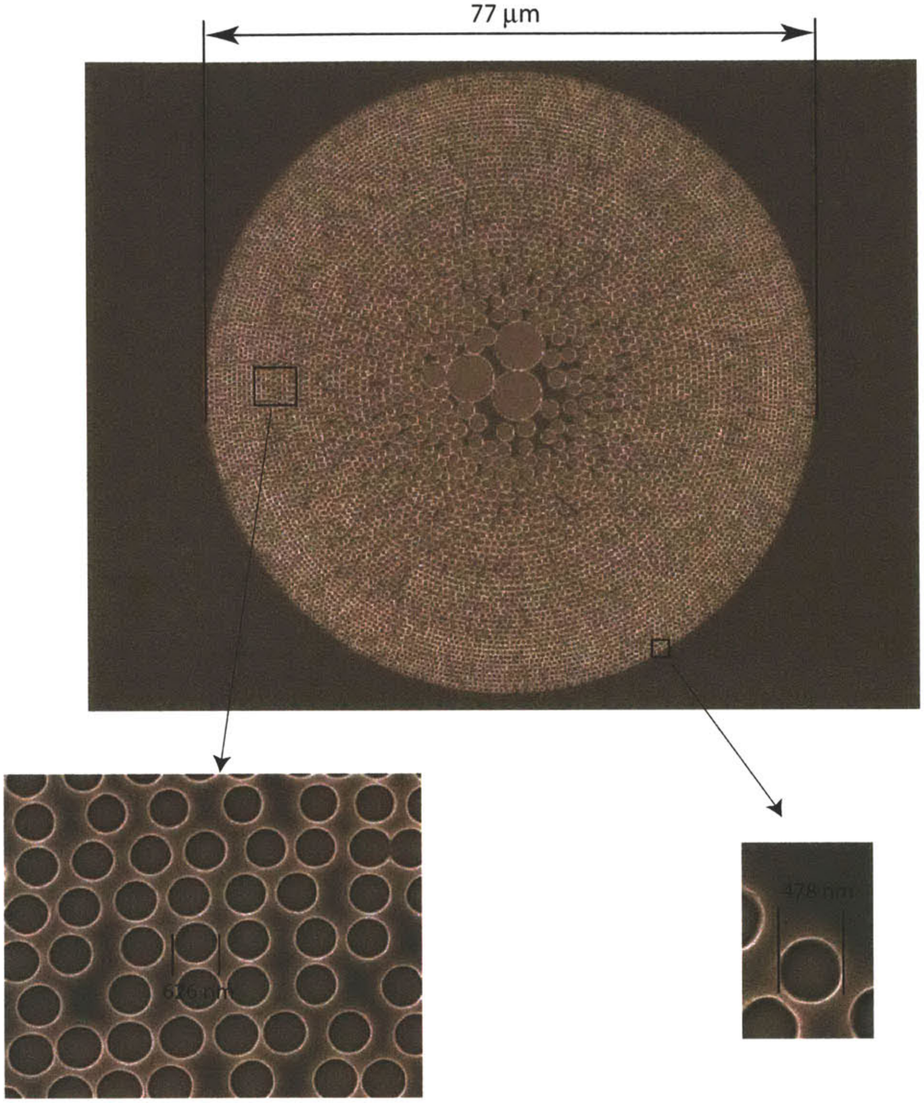


Figure 7-5: *Scanning-electron micrograph of a photon sieve. The pattern was exposed using scanning-electron beam lithography and PMMA as the resist. The pattern is gold-coated to improve contrast in the scanning-electron microscope (SEM).*

7.2.1 Phase-Photon Sieves

In the course of this research, we came up with different versions of the photon sieve which utilize phase-shifting with the goal to improve the focussing efficiency.

The first type of phase-photon sieve is one where the pinholes in alternate zones are phase shifted by π. The phase-shifted pinholes must then have the same K as the other pinholes, since the phase-shift automatically accounts for the path-length difference, and ensures constructive interference of light from all pinholes at the focus (see figure 7-6 (a)). We call this element, the alternating-phase pinhole sieve. In this case, K should be the same for all the zones equal to either one of $\{1.5, 5.5, ...\}$ or one of $\{3.5, 7.5, ...\}$. The main advantage of this phase-shifting photon sieve is the suppression of the zero order. Since the design procedure ensures that the area of the pinholes in any one zone is the same as in any other zone, the total area that is phase-shifted is equal to the total area that is not phase-shifted. This ensures that the zero-order diffraction is suppressed just as in a phase zone plate. Figure 7-6 (a) shows the transmission funtion of a phase photon sieve of NA=0.75 operating at $\lambda = 400\,\text{nm}$, $K = 1.5$, and containing 889 pinholes. Figure 7-6 (b) shows the PSF of this photon sieve. The phase sieve focusses the light to a spot of FWHM diameter of about 300 nm.

The second form of the photon sieve is one where the entire area of the photon sieve is transparent. Pinholes of a phase shift of π are etched into the substrate to define the sieve. This is an interesting element, since like a phase zone plate, 100% of the light is passing through it. This element is termed the phase-photon sieve. If the areas of the phase-shifted pinholes are matched to the rest of the area, the zero order may be suppressed. The random arrangemement of the pinholes will suppress unwanted diffraction orders as well. This suggests that a high percentage of the incident light may be focussed into the spot. The radii of the pinholes are chosen in the same manner as for the amplitude photon sieve. Figure 7-7 shows the transmission function and corresponding PSF of a phase sieve designed to operate at $\lambda = 400\,\text{nm}$, $K = 1.5$ in the even zones, $K = 3.5$ in the odd zones, and containing a total of 739 pinholes. The FWHM of the focussed spot is about 239 nm.

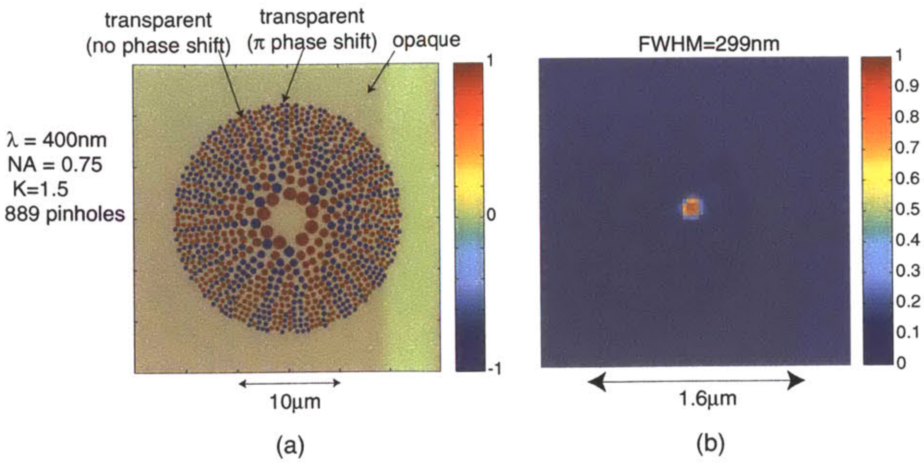


Figure 7-6: *Alternating-phase pinhole sieve. (a) Transmission function of a photon sieve composed of pinholes in an opaque substrate, where pinholes in alternate zones have a relative phase difference of* π. *The parameters used for the design are* $\lambda = 400\,nm$, $K = 1.5$, *NA=0.75, and it contains 889 pinholes. (b) The simulated PSF. The FWHM of the PSF is about 300 nm.*

7.2.2 Size of the Focussed Spot in a Photon Sieve

Focussing by a photon sieve is based on the same principles as focussing by the underlying zone plate. Hence, the size of the focussed spot is approximately equal to the width of the smallest zone in the underlying zone plate. This has an interesting implication. Since, the size of the pinhole is a factor of K times larger than the width of the underlying zone. The focussed spot is smaller than the smallest feature on the photon sieve. This reduction factor, K, is perhaps, the most important advantage of the photon sieve. This relaxes the fabriation requirements for the photon sieve considerably compared to a zone plate of the same NA. This is especially important when the zones of the zone plate become close to the limit of the electron-beam lithography system (about 20 nm).

7.2.3 Depth of Focus

The depth-of-focus (DOF) depends on the angle subtended by the diffracted rays, which in turn depends on the numerical aperture. Hence, one would expect the DOF of the photon sieve to be the same as that of the zone plate. This is indeed found to be true as illustrated

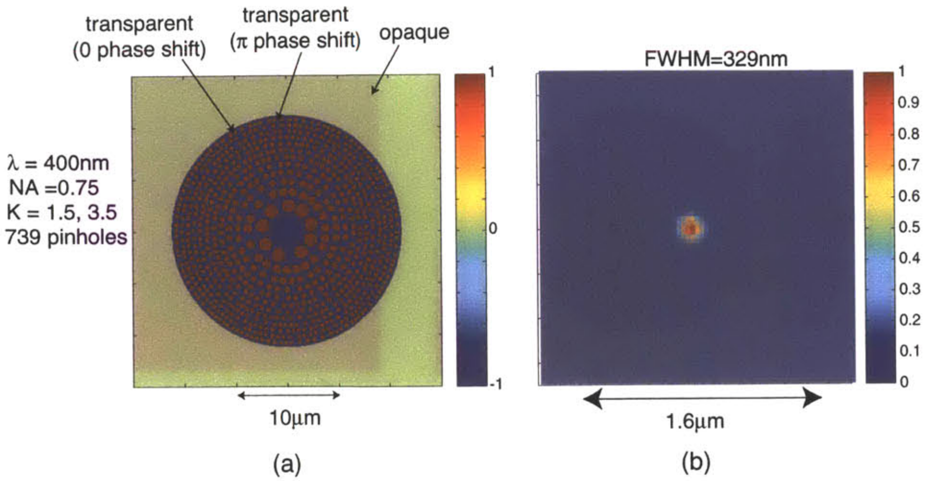


Figure 7-7: *Phase-photon sieve. (a) Transmission function of a photon sieve composed of pinholes in a transparent aperture, where the pinholes have a π-phase shift with respect to the rest of the aperture. The parameters used for the design are $\lambda = 400\,nm$, $K = 1.5$ for even zones, and $K = 3.5$ for the odd zones, NA=0.75, and 739 pinholes. (b) The simulated PSF. The FWHM of the PSF is about 239 nm.*

in figure 7-8 where the normalized intensity along the axial direction is plotted for the 3 different kinds of photon sieves namely amplitude, alternating-phase pinhole sieve (phase 1), and phase-photon sieve (phase 2). The DOF of the simulated data is about 545 nm, which is close to the value for a zone plate.

7.2.4 Apodization

Apodization is a technique used in lenses to suppress undesired sidelobes in the PSF. In a binary diffractive element such as a zone plate, this can be achieved by modulating the duty cycle of the zones, for example, making the open zones larger than the opaque zones with increasing zone radius in an amplitude zone plate will suppress the sidelobes. This phenomenon can be understood by invoking the principles of signal processing. A lens essentially acts as a low-pass filter, where the aperture of the lens limits the spatial frequencies (propagating angles) of the transmitted light. Changing the sharp cut-off imposed by the lens aperture into a smooth transition reduces the sidelobes in the PSF. This is achieved by apodization.

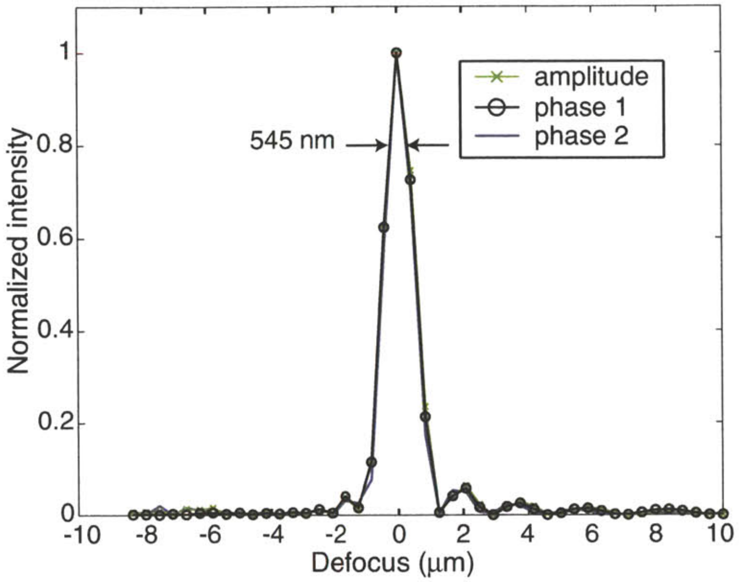


Figure 7-8: *Depth-of-focus of a photon sieve. The intensity at the center of the spot at different defocus from the focal plane is plotted. The depth-of-focus is usually calculated as the defocus where the peak value falls about 20%. The data gives a DOF of about 545 nm. The parameters of the sieve were* $\lambda = 400\,nm$*, and NA=0.75. For the amplitude sieve,* $K = 1.5$ *for even zones and* $K = 3.5$ *for the odd zones. For the phase sieves,* $K = 1.5$ *for all zones for type 1, and* $K = 1.5$ *(for even zones), and* $K = 3.5$ *(for odd zones) for type 2.*

In an amplitude photon sieve, this apodization can be achieved by modulating the total area of the pinholes associated with a zone.

7.2.5 Efficiency of the Photon Sieve

In the case of amplitude-photon sieves, the pinholes are transmissive, whereas the rest of the area of the element is blocked. It is instructive to see what percentage of the area of the sieve is transparent. Recall that in the amplitude zone plate, the transparent area was about half of the total area. We designed amplitude photon sieves operating at $\lambda = 400\,\text{nm}$, and various numerical apertures. The total area of all the pinholes was calculated and expressed as a percentage of the aperture area. This is plotted as a function of NA in figure 7-9. Note that percentage of transmissive area is between 14% and 22%. This suggests that the focussing efficiency of the photon sieve is smaller than that of the underlying zone plate by at least a factor of 2. It is also noticed that the percentage of the transmissive area decreases as the NA of the photon sieve increases. This is due to the fact that as the NA increases, the number of zones and hence the area covered by the entire zone plate increases faster than the area covered by the pinholes. The number of pinholes is shown in parenthesis beside each data point. The design was performed for a photon sieve of focal length equal to 40 μm for NA=0.7, 0.8, 0.9, and 30 μm for NA=0.95. The value of K was designed to be 2 for the odd zones and 4 for the even zones. The rationale for choosing these numbers will be explained shortly.

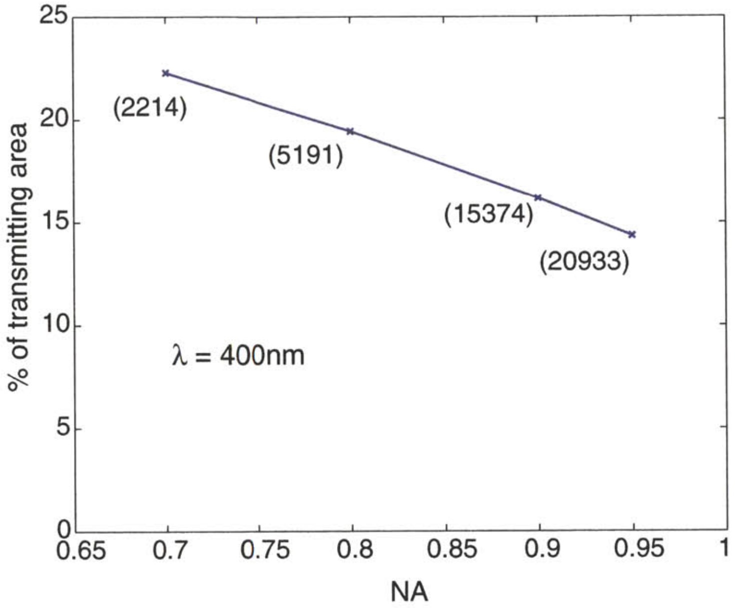


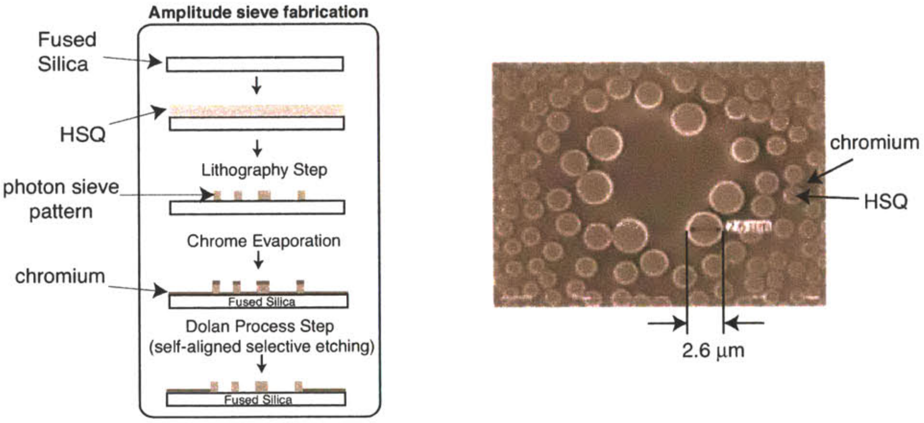
Figure 7-9: *Percentage of transmissive area in amplitude photon sieves. The ratio of the area occupied by the pinholes to the total area of the underlying zone plate is plotted as a percentage for photon sieves of different NAs. The designs were performed assuming* $\lambda = 400$ *nm and focal length =* *40* μ*m (for NA=0.7, 0.8, and 0.9). For NA=0.95, the focal length was 30* μ*m. The* K *was chosen to be equal to 2 for odd zones and 4 for even zones. The total number of pinholes in each design is indicated beside each data point in parenthesis.*

7.3 Fabrication of Photon Sieves

In the previous section, we described the design of 3 different types of photon sieves. Here, we describe techniques that may be used to fabricate these elements.

7.3.1 Amplitude-Photon Sieve

The amplitude photon sieve consists of open pinholes in a transparent substrate. This element can be fabricated by means of the same technique used to fabricate phase zone plates. HSQ, a negative electron-beam resist is spun on top of a fused silica substrate. The pattern of the open pinholes is written using scanning electron-beam lithography. After development, cylinders of HSQ are left standing on the substrate. A layer of chromium is then evaporated on top of the substrate. The chromium covering the raised cylinders is electrically insulated from the chromium covering the rest of the substrate. Using a selective electrochemical wet-etching process (Fulton-Dolan Process), the chromium can be selectively removed from the top of the cylinders. This leaves the final ampitude sieve on the substrate. A scanning-electron micrograph of the central region of an amplitude photon sieve is shown in figure 7-10 (b). The parameters of this design were $\lambda = 400\,\text{nm}$, NA=0.7, f=40 μm, and the sieve contained 2214 pinholes.



(a) (b)

Figure 7-10: *Fabrication of an amplitude-photon sieve. (a) Steps involved in the fabrication process for an amplitude-photon sieve. The method is exactly the same as that used to fabricate a phase zone plate. (b) Scanning-electron micrograph of a fabricated amplitude photon sieve of NA=0.7, operating at* $\lambda = 400\,nm$ *with focal length equal to* $40\,\mu m$ *and containing 2214 pinholes.*

7.3.2 Alternating-Phase Pinhole Sieve

The alternating-phase pinhole sieve is an element consisting of open pinholes with different phase shifts within alternate zones. The different phase shifts are obtained by a difference in the height of the cylinders used to create the photon sieve. The different heights of the cylinders are obtained by exposing HSQ at two different doses using scanning-electron beam lithography. This technique has been proposed to fabricate analog (blazed) diffractive elements [28]. After development, cylinders of two different heights in HSQ are obtained. The rest of the area (not occupied) by the cylinders can be blocked by evaporating a layer of chromium, and using the Fulton-Dolan process to remove the chromium from the top of the cylinders. A schematic illustrating the steps for this process is shown in figure 7-11 (a).

7.3.3 Phase-photon sieve

The phase-photon sieve is an element consisting of pinholes that are phase-shifted with respect to the rest of the aperture of the sieve. The entire aperture of this element is transmissive. Again, the pattern of the sieve can be written on a layer of HSQ using scanning-electron beam lithography. However, in addition to the pinhole pattern, a thin ring marking the aperture of the pinhole is also written. The width of this ring is much smaller than the wavelength, so that it has very little effect on the optical performance of the element. Then, chromium is evaporated over the entire substrate. Now, chromium inside the ring (within the sieve aperture) is electrically isolated from the rest of the chromium. The Fulton-Dolan process can then be used to remove the chromium from only inside the ring. This leaves behind the phase-shifted photon sieve. A schematic illustrating the steps for this process is shown in figure 7-11 (b).

Amplitude-photon sieves were fabricated and tested using the in-house ZPAL system, where the zone plates were replaced with photon sieves. The experimental results are discussed next.

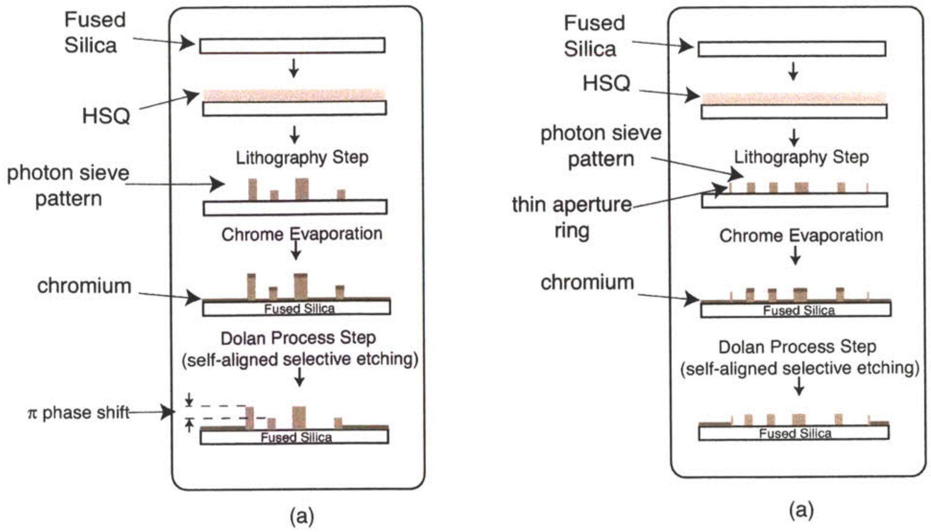
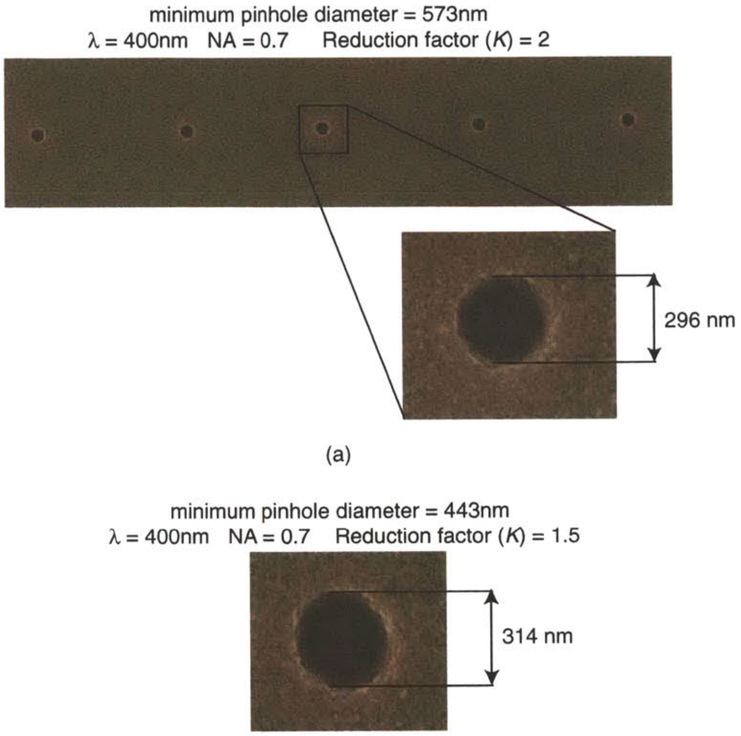


Figure 7-11: *Fabrication of phase-shifted photon sieves. (a) Steps involved in the fabrication of the alternating-phase pinhole sieve. The final sieve consists of pinholes in adjacent zones having π phase shift with respect to one anothers. (b) Steps involved in the fabrication of the phase-photon sieve. The final element consists of pinholes that are phase-shifted with respect to the rest of the aperture. These methods have not yet been put into practise.*

7.4 Lithography using Photon Sieve

Amplitude-photon sieves were fabricated at numerical apertures of 0.7, 0.8 and 0.9, operating at $\lambda = 400\,\text{nm}$. The zone plates in the ZPAL system were replaced by the photon sieves. The procedure for lithography was exactly the same as described in Chapter 4. The patterns after development, were inspected in a scanning-electron microscope.

Figure 7-12 shows scanning-electron micrographs of single-spot exposures performed using two photon sieves of NA=0.7 but different values for K. In figure 7-12 (a), the photon sieve has $K = 2$, i.e. the minimum pinhole diameter of this sieve was 2 times the width of the smallest underlying zone. Accordingly, the exposed spots were about 296 nm in diameter, which is about half of the diameter of the smallest pinhole in the sieve (573 nm). In figure 7-12 (b), the photon sieve has $K = 1.5$. The corresponding exposed spots had a diameter of about 314 nm, which is close to the expected spot size of about 1.5 times smaller than the diameter of the smallest pinhole (443 nm). These are, as far as we aware, the first experimental verification of the performance of a photon sieve at $\lambda = 400\,\text{nm}$, with a reduction factor as high as 2.



(a)

(b)

Figure 7-12: *Patterning of spots with photon sieves. Scanning-electron micrographs of single-exposure spots patterned with two photon sieves of the same NA (0.7) but different* K*. (a)* $K = 2$*. Hence, the exposed spot is half the size of the smallest pinhole on the sieve. (b)* $K = 1.5$*. The exposed spot is about 1.5 times smaller than the smallest pinhole in the sieve. The experiments were performed at* $\lambda = 400\,nm$*.*

Figure 7-13 shows scanning-electron micrographs of dense lines patterned using photon sieves of two NAs (0.7 and 0.8) operating at $\lambda = 400\,\text{nm}$ and the reduction factor, $K = 2$. In figure 7-13 (a), the photon sieve of NA=0.7 has a minimum pinhole diameter of about 570 nm. The smallest lines patterned using this sieve were about 270 nm. The lines in figure 7-13 (b) were patterned using a photon sieve of NA=0.8 with a minimum pinhole diameter of about 500 nm. The smallest lines patterned in this case was about 244 nm. Clearly, the photon sieves work as expected.

Figure 7-14 shows scanning-electron micrographs of dense lines of period 400 nm written using two photon sieves of NA=0.8 and 0.9 operating at $\lambda = 400\,\text{nm}$. The reduction factor,

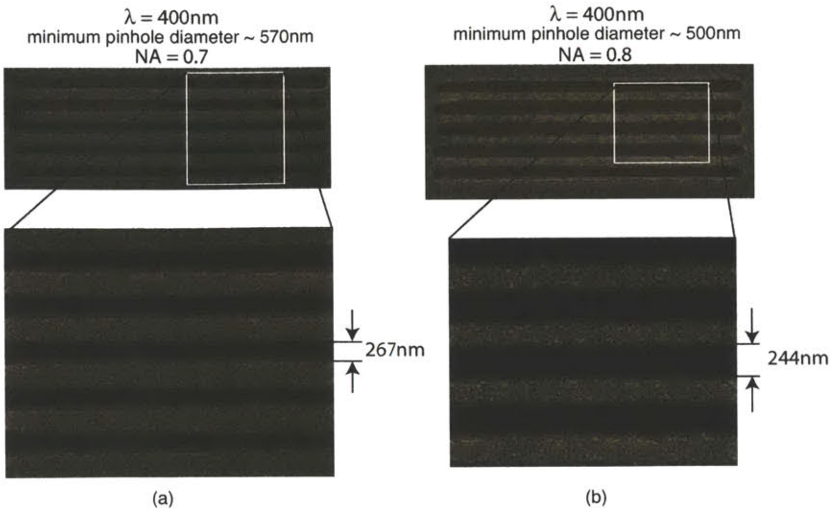


Figure 7-13: *Patterning of dense lines with photon sieves. Scanning-electron micrographs of dense lines patterned with two photon sieves of same K (equal to 2), but different NA. (a) NA=0.7. The exposed lines were about 270 nm, close to half the size of the smallest pinhole on the sieve (570 nm). (b) NA=0.8. The exposed linewidths were 244 nm, again about half the diameter of the smallest pinhole in the sieve (500 nm). The experiments were performed at* $\lambda = 400\,nm$.

K in both cases was 2. The results indicate that photon sieves of NA as high as 0.9 work as expected. Note that these sieves were designed based on scalar diffraction theory. Therefore, we would expect some improvement in their performance, when a full vector theory is used for the design.

In conclusion, we have shown the very first experimental results validating the performance of photon sieves for lithography. We have also shown that a reduction factor upto 2 is possible. This has significant implications in the fabrication and replication of photon sieves compared to zone plates of the same NA. We have also shown results that validate the performance of photon sieves of NA as high as 0.9. Although the efficiency of the photon sieves is smaller than that of phase zone plates, phase-shifting methods offer a means to improve upon this.

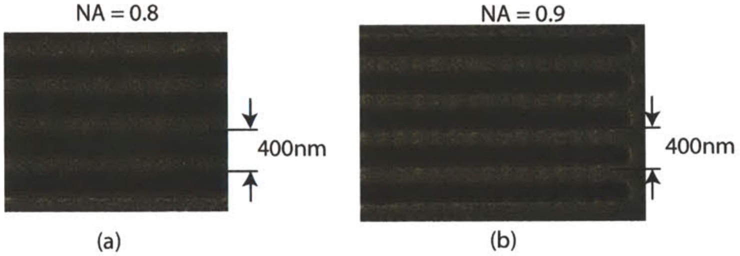


Figure 7-14: *Patterning of dense lines with high-NA photon sieves. Scanning-electron micrographs of dense lines patterned with two photon sieves of same* K *(equal to 2), but different NA. (a) NA=0.8. (b) NA=0.9. In both cases the period of the lines was 400 nm. This is the very first report of focussing by high-NA photon sieves. The experiments were performed at* $\lambda = 400\,nm$.

Appendix A

Scalar Diffraction Theory

Since the full-vector computation of diffracted fields is very time and resource consuming, it is useful to have a simpler model using scalar diffraction theory. For most diffractive lenses of small NA (less than about 0.7), the results are about the same as in the vector case.

A.1 Fresnel-Kirchhoff Diffraction Theory

In this section, we develop a model based on the Fresnel-Kirchhoff diffraction theory.

The field in a plane a distance z away from an aperture is given by

$$f(r,\theta,z) = \frac{e^{ikz}}{2i\lambda} \iint\limits_{-\infty}^{\infty} f_0(\rho,\phi) \frac{e^{iks}}{s} (1+\frac{z}{s}) \rho d\rho d\phi, \tag{A.1}$$

where (r,θ) are the coordinates in the observation plane, (ρ,ϕ) are the coordinates in the plane of the aperture, k is the wave vector and s is given by

$$s = \sqrt{r^2 + \rho^2 - 2r\rho\cos(\theta-\phi) + z^2}. \tag{A.2}$$

$f_0(\rho,\phi)$ is the aperture.

Unfortunately, this equation has to be computed numerically. The presence of the fast varying integrand makes it computationally slow as well. However, it is still faster than the full-vector code especially fine grids are needed. This equation and its numerical implementations are discussed in detail in [16].

In figure A-1, we show the comparison between the intensity computed using the scalar theory and that using the vector theory. This simulation was performed for a zone plate lens with a numerical aperture of 0.7. Note that there is fairly good agreement in the 2 cases.

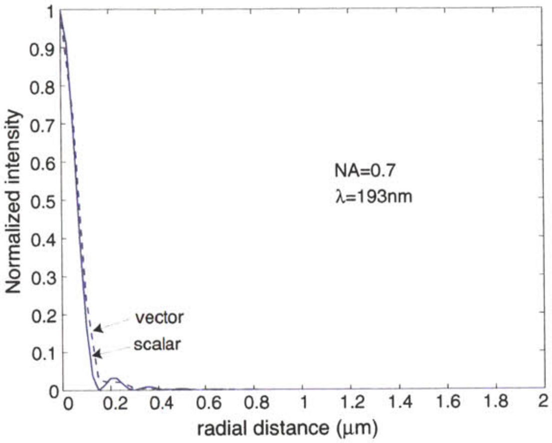


Figure A-1: *The intensity in the focal plane of a zone plate of NA=0.7 and* $\lambda = 193nm$ *is calculated using the FDTD method and the Fresnel-Kirchhoff diffraction integral.*

The same data is plotted in log scale in figure A-2. This shows that the background calculated by the scalar theory is significantly lower than that from the vector code.

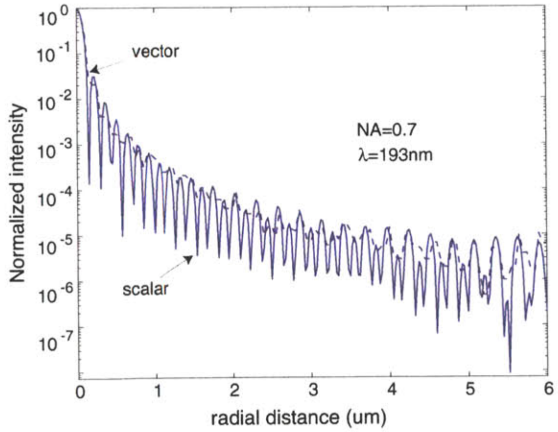


Figure A-2: *The intensity in the focal plane of a zone plate of NA=0.7 and* $\lambda = 193nm$ *is calculated using the FDTD method and the Fresnel-Kirchhoff diffraction integral in log scale.*

Equation (A.1) can be reformulated in cartesian coordinates as

$$f(x,y,z) = \frac{e^{ikz}}{2i\lambda}\iint\limits_{-\infty}^{\infty} f_0(x,y)\frac{e^{iks(x-x',y-y')}}{s(x-x',y-y')}(1+\frac{z}{s(x-x',y-y')})dxdy, \tag{A.3}$$

where

$$s(x-x',y-y') = \sqrt{(x-x')^2+(y-y')^2+z^2}. \tag{A.4}$$

Now, equation (A.3) can further re-written as

$$\begin{aligned} f(x,y,z) &= \frac{e^{ikz}}{2i\lambda}\iint\limits_{-\infty}^{\infty} f_0(x,y)H(x-x',y-y',z)dxdy \\ &= \frac{e^{ikz}}{2i\lambda}f_0 \otimes H, \end{aligned} \tag{A.5}$$

where

$$H(x-x',y-y') = \frac{e^{iks(x-x',y-y')}}{s(x-x',y-y')}(1+\frac{z}{s(x-x',y-y')}), \tag{A.6}$$

and $\otimes$ represents convolution in two dimensions, (x,y).

This makes the numerical computation significantly faster since we can use Fast-Fourier Transform algorithms to compute convolutions more rapidly. After, we implemented the convolution form of the Fresnel-Kirchhof diffraction equation, we have to perform validation simulations. In other words, we compare the results from this software to known results (either analytical or other robust simulation tools such as the FDTD, described earlier).

A.1.1 Validation of numerical Code : Square Aperture

We first start with comparing the results from the numerical code to specific analytical solution. The first case involves diffraction from a rectangular aperture. For a square aperture of side $2w$, the aperture function is

$$f_0(x,y) = \text{rect}(\frac{x}{2w})\text{rect}(\frac{y}{2w}). \tag{A.7}$$

From [14], the diffraction pattern in a plane, z away from the aperture is given by

$$\begin{aligned} f(x,y,z) =& \frac{e^{ikz}}{2i}\{(C(\alpha_2) - C(\alpha_1)) + i(S(\alpha_2) - S(\alpha_1))\} \\ & \{(C(\beta_2) - C(\beta_1)) + i(S(\beta_2) - S(\beta_1))\}, \end{aligned} \tag{A.8}$$

where

$$\begin{aligned} C(z) &= \int_0^z \cos\frac{\pi t^2}{2} dt \\ S(z) &= \int_0^z \sin\frac{\pi t^2}{2} dt, \end{aligned} \tag{A.9}$$

and

$$\begin{aligned} \alpha_1 &= -\sqrt{2}(\sqrt{\frac{\omega^2}{\lambda z}} + \frac{x}{\sqrt{\lambda z}}) \\ \alpha_2 &= \sqrt{2}(\sqrt{\frac{\omega^2}{\lambda z}} - \frac{x}{\sqrt{\lambda z}}) \\ \beta_1 &= -\sqrt{2}(\sqrt{\frac{\omega^2}{\lambda z}} + \frac{y}{\sqrt{\lambda z}}) \\ \beta_2 &= \sqrt{2}(\sqrt{\frac{\omega^2}{\lambda z}} - \frac{y}{\sqrt{\lambda z}}). \end{aligned} \tag{A.10}$$

Figure A-3 shows the intensity along a slice through the center of a square aperture plotted as a function of distance from the optical axis. The results are plotted for both the analytical (solid line) and the numerical (dotted line) case. Note that they agree extremely well with each other.

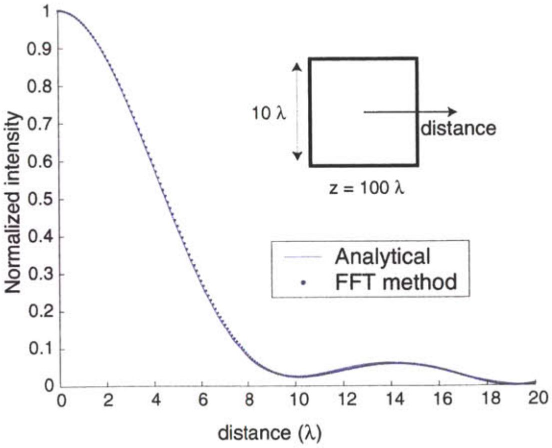


Figure A-3: *Scalar Fresnel-Kirchoff diffraction compared to the vector theory for a square aperture. The normalized intensity in a plane* $z = 100\lambda$ *away from a square aperture of side* $2w = 10\lambda$ *is calculated using the Fresnel-Kirchoff diffraction theory. A slice through the center of the square is taken and the intensity is plotted as a function of the distance from the optical axis. The solid line shows the results using the analytical equations, while the dotted line shows the results using the numerical code.*

A.1.2 Validation of numerical code: Sinusoidal Grating

If the aperture has the form of a one-dimensional sinusoidal grating, we can use the software to predict its self-imaging in the fresnel regime. This is called the Talbot Effect [54]. The grating, with lines in the y-direction is represented as

$$f_0(x,y) = \frac{1}{2}\{1 + \cos\frac{2\pi x}{L}\}, \tag{A.11}$$

where L is the period of the grating.

The diffracted image of this grating at a distance of z is given by

$$f(x,y,z) = \frac{1}{2}\{1 + e^{-i\frac{\pi\lambda z}{L^2}}\cos\frac{2\pi x}{L}\}. \tag{A.12}$$

This predicts that the grating self-images at distances of $z = \frac{2nL^2}{\lambda}$. In addition, the diffracted field repeats the grating with a 180 degree phase shift at distances given by $z = \frac{(2n+1)L^2}{\lambda}$.

In figure A-4, we predict this self-replication of the sinusoidal grating as predicted by the numerical code. This is in agreement with the analytical results.

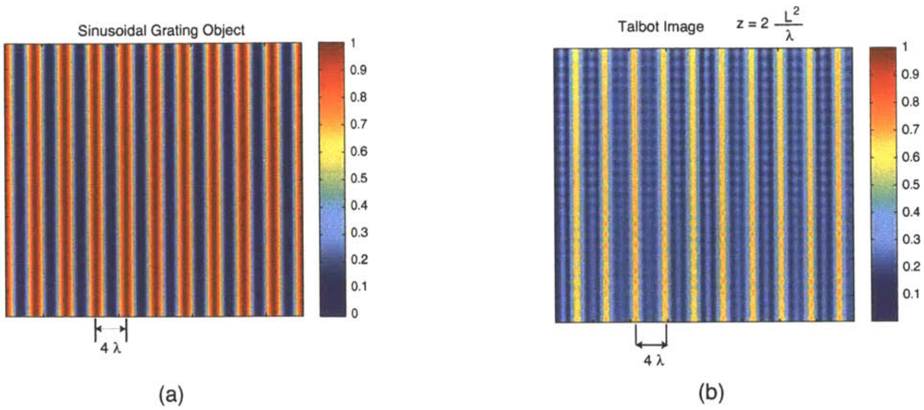


Figure A-4: *Scalar Fresnel-Kirchoff diffraction predicting the Talbot Effect. The normalized intensity in a plane* $z = 2\frac{L^2}{\lambda}$ *away from a sinusoidal grating of period* L *is calculated using the Fresnel-Kirchoff diffraction theory. The code predicts the self-imaging of the grating as expected. (a) is the sinusoidal grating aperture. (b) is its Talbot image.*

A.1.3 Validation of numerical code: Zone Plate

Finally, we compare the diffraction pattern in the focal plane of a zone plate calculated using the numerical Fresnel-kirchoff code to that using the vector FDTD code.

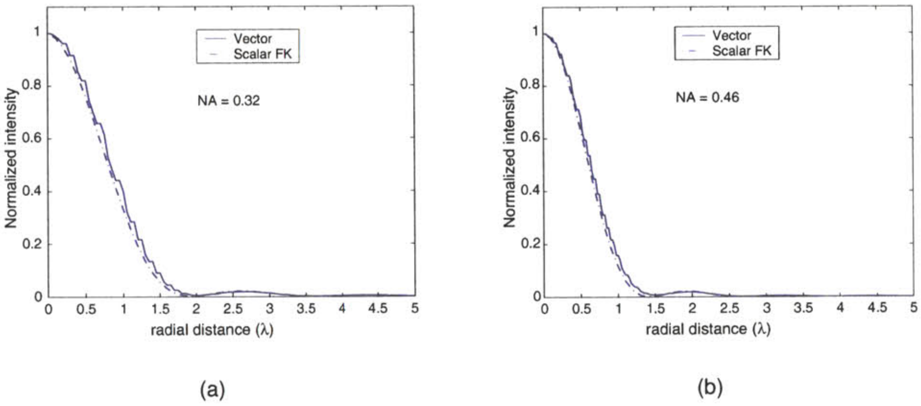


Figure A-5: *Compare scalar Fresnel-Kirchoff diffraction to vector FDTD. The diffraction pattern in the focal plane of a zone plate is plotted as a function of the radial distance. The plots are done with the Fresnel-Kirchoff theory and the FDTD method. They seem to agree quite well. (a) is for a zone plate with NA=0.32 and (b) is for a zone plate with NA = 0.46.*

In figure A-5, we plot the diffracted intensity in the focal plane for zone plates of NA=0.32 (a) and NA=0.46 (b). All distances are normalized to the wavelength used. Note that the vector FDTD method gives results quite close to that from the convolution based Fresnel-Kirchoff code.

A.2 Fresnel Approximation of Fresnel-Kirchhoff Diffraction Equation

Following [55], the scalar diffracted field in a plane perpendicular to the optical axis of the lens is given by

$$f(r,z) = \frac{e^{ikz}}{i\lambda z} e^{ik\frac{r^2}{2z}} \times \Im\{f_0(r') e^{ik\frac{r'^2}{2z}}\}|\rho = \frac{r}{\lambda z}, \tag{A.13}$$

where r is the radial coordinate, z is the axial coordinate, k is the wave number, λ is the wavelength, $\Im$ denotes the Fourier transform (ρ is the spatial frequency coordinate) and f_0 is the diffracting aperture. Note that we assumed circular symmetry.

The fourier transform of a circular symmetric function is given by [56]

$$\Im\{g(r)\} = 2\pi \int_0^\infty r' g(r') J_0(2\pi r' \rho) dr'. \tag{A.14}$$

Therefore, the field at a plane z away from a circular aperture of radius R is given by

$$f_R(r,z) = \frac{e^{ikz}}{i\lambda z} e^{ik\frac{r^2}{2z}} \times 2\pi \int_0^R r' e^{ik\frac{r'^2}{2z}} J_0(2\pi r' \rho) dr'. \tag{A.15}$$

From [57],

$$2 \int_0^1 r e^{-\frac{1}{2} i u r^2} J_0(vr) dr = C(u,v) - iS(u,v), \tag{A.16}$$

where

$$C(u,v) = \begin{cases} \frac{\cos(\frac{u}{2})}{\frac{u}{2}} U_1(u,v) + \frac{\sin(\frac{u}{2})}{\frac{u}{2}} U_2(u,v) & , |\frac{u}{v}| \leq 1 \\ \frac{2}{u}\sin(\frac{v^2}{2u}) + \frac{\sin(\frac{u}{2})}{\frac{u}{2}} V_0(u,v) - \frac{\cos(\frac{u}{2})}{\frac{u}{2}} V_1(u,v) & , |\frac{u}{v}| > 1 \end{cases} \tag{A.17}$$

and

$$S(u,v) = \begin{cases} \frac{\sin(\frac{u}{2})}{\frac{u}{2}} U_1(u,v) - \frac{\cos(\frac{u}{2})}{\frac{u}{2}} U_2(u,v) & , |\frac{u}{v}| \leq 1 \\ \frac{2}{u}\cos(\frac{v^2}{2u}) + \frac{\cos(\frac{u}{2})}{\frac{u}{2}} V_0(u,v) - \frac{\sin(\frac{u}{2})}{\frac{u}{2}} V_1(u,v) & , |\frac{u}{v}| > 1 \end{cases} \tag{A.18}$$

Here U_n and V_n are Lommel functions expressed in terms of rapidly converging series.

$$U_n(u,v) = \sum_{s=0}^{\infty} (-1)^s (\frac{u}{v})^{n+2s} J_{n+2s}(v). \tag{A.19}$$

$$V_n(u,v) = \sum_{s=0}^{\infty} (-1)^s (\frac{v}{u})^{n+2s} J_{n+2s}(v). \tag{A.20}$$

From equations A.15 and A.16, we can write

$$f_R(r,z) = \frac{e^{ikz}}{i\lambda z} e^{ik\frac{r^2}{2z}} \times \pi R^2 (C(u,v) - iS(u,v)), \tag{A.21}$$

where

$$\begin{aligned} u &= -\frac{2\pi}{\lambda}\frac{R^2}{z} \text{ and} \\ v &= 2\pi \frac{rR}{\lambda z}. \end{aligned} \tag{A.22}$$

For a phase zone plate lens with alternate zones having a phase shift of ϕ_0, the field is given by

$$f_{zp}(r,z) = (1 - e^{i\phi_0}) (\sum_{n=1}^{N-1} (-1)^n f_{R_n}(r,z)) + T_N(r,z), \tag{A.23}$$

where

$$T_N = \begin{cases} e^{i\phi_0} f_{R_N}(r,z) & \text{, if N is odd} \\ f_{R_N}(r,z) & \text{, if N is even} \end{cases} \tag{A.24}$$

and R_n are the radii of the zones.

Thus, we have derived an analytical expression for the scalar diffraction from a diffractive lens. In order to understand the scope of this approximation, we compare the results from this approximation to that from the FDTD code. In order to study the validity of this method, we compare the normalized intensity in the focal plane of a zone plate lens calculates using the fresnel approximation and using the FDTD method. This comparison is performed for a range of NA's. In figure A-6, we plot the data for a zone plate lens of NA=0.22. The other parameters used were $\lambda = 193$nm, $f = 40\mu$m and $N = 10$.

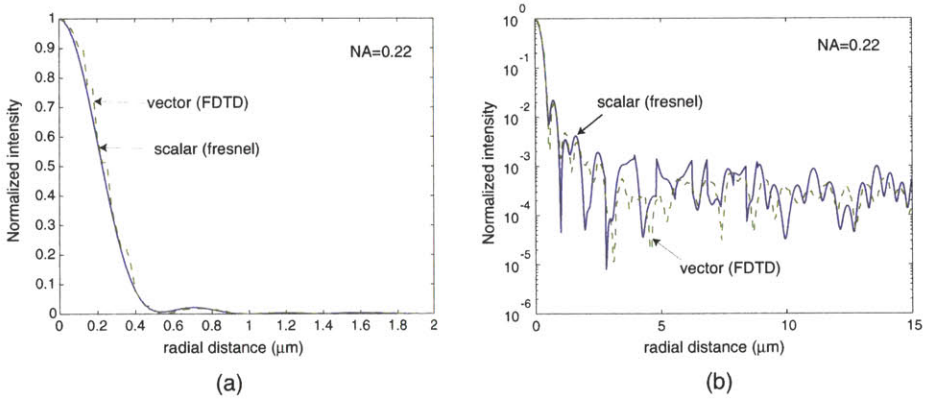


Figure A-6: *Scalar Fresnel approximation compared to the vector theory. The normalized intensity in the focal plane of a zone plate is plotted for the fresnel approximation theory and the full-vector theory in linear (a) and log (b) scales. The calculations were performed for a zone plate with a π phase shift, $\lambda = 193nm$ and $NA = 0.22$.*

Note that in this case, the agreement between the 2 methods is extremely good. The linear plot shows the agreement in detail, whereas the log plot gives a sense for the long range behavior of the data. In order to vary the NA, we only had to change the number of zones.

The parameters in figure A-7 were $\lambda = 193$nm and $NA = 0.7$. Note that the fresnel ap-

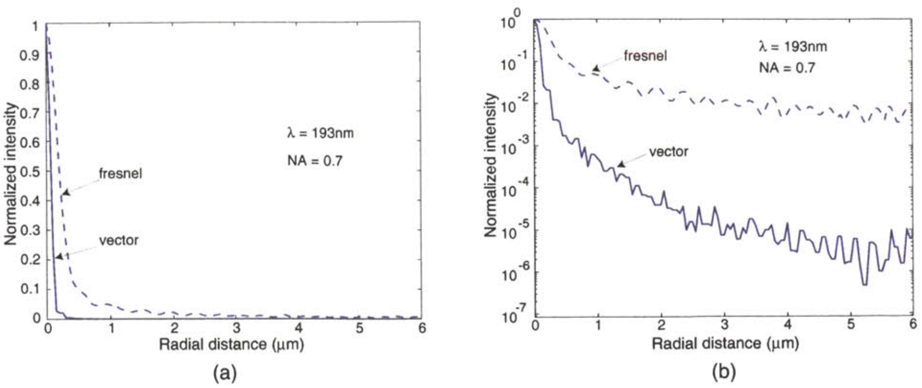


Figure A-7: *Scalar Fresnel approximation compared to the vector theory. The normalized intensity in the focal plane of a zone plate is plotted for the fresnel approximation theory and the full-vector theory in linear (a) and log (b) scales. The calculations were performed for a zone plate with a* π *phase shift,* $\lambda = 193nm$ *and* $NA = 0.7$.

proximation gives an unusually large background. This is expected since the approximation tends to be less accurate for observation points farther from the optical axis. Even though this is a quick, analytical method, it is not suitable for contrast calculation, where large area intensities need to be accurately calculated.

Appendix B

Derivation of Diffraction Efficiency for Zone Plate with general Phase Shift

In this appendix, we derive the diffraction efficiency for a zone plate with a general phase shift. The derivation is based on the assumption of low-NA and long focal lenghts. However, this analysis provides an analytical expression, which gives insight into the numerical results presented in Chapter 4.

We begin with the transmission function of the zone plate, T. Under the assumptions, T is a periodic function in r^2 with a period equal to $2\lambda f$, where λ is the wavelength and f is the focal length of the zone plate. In the subsequent analysis, we use the coordinate transformation

$$u = r^2 \tag{B.1}$$

For a general phase shift, $\Delta\phi$, T is a complex function. We can break this function into its real and imaginary parts in order to use the Fourier series representation.

$$T = T_R + iT_I \tag{B.2}$$

where

$$T_R = \begin{cases} 1 & -\frac{\lambda}{2} \leq u < \frac{\lambda}{2} \\ \cos(\Delta\phi) & \frac{\lambda}{2} < u \leq 3\frac{\lambda}{2} \end{cases} \tag{B.3}$$

and

$$T_I = \begin{cases} 0 & -\frac{\lambda}{2} \leq u < \frac{\lambda}{2} \\ \sin(\Delta\phi) & \frac{\lambda}{2} < u \leq 3\frac{\lambda}{2} \end{cases} \tag{B.4}$$

Using the Fourier series represenation,

$$T_R = \sum_{m=-\infty}^{\infty} c_{Rm} e^{i2m\pi \frac{u}{2\lambda f}} \tag{B.5}$$

and

$$T_I = \sum_{m=-\infty}^{\infty} c_{Im} e^{i2m\pi \frac{u}{2\lambda f}} \tag{B.6}$$

The Fourier-series coefficients are given by

$$c_{Rm} = \begin{cases} \frac{\sin(m\pi/2)}{m\pi}(1 - \cos(\Delta\phi)) & m \text{ odd} \\ 0 & m \text{ even} \\ \frac{1+\cos(\Delta\phi)}{2} & m = 0 \end{cases} \tag{B.7}$$

and

$$c_{Rm} = \begin{cases} -\frac{\sin(m\pi/2)}{m\pi} \sin(\Delta\phi)) & m \text{ odd} \\ 0 & m \text{ even} \\ \frac{\sin(\Delta\phi)}{2} & m = 0 \end{cases} \tag{B.8}$$

The diffraction efficiency of the m^{th} order is given by η_m,

$$\eta_m = c_{Rm}^2 + c_{Im}^2 \tag{B.9}$$

$$\eta_m = \begin{cases} \frac{2}{m^2\pi^2}(1-\cos(\Delta\phi)) & m \text{ odd} \\ 0 & m \text{ even} \\ \frac{1}{2}(1+\cos(\Delta\phi)) & m = 0 \end{cases} \tag{B.10}$$

Thus, we can see that for $\Delta\phi = \pi$, the diffraction efficiency of the 0^{th} order is zero. Substituting $m = 1$, we get

$$\eta_m = \frac{2}{\pi^2}(1-\cos(\Delta\phi)) \tag{B.11}$$

The phase-etch error, Δh can be related to the phase shift by

$$\Delta\phi = \pi + \Delta h(n-1)\frac{2}{\pi} \tag{B.12}$$

Here, we assume π as the desired phase shift. Thus, we can relate the first-order diffraction efficiency to the phase-etch error as follows.

$$\eta_1 = \frac{2}{\pi^2}(1+\cos(\Delta h(n-1)\frac{2\pi}{\lambda})) \tag{B.13}$$

Appendix C

Relation of knife-edge data to Radon Transform

The knife-edge scan is related to the PSF by equation (6.8).

$$\frac{dU}{dx'} = -\int_{-\infty}^{\infty} dy' h^2(x', y') \tag{C.1}$$

where x' is the scan-coordinate, y' is the coordinate in the direction perpendicular to the scan, and h^2 is the PSF.

The radon transform of a function, $f(x,y)$, is given by [58]

$$D_\theta(r) = \iint_{-\infty}^{\infty} f(x - x_0, y - y_0)\delta(x\cos(\theta + \frac{\pi}{2}) + y\sin(\theta + \frac{\pi}{2}))dxdy \tag{C.2}$$

where

$$r = |x\cos\theta + y\sin\theta - r_0| \tag{C.3}$$

The integral is taken along the line

$$(x - x_0)\cos(\theta + \frac{\pi}{2}) + (y - y_0)\sin(\theta + \frac{\pi}{2}) = 0 \tag{C.4}$$

When the scan is along the X-axis, we can set $\theta = \pi/2$ and $y_0 = 0$,

$$D_{\frac{\pi}{2}}(r) = \int_{-\infty}^{\infty} f(-x_0, y)dy \tag{C.5}$$

If we make the assumption

$$f(-x, y) = f(x, y) \tag{C.6}$$

which is true for radially symmetric functions, we get

$$D_{\frac{\pi}{2}}(r) = \int_{-\infty}^{\infty} f(x_0, y)dy \tag{C.7}$$

Comparing equations (C.1) and (C.7), we can see that the left-hand side of equation (C.1) is the radon transform of h^2 where the integral for the transform is carried out along the knife-edge scan axis.

Appendix D

Derivation of the optimum pinhole-radius in a photon sieve

A photon sieve consists of pinholes centered about the zones of an underlying zone plate. The radii of the pinholes must be chosen such that the field at focus is maximised for all the pinholes. Here, we begin with the assumption of far-field, which gives simple analytic expressions, before proceeding to numerical methods. The far-field approximation implies that

$$f >> \frac{2(2a_n)^2}{\lambda} \tag{D.1}$$

where a_n is the radius of the pinhole that is centered on the nth zone. Under this assumption, the field is given by

$$U(\rho_n, a_n) \propto a_n^2 \frac{J_1(\frac{ka_n\rho_n}{f})}{(\frac{ka_n\rho_n}{f})} \tag{D.2}$$

where J_1 is the bessel function of the first order, $k = 2\pi/\lambda$, and the first term, a_n^2 represents the increase in area with increasing a_n. Assuming low-NA and $n >> 1$,

$$\rho_n \approx \sqrt{n\lambda f} \tag{D.3}$$

Under the same assumptions, the zone-width is given by

$$w_n = \sqrt{(n+1)\lambda f} - \sqrt{n\lambda f} \tag{D.4}$$

$$\approx \frac{1}{2}\sqrt{\frac{\lambda f}{n}} \tag{D.5}$$

Subsituting the above expressions into equation (D.2), we get

$$U(\rho_n, a_n) \propto (K)^2 \frac{J_1(\frac{\pi K}{2})}{(\frac{\pi K}{2})} \tag{D.6}$$

where $K = \frac{2a_n}{w_n}$ is the ratio of the diameter of the pinhole to the underlying zone width. The extremum points of this equation can be analytically determined and they are $K = 1.531, 3.515, 5.509, 7.507, \ldots.$

Bibliography

[1] Semiconductor Research Corporation. A brief overview of src/darpa's research initiative in maskless lithography. *Maskless Lithography Workshop*, August 2001.

[2] P. J. Silverman. The intel lithography roadmap. *Intel Technology Journal*, May 2002.

[3] Chiang Yang. Challenges of mask cost and cycle time. *Mask Supply Workshop*, October 2001.

[4] Official website of Micronic Laser Systems. http://www.micronic.se/.

[5] Official website of Power Technology, Inc. http://www.powertechnology.com/.

[6] Official website of the DMD technology from Texas Instruments. http://www.dlp.com/.

[7] T. Bifano et. al. Microelectromechanical deformable mirrors. *IEEE Journal of Selected Topics in Quantum Electronics*, pages 83–89, 1999.

[8] Boston Micromachines Corporation. http://www.bostonmicromachines.com/.

[9] M. J. Li et. al. Fabrication of microshutter arrays for space application. *Proc. of the SPIE*, 4401:295–303, 2001.

[10] Official website of Axsun Technologies, Inc. http://www.axsun.com/.

[11] Official website of Physik Instrumente. http://www.physikinstrumente.de.

[12] Lord Rayleigh. Wave theory. In *Encyclopaedia Brittanica*, volume 24. 1988. Rayleigh's first entry in his notebook is dated 11 April, 1871.

[13] Alan G. Michette, editor. *Optical Systems for soft x rays*, page 217. Plenum Press, 1986.

[14] Joseph W. Goodman. *Introduction to Fourier Optics*, chapter 4, page 84. McGraw-Hill, 1996.

[15] David Attwood. *Soft X-rays and Extreme Ultraviolet Radiation*, chapter 9, pages 347–349. Cambridge University Press, 2000.

[16] Rajesh Menon. Zone plate array lithography (zpal): Simulations and system design. Master's thesis, Massachusetts Institute of Technology, 2000.

[17] K. S. Yee. *IEEE Trans. Antennas and Propagation*, 14:302–307, 1966.

[18] J. P. Beregnger. A perfectly matched layer for the absorption of electromagnetic waves. *J. Comput. Phys.*, 114:185–200, 1994.

[19] Allen Taflove and Susan C. Hagness. *Computational Electrodynamics the finite-difference time-domain method*, chapter 1, pages 18–19. Artech House, 2000.

[20] Allen Taflove and Susan C. Hagness. *Computational Electrodynamics the finite-difference time-domain method.* Artech House, 2000.

[21] http://cuervo.eecs.berkeley.edu/Volcano/ The website of Tempest at the University of California, Berkeley.

[22] Dennis W. Prather and Shouyuan Shi. Formulation and application of the finite-difference time-domain method for the analysis of axially symmetric diffractive optical elements. *J. Opt. Soc. Am. A.*, 16(5):1131–1142, 1999.

[23] T. Rong-Chung et. al. Design, fabrication, and characterization of form-birefringent multiplayer polarizing beam splitter. *J. Opt. Soc. Am. A.*, 14(7):1627–1636, 1997.

[24] S. Quabis, R. Dorn, M. Eberler, O. Glöckl, and G. Leuchs. Focusing light to a tighter spot. *Optics Communications*, 179:1–7, May 2000.

[25] S. Quabis, R. Dorn, O. Glöckl, M. Reichle, and M. Eberler. Reduction of spot size by using a radially polarized laser beam. *Proc. of the SPIE*, 4429:105–111, 2001.

[26] P. Lalanne. Waveguiding in blazed-binary diffractive elements. *J. Opt. Soc. Am. A.*, 16(10):2517–2520, October 1999.

[27] P. Lalanne et. al. Design and fabrication of blazed binary diffractive elements with sampling periods smaller than the structural cutoff. *J. Opt. Soc. Am. A.*, 16(5):1143–1156, May 1999.

[28] Darío Gil. *Maskless nanolithography and imaging with Diffractive Optical Arrays.* PhD thesis, Massachusetts Institute of Technology, 2003.

[29] G. J. Dolan and T. A. Fulton. Canyon lithography. *IEEE Electron device Letters*, 4(6):178–180, 1983.

[30] J. T. Hastings. *Nanometer-Precision Electron-beam lithography with applications in Integrated Optics.* PhD thesis, Massachusetts Institute of Technology, 2003.

[31] C. Grant Willson, Ralph A. Dammel, and Arnost Reiser. Photoresist materials: A historical perspective. *Proc. of the SPIE*, 3049:28–41, 1997.

[32] Katsuji Douki, Toru Kajita, and Tsutomu Shimokawa. High-performance 193-nm positive resist using alternating polymer system of functionalized cyclic olefins / maleic anhydride. White paper from JSR corporation.

[33] P. Rai-Choudhury, editor. *Handbook of Microlithography, Micromachining and Microfabrication.* SPIE Optical Engineering Press, 1997.

[34] Max Born and Emil Wolf. *Principles of Optics*, chapter 8. Cambridge University Press, 1999.

[35] D.J.D Carter et. al. Zone-plate array lithography (zpal): A new maskless approach. *Proc. of the SPIE*, 3676:324–332, 1999.

[36] Alan G. Michette, editor. *Optical Systems for soft x rays*, chapter 9, pages 217–224. Plenum Press, 1986.

[37] Commercial 193nm laser from Cymer. Model ELS-6010A. http://www.cymer.com/products/.

[38] J. T. Hastings, Feng Zhang, and Henry I. Smith. Nanometer-level stitching in raster-scanning e-beam lithography using spatial-phase locking. *J. vac. Sci. Technol. B*, 2003. To be published.

[39] Gilbert Strang. *Introduction to Applied Mathematics*, chapter 8, pages 665–671. Wellesley-Cambridge Press, 1986.

[40] T. H. P. Chang. Proximity effect in electron-beam lithography. *J. Vac. Sci. Technol.*, 12(6):1271–1275, November 1975.

[41] S. Y. Lee and J. Laddha. Application of neural network to enhancing accuracy of e-beam proximity effect correction. *Microelectronic Engineering*, 53:317–320, 2000.

[42] Soo-Young Lee et. al. Proximity effect correction in electron-beam lithography: A heirarchical rule-based scheme-pyramid. *J. Vac. Sci. Technol. B*, 9(6):3048–3053, 1991.

[43] M. Peckerer, R. Bass, and K. W. Rhee. Sub-0.1 μm electron-beam lithography for nanostructure development. *J. Vac. Sci. Technol. B*, 18(6):3143–3149, November 2000.

[44] M. Minsky. Microscopy apparatus. US patent 3,013,467, December 1961.

[45] H. J. Tiziani, R. Achi, R. N. Kramer, and L. Wiegers. Theoretical analysis of confocal microscopy with microlenses. *Applied Optics*, 35(1):120–125, January 1996.

[46] T. Wilson and C. J. R. Sheppard. *Theory and Practise of Scanning Optical Microscopy*, chapter 3. Academic Press, 1984.

[47] T. Wilson and A. R. Carlini. Size of the detector in confocal imaging system. *Optics Letters*, 12(4):227–229, April 1987.

[48] T. Wilson and A. R. Carlini. Three-dimensional imaging in confocal imaging systems with finite sized detectors. *Journal of Microscopy*, 149(1):51–66, January 1988.

[49] Rajesh Menon. An optimum volume-holographic microscope. Area exam report, Massachusetts Institute of Technology, Electrical Engineering and Computer Science, Cambridge, Massachusetts, 2002.

[50] Anthony Yen. *Fabrication of large area 100-nm period gratings using achromatic holographic lithography.* PhD thesis, Massachusetts Institute of Technology, 1991.

[51] A. Ashkin, J. M. Dziedzic, J. E. Bjorkholm, and Steven Chu. Obsevration of a single-beam gradient forcce optical trap for dielectric particles. *Optics Letters*, 11:288–290, 1996.

[52] B. Kress and P. Meyrueis. *Digital Diffractive Optics. An introduction to Planar Diffractive Optics and Related Technology*, chapter 4. John Wiley and Sons, 2000.

[53] L. Kipp, M. Skibowksi, R. L. Johnson, R. Berndt, R. Adelung, and R. Seemann. Sharper images by focusing soft x-rays with photon sieves. *Nature*, 414(184), 2001.

[54] Joseph W. Goodman. *Introduction to Fourier Optics*, chapter 4, pages 87–89. McGraw-Hill, 1996.

[55] Joseph W. Goodman. *Introduction to Fourier Optics*, chapter 4, pages 66–67. McGraw-Hill, 1996.

[56] Joseph W. Goodman. *Introduction to Fourier Optics*, chapter 2, pages 12–16. McGraw-Hill, 1996.

[57] Max Born and Emil Wolf. *Principles of Optics*, chapter 8, pages 435–439. Cambridge University Press, 1999.

[58] J. E. R. Costa, A. V. R. Silva, A. Lüdi, and A. Magun. Beam profile determination by tomography of solar scans. *Astronomy and Astrophysics*, 387:1153–1160, 2002.