Lithography

5

5.1 Introduction

Lithography is the cornerstone of modern IC manufacturing. The ability to print patterns with submicron features and to place those patterns on a silicon substrate with better than 0.1-μm precision is what makes today's chips possible. Virtually all ICs are manufactured today with optical lithography, the basic process introduced in Figure 1–9 in Chapter 1. The concept is simple. A light sensitive photoresist is spun onto the wafer forming a thin layer on the surface. The resist is then selectively exposed by shining light through a mask which contains the pattern information for the particular layer being fabricated. The resist is then developed which completes the pattern transfer from the mask to the wafer. As we saw in the process flow in Chapter 2, the resist may then be used as a mask to etch underlying films or it may be used as a mask for an ion implantation doping step.

While the concept is simple, the actual implementation is very expensive and very complex, primarily because of the demands placed on this process for resolution, exposure field, placement accuracy, throughput, and defect density. Resolution requirements result from the ever-increasing demand for smaller device structures. Exposure field requirements result from ever-increasing chip sizes and the need to expose at least one full chip (preferably more) with each exposure. Placement accuracy is an issue because generally each mask layer needs to be carefully aligned with respect to the existing patterns already on the wafer. Throughput and defect density are of course issues because of the competitive nature of the semiconductor industry. Throughput translates directly into manufacturing cost. Defects translate directly into yield loss and therefore higher costs for the finished chips. Defects introduced during the lithographic process are a significant contributor to final chip yields.

The SIA NTRS defines the needs for future generations of silicon technologies. A portion of this roadmap relating to lithography is shown in Table 5–1 [5.1]. There are a number of interesting observations that can be made from this chart. First, as we saw in Chapter 1, the principal driving force behind this roadmap is the steady reduction in device dimensions or minimum feature size. The feature size reduction corresponds to a factor of 0.7X in linear dimension or a decrease in the area required per transistor by 50% approximately every three years. It is also interesting to note in the table that the most commonly cited minimum feature sizes in the NTRS (the bold entries in the table corresponding to the dense lines and spaces in DRAM chips) are actually not as small as the isolated lines (MOS gates) required on microprocessor (MPU) chips. These features are generally 20% – 30% smaller than the usually cited "minimum feature size."

Table 5–1 Lithography requirements for future generations of silicon technology [5.1]

Year of first DRAM Shipment	1997	1999	2003	2006	2009	2012
DRAM Bits/Chip	256M	1G	4G	16G	64G	256G
Minimum Feature Size nm						
Isolated Lines (MPU)	200	140	100	70	50	35
Dense Lines (DRAM)	**250**	**180**	**130**	**100**	**70**	**50**
Contacts	280	200	140	110	80	60
Gate CD Control 3σ (nm)	20	14	10	7	5	4
Alignment (mean + 3σ) (nm)	85	65	45	35	25	20
Depth of Focus (μm)	0.8	0.7	0.6	0.5	0.5	0.5
Defect Density (per layer/m^2)	100	80	60	50	40	30
@ Defect Size (nm)	@ 80	@ 60	@ 40	@ 30	@ 20	@ 15
DRAM Chip Size (mm^2)	280	400	560	790	1120	1580
MPU Chip Size (mm^2)	300	360	430	520	620	750
Field Size (mm)	22x22	25x32	25x36	25x40	25x44	25x52
Exposure Technology	248 nm DUV	248 nm DUV	248 nm or 193 nm DUV	193 nm DUV or ?	193 nm DUV or ?	?
Minimum Mask Count	22	22/24	24	24/26	26/28	28

Minimum features must not only decrease in average size with each technology generation, but the variation of these feature sizes must also decrease with time. Generally the CD (Critical Dimension) control is required to be about 10% of the smallest feature size. This requirement is usually expressed in terms of the 3σ control (three standard deviations of the feature size population need to be within the specified 10% of the mean).

The corresponding placement or alignment accuracy for these features (overlay) remains at about one-third of the minimum feature size for each generation of technology. The size of chips is projected in the NTRS to increase significantly over time. The corresponding lithographic printing area (field size) therefore also needs to increase in order to be able to print at least one chip with each exposed field. Typically about one-third of the total wafer manufacturing costs are consumed by lithography. Thus with a wafer manufacturing cost of about $1000 for an 8″ wafer, only a few hundred dollars can be spent on lithography. Wafer exposure tools today cost about $10 million and thus must be capable of printing on the order of 50 wafers per hour in order to meet these cost objectives.

Optical lithography will certainly be used through the 0.18- and 0.13-μm generations, and it is likely that it will be extendible to the 0.1-μm generation as well. Beyond that point, there is considerable uncertainty, for reasons that we will discuss in detail in this chapter. Potential candidates to replace optical lithography include X-ray, e-beam direct write, projection e-beam, and Extreme Ultraviolet (EUV), each of which is being explored in a variety of approaches. The last two are currently the most likely "winners,"

but there is still considerable debate as to which (if any) of these techniques will prove to be successful in manufacturing beyond 0.1 μm. This uncertainty is perhaps the greatest single challenge facing the semiconductor industry today in terms of continuing the historic trends represented by the NTRS.

In this chapter we will explore two major areas associated with the lithographic process. The first of these is the exposure system used to print the patterns on the wafer. We will focus on optical exposure systems because these dominate the industry today. Typical systems today use "step and repeat" or "step and scan" reducing projection printing with complex lens systems. The second major area we will explore is the photoresist material itself. Here the primary issues are sensitivity, resolution, and ruggedness. Typical resists today are carbon-based organic materials that have been engineered for optimum performance in IC manufacturing.

5.2 Historical Development and Basic Concepts

The overall lithography process is conceptually illustrated in Figure 5–1. The patterns that comprise the various layers in an integrated circuit are designed using Computer-Aided Design (CAD) systems. Today these systems contain many advanced capabilities that greatly improve the efficiency of designing chips with many millions of components. Libraries of previous designs that are known to work are usually available, from which basic functions or circuits can be cut and pasted into new designs. Software tools are used to help route or wire the connections between functional blocks. Additional tools check the design to make sure that there are no violations of design rules. And finally, circuit- and system-level simulation tools are available to predict the performance of the new design.

Once the design is complete and ready to transfer to the fabrication facility, the information for each mask level is transferred to a mask making machine that is either an

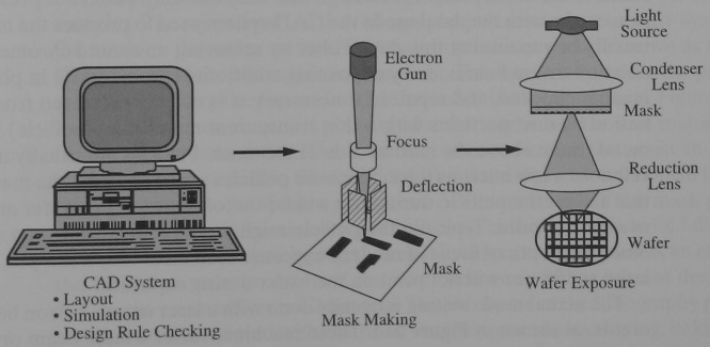


Figure 5–1 Lithography process from mask design to wafer printing.

electron beam or laser pattern generator. The pattern for each mask is literally written on a mask blank using a scanning electron or laser beam. The mask itself is usually a fused silica plate covered with a thin layer ($\approx$ 80 nm) of chromium and a layer of photoresist. A thin antireflection coating (ARC) layer (10 – 15 nm) is also often used between the chrome and the resist to prevent reflections from the chrome layer which can degrade pattern resolution. The electron or laser beam exposes the resist which is then developed and used as an etch mask to transfer the mask pattern into the chrome. The chrome layer is usually wet etched because the process is simple although dry etching may be more widely used in the future as dimensions continue to decrease. Tight dimensional control can be maintained because the chrome layer is so thin. Once the chrome is etched, the photoresist is removed. The fused silica substrate has highly polished surfaces so that light is not scattered as it passes through the mask and ideally has a small thermal expansion coefficient so that mask dimensions are stable over small temperature changes. It is also obviously important that the clear areas of the mask where the chrome has been etched away have high transparency at the wavelength of light used in the wafer exposure system. This is becoming more difficult to achieve in practice as the exposing wavelength moves deeper into the UV region to achieve higher resolution. Even though the mask material (basically SiO_2) has a large bandgap ($\approx$ 9 eV), trace impurities in the glass can degrade optical transparency because they can absorb light at these short wavelengths.

Usually the mask is fabricated with pattern dimensions 4X to 5X larger than the features actually desired on the wafer because the wafer exposure system reduces the image by the same factor. Demagnification in the wafer exposure system makes the mask easier to fabricate and easier to check for imperfections. This is a crucial step because any defects on the mask will be directly imaged on every wafer exposed with the mask and contribute to yield loss. Since step and repeat exposure systems typically expose only a one- or two-chip area at a time, a mask defect which prints on the wafer can easily make the yield zero. Mask inspection is performed by comparing the mask chrome pattern with another identical pattern when there are two or more chip patterns on the mask. If the mask pattern is so large that only one chip pattern is present, the comparison can be with the database in the CAD system used to produce the mask. Defects can usually be repaired at this stage either by removing unwanted chrome areas with laser ablation or ion beams, or by depositing additional chrome to fill in pinholes. Once the mask is checked, and repaired if necessary, it is usually protected from later contamination by dust particles with a thin transparent membrane (pellicle) stretched over a metal frame above the chrome side of the mask. Pellicles are usually made from nitrocellulose a few microns thick. Since the pellicles are offset from the mask, dust particles that fall on the pellicle during use will be out of focus on the wafer and therefore will usually not print. Typically the pellicle might be mounted a few mm above the mask. Since the depth of focus of modern exposure systems is only a few μm, dust particles as large as 100 μm will not print on the wafer during exposure.

The actual mask writing is usually done with a laser or an electron beam pattern generator, as shown in Figure 5–1. These machines raster a laser beam or a beam of electrons in an X-Y pattern across the mask with the beam blanked on and off as necessary to generate the appropriate mask pattern. Generally the beam size is on the order of 0.1 – 0.5 μm for e-beam systems and slightly larger for laser-based systems, which is ade-

quate for writing 4X or 5X masks for today's steppers. As minimum feature sizes continue to decrease, the beam size will have to shrink as well in order to produce masks with adequate resolution. It is perhaps obvious that the e-beam system could be used directly to write images on a wafer simply by putting the electron sensitive resist on the wafer rather than on the mask. The reason this is not done in high-volume manufacturing is simply because the wafer throughput would be too slow. Typical e-beam pattern generators would require tens of minutes to expose a full wafer. This is much slower than optical steppers that typically have throughputs on the order of 50 wafers per hour. However, as device dimensions continue to shrink, e-beam approaches to wafer exposure at least offer the required resolution, if means could be found to improve throughput. We will return to this issue later in the chapter.

The pattern information is transferred to the wafer by printing the mask pattern in a layer of photoresist on the wafer surface. Generally today this is done with a projection exposure system as also illustrated in Figure 5–1. Light from a high-intensity source is collimated and passed through the mask. Each clear area on the mask transmits the light which is then collected and focused by a second lens system. This second lens system also serves to reduce the image size by 4X – 5X. The field of view of such systems is typically only a few cm on a side, so that only a few chips are printed during each exposure. The wafer is physically moved ("stepped") to the next exposure field and the process repeated so these systems are commonly called step and repeat tools or simply steppers. In some systems ("scanners") a narrow slit of light is produced by the light source system and the mask and wafer are simultaneously scanned mechanically so that the mask image is scanned across the wafer. Note that in systems of this type, the mechanical system must scan the mask 4X or 5X times faster than the wafer, corresponding to the optical reduction factor. Scanners greatly increase the field size up to the limit of the mask dimension, at least in the scanning direction.

It is convenient to separate the lithography process into three parts. The first, which is the most straightforward, is the light source used to generate the photons which ultimately expose the resist. While this might seem to be a simple task, higher-resolution lithography demands shorter wavelength photons and this makes the source issue more complex. The second part of the lithography process is the exposure system that images the mask on the wafer surface. This system produces what is known as the aerial image (Figure 5–2) at the top surface of the resist. This image is basically the pattern of optical radiation striking the top surface of the resist. In the example in Figure 5–2, a positive resist is illustrated since this polarity of resist represents the vast majority of manufacturing applications today. The incoming photons strike the resist in the light areas, changing its properties in those regions. The three-dimensional optical intensity pattern through the photoresist layer produces a 3D latent image of the mask pattern that can then be developed. The exposed positive resist in this example would dissolve in the resist developer. After developing, the resist in this example is used as an etch mask so that the TiN local interconnect level can be etched away where it is not desired (see Figure 2–37 in the process flow in Chapter 2). The third part of the lithography process contains all the issues associated with the resist itself: exposure, developing, baking, and so on.

The aerial image thus serves as the dividing line between the major parts of the lithography system (the exposure tool and the resist). The job of the exposure tool is to produce the best aerial image possible. Best is defined in terms of resolution, exposure

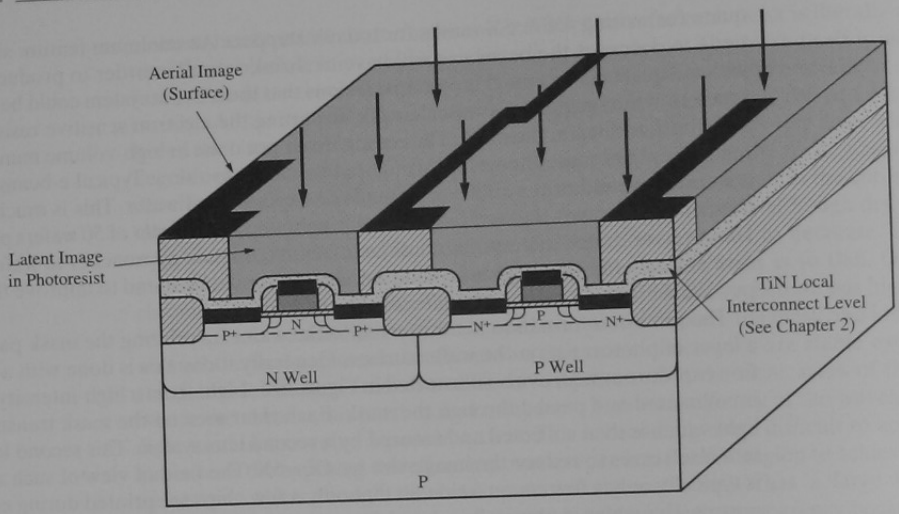


Figure 5–2 Separation of the lithography process into exposure and wafer features. The dividing line is the aerial image of the mask, which is the pattern of radiation that strikes the surface of the photoresist. This example corresponds to Figures 2–36 and 2–37 in Chapter 2.

field, depth of focus, uniformity, and lack of aberrations across the field, photon intensity, and so forth. The job of the photoresist is to translate this aerial image into the best thin film 3D replica of the aerial image possible. Here best is defined in terms of geometric accuracy, exposure speed, and resist resistance to subsequent processing (ruggedness). In the following sections we will consider many of the basic concepts associated with exposure tools and resists and also briefly consider the issues associated with the light source itself.

5.2.1 Light Sources

It is probably obvious to most readers that higher-resolution lithography requires shorter wavelength photons. Modern exposure systems produce diffraction limited images and diffraction effects are strongly related to the wavelength of the exposing radiation. We will discuss this issue in more detail shortly, but in this section we consider first the issues associated with generating the light in modern lithography systems. Historically, most lithography systems have used arc lamps as the primary light source. These lamps usually contain Hg vapor inside a sealed glass envelope. Two conducting electrodes inside the envelope are separated by several mm. An arc is struck between the electrodes by applying a high enough voltage to ionize the gas (typically several kV). Once the gas is ionized, it behaves like a plasma, conceptually similar to the plasmas used in the deposition and etching systems described in Chapters 9 and 10. In the arc

lamp, the plasma is conducting and consists of ions, electrons, and neutral species. Typical lamps used for photolithography consume about one kW of power. At room temperature, the Hg vapor pressure is on the order of 1 atm, but once the lamp is operating, power dissipation quickly raises the temperature and therefore the internal pressure to $\approx 20 - 40$ atm.

Light emission occurs through two processes. The free electrons in the plasma have an effective temperature on the order of 40,000K and emit black body radiation according to Planck's law. The wavelength corresponding to this temperature is very deep in the ultraviolet and is above the bandgap of the fused silica used to form the glass envelope of the lamp. Thus this radiation is mostly absorbed before it exits the lamp.

The second source of emitted light comes from the Hg atoms themselves. Collisions in the gas between the free electrons and the Hg atoms provide energy to some of the electrons in the Hg atoms, raising them to higher energy levels. When these electrons drop back to their lower energy states, they radiate photons at specific energies (frequencies) characteristic of the allowed energy levels in the Hg atom. This process is very much like that shown in Figure 4–18 in Chapter 4, where we discussed the use of these energy transitions in identifying the specific elements from which photons are emitted. In the case of Hg, there is strong emission at a number of UV wavelengths. Most steppers use a single wavelength by simply filtering out the unwanted emissions. The complex optical systems in these exposure tools are much easier to design if they need to focus only a single wavelength. Two that are commonly used are 436 nm (g-line) and 365 nm (i-line). Most exposure tools in the early 1990s used the g-line. However, as linewidths shrunk, use of the i-line became dominant because of the better resolution it provides. Some exposure tools, particularly simple contact and proximity printers, sometimes use broader wavelength light (several lines from the Hg source) for exposure since there are no focusing optics in these systems.

i-line steppers dominated production for the 0.35-μm generation of technology. Generations beyond 0.35 μm require shorter exposure wavelengths and new light sources. The brightest sources in the deep UV part of the spectrum are excimer lasers. Two that are of specific interest for lithography are KrF (248 nm) and ArF (193 nm). In excimer lasers two elements are present which do not normally react in their unexcited state (often a noble gas and a halogen containing compound). However if these elements (Kr and NF_3, for example) are excited, a chemical reaction forming, for example, KrF, is possible. When the excited molecule returns to its ground state, a photon is emitted in the deep UV and the molecule breaks up. Energy must be continually provided, therefore, usually from an internal pulsed discharge, to replenish the population of the excited species. Usually the laser source is strobed at a frequency of several hundred Hz and generally about 100 flashes are used at each exposure site on the wafer to minimize speckle noise. A total energy of several hundred mJ is provided to each exposure site. In steppers, this energy is focused on a wafer area of a few cm^2, providing reasonable exposure times. A number of problems exist with these exposure sources, including reliability and longevity of the laser, transparency of optical components in the lens systems at these wavelengths, and finding suitable resists. These problems have been solved at 248 nm (KrF) and this source is now being used in commercial production for the 0.25-μm and 0.18-μm generations. ArF is the most likely source for the 0.13-μm and 0.10-μm technology generations. The picture is much less clear beyond the 0.1-μm generation.

5.2.2 Wafer Exposure Systems

We now consider in more detail one of the critical elements in modern lithography, the system used to create the aerial image at the photoresist surface. There are three general classes of optical wafer exposure tools, contact, proximity, and projection systems, although only the last is in widespread high-volume manufacturing use today. These are conceptually illustrated in Figure 5–3.

Contact printing is the oldest and the simplest printing process. The mask is placed chrome side down in direct contact with the resist layer on the wafer. The exposure of the resist then takes place by shining light through the mask. The aligning of the mask to patterns already on the wafer takes place prior to exposure, by observing both the mask and wafer patterns through a microscope with the mask slightly separated from the wafer. Contact printing systems are actually capable of high-resolution printing because with the mask and wafer in contact, diffraction effects are minimized. In addition, the machines to perform contact printing are relatively inexpensive. However these types of systems cannot be used in high-volume manufacturing of complex chips for one very simple reason. The hard contact between the mask and the wafer results in damage to both the mask and the resist layer and therefore results in high defect densities. The resulting chip yields are not compatible with economic manufacturing of today's chips. However, contact printing systems are used in some manufacturing applications where low volumes or small chip sizes make the economics of these systems more attractive.

Proximity printing largely solves the defect issues associated with contact printing, because the mask and wafer are kept separated by 5 – 25 μm. However, these systems are also not suitable for manufacturing most of today's chips because the separation of the mask and wafer degrades the resolution of the printed patterns due to diffraction effects. We will quantify these effects shortly, but in practice it is not possible to print features smaller than a few μm with UV exposure and gaps on the order of 20 μm. Interestingly, the resolution of these systems improves as the exposure wavelength de-

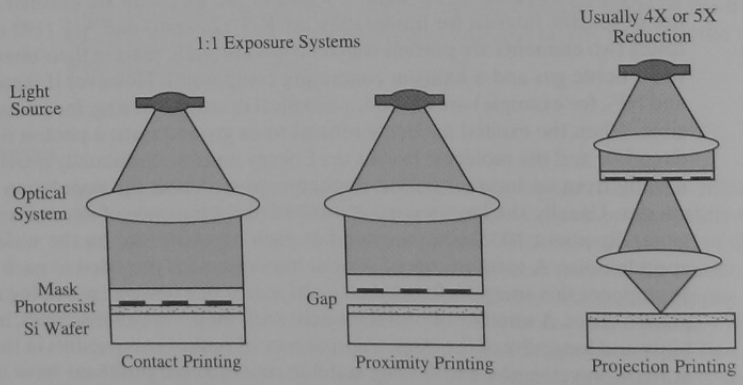


Figure 5–3 Three basic methods of wafer exposure.

creases and X-ray lithography systems can use proximity printing and achieve high resolution because of the very short exposure wavelength (1 – 2 nm). We will return to X-ray systems at the end of this chapter. Both contact and proximity systems require 1X masks, which are more difficult to produce than masks for reduction systems.

The dominant method of wafer exposure today is projection printing. These systems provide high resolution but without the defect problems of contact printing. In projection exposure tools, the mask is physically separated from the wafer and an optical system is used to image the mask on the wafer. This obviously solves the defect issues associated with contact printing. The resolution of projection printers is generally limited by diffraction effects that we will describe in detail. Generally the optical system reduces the mask image by 4X to 5X which means that only a small portion of the wafer is printed during each exposure. Typically such steppers today are capable of printing 0.25-μm or smaller features over an exposure field of several cm^2 and have a throughput of 25 – 50 wafers per hour. They also cost many millions of dollars.

5.2.2.1 Optics Basics—Ray Tracing and Diffraction

In order to understand and quantify the capabilities of modern wafer exposure systems we will need to review some basic concepts about light and optical systems. Many standard references exist on these topics (see, for example, [5.2]) and the reader is referred to these texts for a more detailed treatment of the concepts described here. The application of basic optical principles to lithography systems is also described in detail in a number of books, for example [5.3, 5.4]. In this section we will introduce the basic ideas in a somewhat qualitative way. In Section 5.5, we will be more quantitative.

Light travels as an electromagnetic wave through space. When one is interested in the behavior of an optical system in which all of the dimensions are very large compared to the wavelength of light, the light can usually be treated as particles traveling in straight lines between the optical components. This simplifies the problem to one of "ray tracing." In projection lithography systems, for example, this approach can often be used in describing the optical source and condenser lens, but it fails when the light passes through the mask because the feature dimensions on the mask are comparable to the wavelength of the light. To understand the behavior of the light as it passes through the mask, the objective lens, and on to the wafer, we must include the wave nature of light. The single most important effect that we must account for is diffraction.

Diffraction effects occur because light does not in fact travel in straight lines. Many simple experiments can demonstrate this fact. Figure 5–4 illustrates the passage of light through a small aperture. The light pattern which strikes the screen (image plane) covers a much larger area than can be accounted for by simply drawing straight lines (ray tracing) to describe the light propagation. In fact the smaller the aperture becomes, the more spread out the screen image becomes. This is very much like the situation we find in modern lithography where light passes through a mask with apertures (clear regions) with dimensions on the order of the light wavelength. (We actually illustrated these effects in Figures 5–1 and 5–3 by showing the light spreading out after passing through the mask, but did not explain why this happens in connection with those figures.)

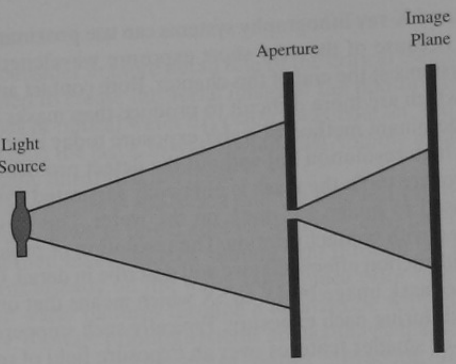


Figure 5–4 Simple example of diffraction effects. Light passes through a narrow aperture. The image formed covers a much larger area than can be explained based on simple straight line ray tracing.

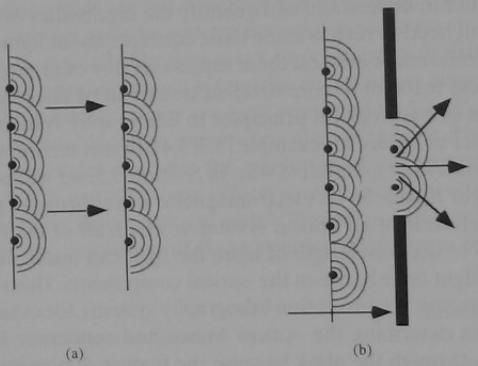


Figure 5–5 Propagation of a plane wave in (a) free space and (b) through a small aperture, illustrating the use of the Huygens-Fresnel principle to construct the wavefront as it propagates.

Consider Figure 5–5. A plane wave is shown propagating through space, unobstructed in (a) and passing through an aperture in (b). The Huygens-Fresnel principle can be used to construct the wavefront versus position as it propagates. This principle states that every unobstructed point of a wavefront at a given instant in time acts as a source of a spherical secondary wavelet of the same frequency as the primary wave. The amplitude of the optical field can be found by superimposing all these wavelets, considering their relative magnitudes and phases. In (a) this superposition simply results in a propagating plane wave. In (b), only the source points in the aperture serve as sources

of Huygens wavelets and the resulting propagating pattern beyond the aperture involves diffraction. As the arrows in (b) illustrate, the light spreads out beyond the aperture. In fact the smaller the aperture becomes, the more the light spreads out because there are fewer wavelets able to pass through the aperture and thus more spherical propagation results beyond the aperture.

Diffraction can be thought of simply as the "bending" of light when it passes through an aperture. The light that passes through the aperture carries with it the information on the size and shape of that aperture. If, for example, the aperture were part of a mask that we wished to print on a wafer, then the information about the aperture size and shape needs to be carried by the light to the photoresist on the wafer. The problem is that this information spreads out in space because of diffraction and it must all be collected to convey perfect information about the aperture to the resist on the wafer. Figure 5–6 illustrates this in a qualitative way. Because of its finite size, the focusing lens collects only part of the total diffraction pattern associated with the light passing through the aperture. The light diffracted to wider angles carries the information about the finer details of the aperture, so it is those details that are lost first when a lens of finite size is used to collect and focus the light.

The actual image produced in this simple example is shown in Figure 5–7 for a small circular aperture and is known as Airy's disk after Sir George Biddell Airy who first derived the expression describing the central intensity maximum. The central maximum approximates the image of the circular aperture. Because of diffraction effects, the image is composed of a bright center disk surrounded by a series of faint rings. The image intensity can be described mathematically by Bessel functions, and the approximate size of the image is given by

$$\text{Radius of central maximum} = \frac{1.22\lambda f}{d} \tag{5.1}$$

where d is the focusing lens diameter, f is the focal length, and λ is the wavelength of the light. Note that a point source only produces a point image if $d \to \infty$ (or if λ or $f \to 0$).

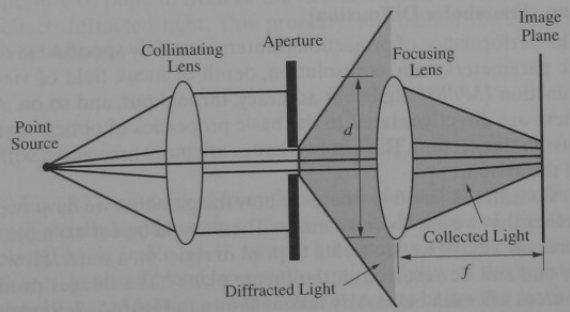


Figure 5–6 Qualitative example of a small aperture being imaged.

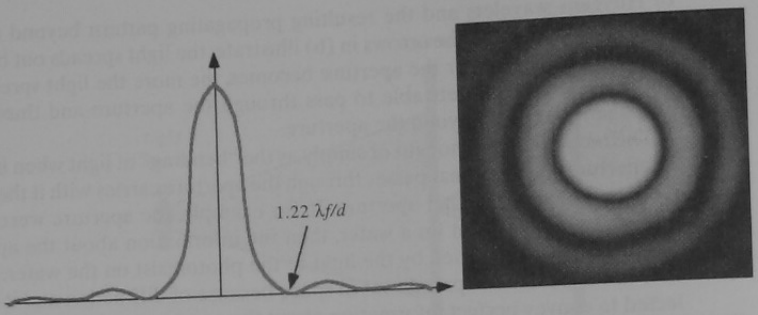


Figure 5–7 Image intensity of a circular aperture in the image plane (Fraunhofer diffraction pattern). The intensity is sketched along any diameter on the left. The pattern on the right illustrates the 2D image. Photo courtesy of J. Goodman. Reprinted with permission of McGraw-Hill [5.2].

Two types of diffraction are usually distinguished: Fresnel or near field diffraction and Fraunhofer or far field diffraction. Both are fundamentally due to the same effect—the wave nature of light. In Fresnel diffraction, the image plane is close to the aperture. The light travels directly from the aperture to the plane where the image is formed, as in Figure 5–4 with no intervening lens system. In Fraunhofer diffraction, the image is far from the aperture and a lens is normally placed between the aperture and the image plane to capture and focus the image, as in Figure 5–6. In modern lithography systems, Fresnel diffraction applies to contact and proximity exposure systems and Fraunhofer diffraction applies to projection systems as should be apparent by considering Figure 5–3. Mathematical descriptions and detailed models have been developed for both of these regimes. Powerful simulation tools based on these models have also been developed, which allow the calculation of the aerial image formed by the wafer exposure system. We will discuss these issues in more detail in Section 5.5.

5.2.2.2 Projection Systems (Fraunhofer Diffraction)

The performance of projection printers is usually specified in terms of a number of basic parameters such as resolution, depth of focus, field of view, Modulation Transfer Function (*MTF*), alignment accuracy, throughput, and so on. At least the first four of these are directly related to the basic properties of optical systems, which we will discuss in this section. The last two issues are more associated with the mechanical design of the system.

Consider Figure 5–8 where we now imagine that we have two point sources close together that we are trying to image. These could be, for example, two small adjacent features on a mask we are trying to print in resist on a wafer. How close together can they be and still be resolved in the image plane? The images produced by the two point sources will each be an Airy disk as shown in Figure 5–7. Rayleigh suggested that a reasonable criterion for resolution was that the central maximums of each point image lie at the first minima of the adjacent point image, the distance given by Eq. (5.1). While

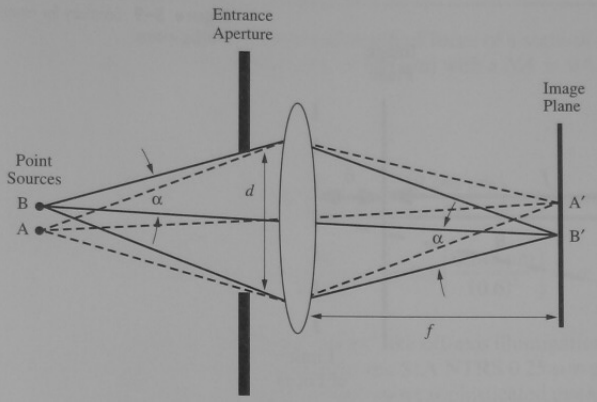


Figure 5–8 Illustration of the resolving power of a lens when two point sources are to be separated in the image.

this definition is somewhat arbitrary, it is useful and has been widely adopted. With this definition, the resolution R of the lens is given by [5.5]

$$R = \frac{1.22\lambda f}{d} = \frac{1.22\lambda f}{n(2f\sin\alpha)} = \frac{0.61\lambda}{n\sin\alpha} \tag{5.2}$$

where n has been included for generality and is the index of refraction of the material between the object and the lens (usually air with $n = 1$ in lithography systems). α is the maximum half-angle of the diffracted light that can enter the lens or the acceptance angle of the lens. α may be limited by the physical size of the lens itself, or by an entrance aperture or pupil in front of the lens. α is really a measure of the ability of the lens to collect diffracted light. This property was named the Numerical Aperture (NA) by Ernst Abbe.

$$NA \equiv n\sin\alpha \tag{5.3}$$

$$\therefore\ R = \frac{0.61\lambda}{NA} = k_1\frac{\lambda}{NA} \tag{5.4}$$

Eq. (5.2) was derived from the Fraunhofer diffraction pattern for an Airy disk and as such strictly applies only to point sources. Because of this the 0.61 factor is often replaced by k_1 in Eq. (5.4). In real lithography systems the mask contains a variety of shapes. k_1 also depends in practice on the ability of the resist chemistry to distinguish closely spaced features on the wafer structure below the resist (topography, reflectivity, etc.) and on defocusing at the image plane. Actual k_1 values achieved in practical optical lithography systems are 0.6 to 0.8.

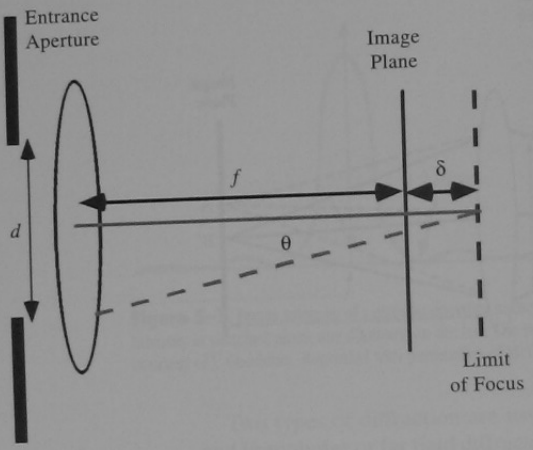


Figure 5–9 Geometry for estimating the depth of focus of an imaging system.

It is obvious from Eq. (5.4) that shorter exposure wavelengths lead to better image resolution. It is also clear that lenses with higher numerical apertures also achieve better resolution, basically because they are able to capture more of the diffracted light and therefore construct a better image. Aside from the difficulty of building larger (higher *NA*) lenses, there is also a significant drawback to using higher *NA* lenses. This is the depth of focus of the lens, which can be estimated as shown in Figure 5–9.

If δ is the on-axis path length difference at the limit of focus, then the path length difference for a ray from the edge of the entrance aperture is simply $\delta\cos\theta$. The Rayleigh criteria for depth of focus is simply that these two lengths not differ by more than $\lambda/4$. Thus

$$\lambda/4 = \delta - \delta\cos\Theta \tag{5.5}$$

Assuming θ is small,

$$\lambda/4 = \delta\left[1 - \left(1 - \frac{\Theta^2}{2}\right)\right] \equiv \delta\frac{\Theta^2}{2} \tag{5.6}$$

$$\Theta \equiv \sin\Theta = \frac{d}{2f} = NA \tag{5.7}$$

$$\therefore\ DOF = \delta = \pm\frac{\lambda}{2(NA)^2} = \pm k_2\frac{\lambda}{(NA)^2} \tag{5.8}$$

The factor of $^1/_2$ is often replaced by k_2 in Eq. (5.8) because the $^1/_2$ value is appropriate at the Rayleigh resolution limit but does not take into account the increase in depth of focus for larger features nor the dependence in practice on other parameters like the resist process.

Example Estimate the resolution and depth of focus of a state-of-the-art excimer laser stepper using a KrF light source ($\lambda = 248$ nm) with a $NA = 0.6$. Assume $k_1 = 0.75$ and $k_2 = 0.5$.

Answer

$$R = k_1 \frac{\lambda}{NA} = 0.75 \left(\frac{0.248\ \mu m}{0.6} \right) = 0.31\ \mu m$$

$$DOF = \pm k_2 \frac{\lambda}{(NA)^2} = \pm 0.5 \left[\frac{0.248\ \mu m}{(0.6)^2} \right] = \pm 0.34\ \mu m$$

Using additional technical "tricks" like off-axis illumination, the resolution can be pushed below 0.25 μm, suitable for the SIA NTRS 0.25-μm generation. Further improvements can be obtained through more sophisticated mask designs using concepts like optical proximity correction and phase shift masks, which we will describe later. The depth of focus is on the same order as the resist layer thickness itself and therefore requires very flat topography and careful attention in the stepper to keeping the image plane focused by adjusting the height of the wafer with respect to the lens.

There is an additional basic concept regarding optical exposure systems that will be useful to us. This is the Modulation Transfer Function or *MTF* illustrated in Figure 5–10. This concept strictly applies only to incoherent illumination and as a result is not really applicable to modern steppers, which generally use partially coherent illumination. However the basic idea is a useful one in understanding lithography issues.

Figure 5–10 shows a generic projection lithography system in which a reducing lens system is used to image a mask pattern in resist. Since diffraction effects are only important after the light passes through the mask, the optical intensity pattern as the light exits the mask will be almost an ideal representation of the mask. Because of diffraction effects and other nonidealities in the optical system however, the aerial image produced at the resist plane will not be perfectly black and white. If the features are widely separated, the aerial image may approach the ideal shown on the left in the Figure 5–10, but as the features move closer together, the aerial image will look more like that sketched on the right. As a simple example, imagine two Airy disks (Figure 5–7) partially overlapping. A useful measure of the quality of the aerial image is the *MTF* that can be defined as

$$MTF = \frac{I_{MAX} - I_{MIN}}{I_{MAX} + I_{MIN}} \tag{5.9}$$

where I is the intensity of the light.

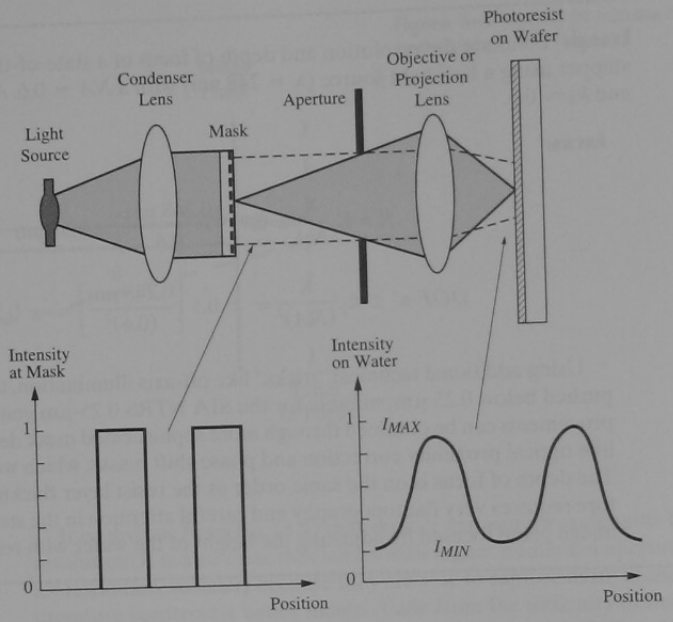


Figure 5–10 Modulation Transfer Function (*MTF*) concept. A generic lithography system is shown at the top with a mask being imaged on photoresist on a wafer. The mask *MTF* is almost ideal ($MTF = 1$) since the feature sizes are 4 – 5 X larger than those imaged in the resist and diffraction effects are minimal. The aerial image *MTF* is much lower ($MTF \approx 0.6$) because of diffraction effects in the optical system.

The *MTF* is really a measure of the contrast in the aerial image produced by the exposure system. Generally an exposure system needs to achieve a *MTF* value of 0.5 or larger in order for the resist to properly resolve the features. Recently developed Deep UV (DUV) resists can work with somewhat smaller *MTF* values.

The *MTF* obviously depends on the feature size in the image and generally has the behavior illustrated in Figure 5–11. For large features, the resulting aerial image produced by the exposure system has excellent contrast and the *MTF* is unity. As the feature size decreases, diffraction effects cause the *MTF* to degrade and to finally reach zero when the features are so closely spaced that there is no remaining contrast in the aerial image.

The *MTF* is also affected by a parameter known as the spatial coherence of the light source. Figure 5–12 illustrates the concept of spatial coherency. An ideal point source produces light in which the waves are in phase at all points along the emitted wavefronts. A condenser lens can then convert these waves to plane waves, all of which strike

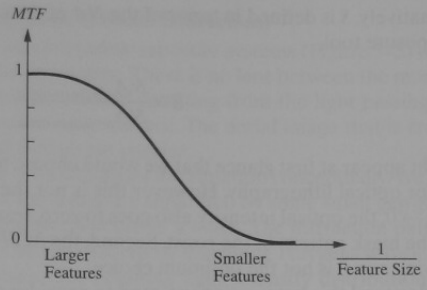


Figure 5–11 Modulation Transfer Function (*MTF*) versus feature size in the image.

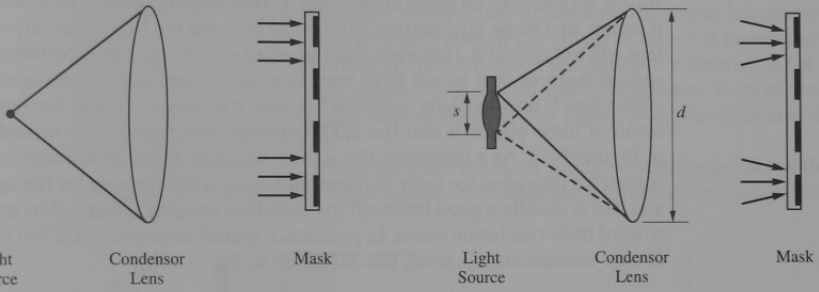


Figure 5–12 Examples of spatially coherent (left) and partially coherent (right) light sources.

the mask at exactly the same angle as illustrated in the left of the figure. Such a source is an ideal coherent source. As the physical size of the source increases as shown in the right of Figure 5–12, light is emitted from a volume rather than a point and the waves will not be perfectly in phase everywhere. If the same condenser lens is used to convert the light to plane waves, the result will be light arriving at the mask from a variety of angles as illustrated. Such a source is a partially coherent source. In the limit, if the size of the source became infinite (and the condenser lens were also infinite to collect all the light), the source would become completely incoherent. A useful definition of the spatial coherence of practical light sources for lithography is simply

$$S = \frac{\text{light source diameter}}{\text{condenser lens diameter}} = \frac{s}{d} \tag{5.10a}$$

Alternatively, S is defined in terms of the NA of the condenser and projection optics in the exposure tool,

$$S = \frac{NA_{condenser\ optics}}{NA_{projection\ optics}} \tag{5.10b}$$

It might appear at first glance that we would choose to have an ideal coherent source ($s = 0$) for optical lithography. However this is not the case for the following reasons. First, as $s \to 0$, the optical intensity also goes to zero, resulting in infinite exposure times to print the mask pattern in the resist. Second, the MTF is also affected by the value of s and having $s = 0$ is not the optimum choice.

To understand this latter point, we need to consider diffraction effects once again. If the light passing through the mask is partially coherent (i.e., coming in at a variety of angles), then the diffraction patterns resulting from mask features will be smeared out. If we refer again to Figure 5–8, this means that the diffraction pattern for a given feature will be spread over an angle larger than α. This initially seems like a bad thing, because it means that some information will be lost because the finite aperture of the objective lens will not collect it. However, this smearing also means that information from closely spaced features that would have been completely lost outside the aperture of the focusing lens is now partially collected because it is smeared inside the lens aperture. The result of these effects is that the MTF behavior with feature size is modified, as shown in Figure 5–13. As s increases, the source becomes more incoherent, and the MTF is somewhat degraded for large features. However, it is improved for the smallest features and this is usually a good trade-off in projection imaging systems that are being pushed toward their resolution limits. In practice, a spatial coherence of 0.5 to 0.7 is often used in chip manufacturing, using the definition in Eq. (5.10b).

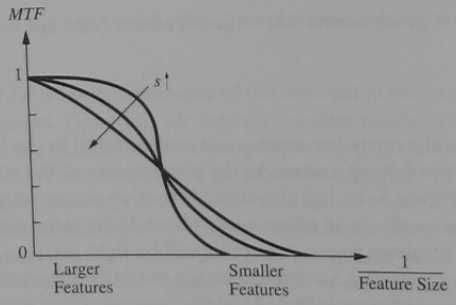


Figure 5–13 Modulation Transfer Function (MTF) versus feature size in the image. As s increases (more incoherent source), the MTF degrades for larger features but improves for very small features.

5.2.2.3 Contact and Proximity Systems (Fresnel Diffraction)

Contact and proximity exposure systems (Figure 5–3) operate in the near field or Fresnel diffraction regime. There is no lens between the mask and the resist on the wafer, so the diffraction pattern resulting from the light passing through the mask directly impinges on the resist surface. The aerial image that is created therefore depends on the near field diffraction pattern.

Figure 5–14 illustrates the situation conceptually. We assume for the moment that the mask and the wafer are separated by some small gap g. A plane wave is assumed to be incident on an aperture in a mask. The diffraction pattern on the opposite side of the mask can be constructed by imagining Huygens wavelets emanating from each point in the aperture. The resulting light intensity distribution striking the top surface of the resist is also illustrated in Figure 5–14. There are several features in the light distribution of interest. First, notice that the intensity rises gradually near the edges of the mask aperture. Because of diffraction effects, the light "bends" away from the aperture edges, producing some resist exposure outside the aperture edges. Second, note the "ringing" in the intensity distribution within the aperture dimension. This arises because of constructive and destructive interference between the Huygens wavelets emanating from the aperture. Contact and proximity printers often use multiple wavelengths for exposure and also do not use light sources with perfect spatial coherence. Both of these approaches minimize the ringing effects illustrated in Figure 5–14, but of course do not eliminate diffraction effects.

As the separation g increases between the mask and the resist, the quality of the aerial image produced at the resist surface will degrade because diffraction effects will

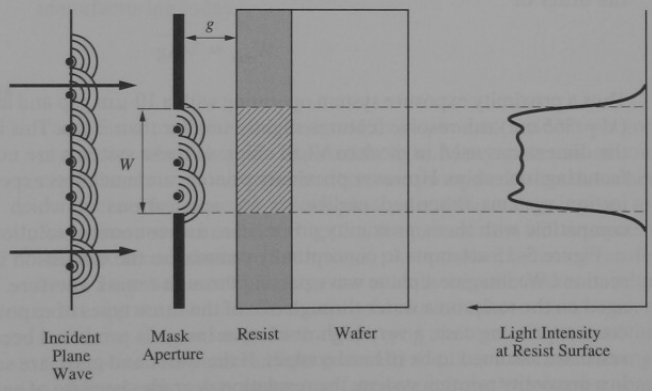


Figure 5–14 Basic contact or near field exposure system illustrating the use of Huygen's wavelets emanating from an aperture in the mask. A mask feature size of W is assumed, along with a mask to resist separation of g. The resulting light intensity distribution (aerial image) at the resist surface is shown on the right.

become even more important. Contact printers minimize these effects by attempting to reduce g to zero, but in most practical systems g is not strictly zero because the top surface of the resist is not perfectly flat due to topography on the wafer surface. In the limit when the mask and wafer are pressed into hard contact, the resolution of such systems can be very good (well below 0.1 μm). However, even in this case, the resist itself still has a finite thickness and the resolution will still be limited by light scattering in the resist and by light reflection from surface features on the underlying wafer which scatter the light laterally into regions adjacent to the mask apertures. We will discuss some of these scattering issues more carefully in the next section on resist exposure.

Generally the aerial image can be calculated using Fresnel diffraction theory whenever the gap g falls within the limits

$$\lambda < g < \frac{W^2}{\lambda} \tag{5.11}$$

where W is the size of the mask aperture (feature size). The lower limit on g is certainly satisfied by proximity printing systems and is often satisfied by contact printing systems unless hard contact is used. If g falls below the wavelength of the light used for the exposure, the resulting light intensity distribution can still be calculated but only by a full numerical solution to Maxwell's equations. Such solutions are very complex, but fortunately are rarely needed [5.6]. The upper bound on g arises because when g increases, Fresnel diffraction theory must be replaced by far field (Fraunhofer) diffraction theory in order to calculate the aerial image.

Within the Fresnel diffraction range, the minimum resolvable feature size is on the order of

$$W_{min} \approx \sqrt{\lambda g} \tag{5.12}$$

Thus a proximity exposure system operating with a 10-μm gap and an i-line light source (λ = 365 nm) can resolve features slightly smaller than 2 μm. This is much larger than the dimensions used in modern VLSI chips, so these systems are not useful for manufacturing such chips. However proximity printers are much less expensive than the projection systems described earlier, so for applications in which features sizes are compatible with them, proximity printers are an economical solution.

Figure 5–15 attempts to conceptually summarize the discussion in the previous two sections. We imagine a plane wave passing through a mask aperture. The aperture is imaged on the resist on a wafer through one of the three types of exposure systems. In the contact printing case, a very high resolution image is produced because the mask and resist are assumed to be in hard contact. If the wafer and mask are separated slightly as in a proximity printing system, the resolution degrades because of near field Fresnel diffraction effects. Finally, if we place a lens between the mask and wafer and focus the aperture on the wafer, an image characterized by Fraunhofer diffraction is produced. In the example in this figure, the resolution of the proximity system image is illustrated as inferior to both of the other systems. This is usually the case in actual systems, as we saw

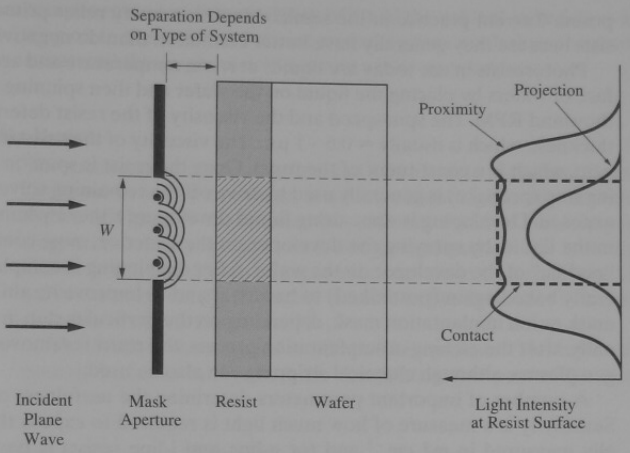


Figure 5–15 Aerial images produced by the three types of optical lithography tools. The mask and wafer would be in hard contact in a contact aligner, separated by a gap g in a proximity aligner, and far apart with an intervening focusing lens in a projection system.

in the numerical example earlier and is the reason why projection systems are used in manufacturing today.

5.2.3 Photoresists

Photoresist materials are designed to respond to incident photons by changing their properties when they are exposed to light. Many materials of course absorb light, but often the absorption results in electronic processes rather than chemical changes. Semiconductors, for example, absorb photons and the energy is given to electrons and holes. As we have seen in earlier chapters, the free carriers will dissipate the absorbed energy through recombination or through phonon interactions (transferring the energy to heat). In some cases the energy can actually be collected as in solar cells. None of these processes are useful in lithography because in photoresists we require a material that maintains a latent image of the impinging photons at least until the resist is developed. A long-lived response to light generally requires a chemical change in the material.

Almost all resists today are fabricated from hydrocarbon-based materials. When these materials absorb light, the energy from the photons generally breaks chemical bonds. After this happens, the resist material chemically restructures itself into a new stable form. Positive resists respond to light by becoming more soluble in the developer solution. Negative resists do the opposite. They become less soluble where they are ex-

posed. Current practice in the semiconductor industry relies primarily on positive resists because they generally have better resolution than do negative resists.

Photoresists in use today are liquids at room temperature and are applied to the surface of wafers by placing the liquid on the wafer and then spinning the wafer at several thousand RPM. The spin speed and the viscosity of the resist determine the final resist thickness, which is usually $\approx 0.6 - 1\ \mu m$. The viscosity of the resist is controlled by a solvent, which is a constituent of the resist. Once the resist is spun on to the wafer, a baking step (prebake) is generally used to drive off the remaining solvent. The resist is then exposed. Developing is done using liquid developers either by immersion of the wafer in the liquid, by spraying the developer on the wafer or, most commonly, by placing a "puddle" of the developer on the wafer. After developing is complete, the resist is generally baked again (postbaked) to harden it and to improve its ability to act as an etch mask or ion implantation mask, depending on the particular step in the process flow. Finally, after the etching or implantation process, the resist is removed, usually in an oxygen plasma, although chemical stripping can also be used.

A number of important parameters determine the usefulness of a particular resist. Sensitivity is a measure of how much light is required to expose the resist. This is usually measured in mJ cm^{-2} and for g-line and i-line resists is typically 100 mJ cm^{-2}. Newer Deep UV (DUV) resists often achieve sensitivities of 20–40 mJ cm^{-2} because they use chemical amplification, a process we will describe later. Generally, high sensitivity is desired because this decreases the exposure time of the resist and therefore improves throughput in the lithography process. However, extremely high sensitivity is usually not desired because this tends to make the resist material unstable, makes it very sensitive to temperature, and can also create problems with statistical variations due to shot noise during the exposure. Usually higher contrast and more process latitude are achieved at lower sensitivities. However, DUV chemically amplified resists achieve both higher sensitivity and higher contrast than the older g-line and i-line resists. Since the light used to expose the resist is at a specific wavelength, it is also important that the sensitivity of the resist be optimized for the exposure wavelength.

Resolution is obviously important in resists. The quality of resist patterns today is generally limited by the exposure system (aerial image) and not by the resist itself. However the resist materials and the process steps (exposure dose, baking, and developing cycles) must be carefully controlled to achieve diffraction limited resolution in the resist images.

The third parameter of importance has to do with the "resist" function of photoresists. The term "resist" describes the need for the photoresist to withstand etching or ion implantation after the mask pattern is transferred to the resist. Practical resists need to have reasonable robustness with respect to these processes. In practice this means that resists must be able to resolve small features even when the resist is of reasonable thickness.

g-line and i-line photoresists generally consist of three components, an inactive resin which is usually a hydrocarbon which forms the base of the material, a photoactive compound (*PAC*) which is also a hydrocarbon, and a solvent which is used to adjust the viscosity of the resist. DUV resists replace the *PAC* component with a Photo-Acid Generator (*PAG*) which acts as a chemical amplifier or catalyst, and often add other

components as well for stability. Most of the solvent in resists evaporates during spinning on to the wafer and during prebake processes before the resist is exposed, leaving a material that is about 1:1 base and active components.

5.2.3.1 g-line and i-line Resists

The most commonly used g-line and i-line resists today are diazonaphthoquinones or DNQ materials. A base resin is added which is generally novolac, with the structure shown in Figure 5–16. Novolac is a polymer material consisting of basic hydrocarbon rings with 2 methyl groups and 1 OH group attached. The basic ring structure which is shown in the figure may be repeated many times to form a long chain polymer material. Novolac by itself will readily dissolve in the developer solution at a typical dissolution rate of about 15 nm sec^{-1}.

The *PAC*s in these resists are often diazoquinones. The basic structure of these compounds is shown in Figure 5–17. The photoactive part of the structure is the portion

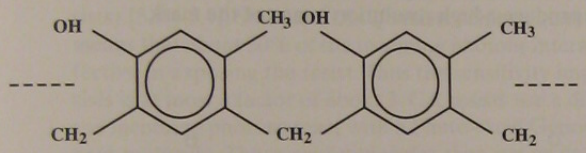


Figure 5–16 Basic structure of novolac, a thick resin used as the base material in positive photoresists.

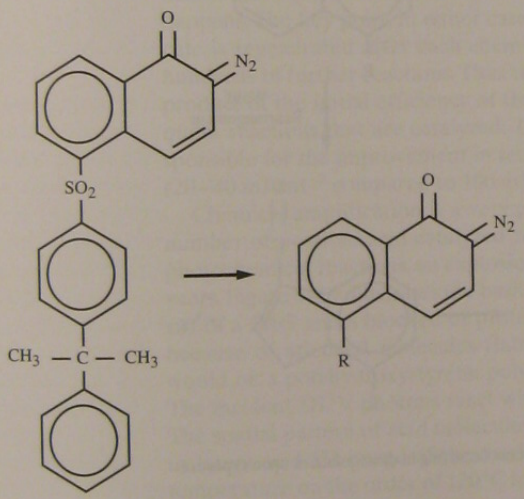


Figure 5–17 Basic structure of diazoquinone, a commonly used photoactive compound in positive photoresists. R represents the bottom part of the molecule.

above the SO_2. The remainder of the molecule is often abbreviated as shown in the figure. The role of the *PAC* is to inhibit the dissolution of the resist material in the developer. Diazoquinones are insoluble in typical developers and they reduce the overall dissolution rate of the resist to approximately 1–2 nm sec^{-1}. Thus the DNQ material is essentially insoluble in the resist developer before it is exposed to light.

When the resist is exposed to light, the diazoquinone molecules chemically change as illustrated in Figure 5–18. The N_2 molecule is weakly bonded in the *PAC* and the first part of the photochemical reaction involves the light breaking this bond. This leaves behind a highly reactive carbon site. The *PAC* structure can stabilize itself by moving a carbon atom outside the ring with the oxygen atom covalently bonded to it. This is known as the Wolff rearrangement. The resulting ketene molecule finally transforms into carboxylic acid (lower left in Figure 5–18) in the presence of water. The carboxylic acid is now readily soluble in a basic developer [typically TMAH (tetramethyl ammonium hydroxide), KOH, or NaOH dissolved in H_2O]. The novolac matrix material is also readily soluble in this solution. The exposed resist material thus dissolves at a rate of 100 – 200 nm sec^{-1}. The unexposed regions of the resist are essentially unaffected by the developer and so if the exposed pattern accurately reproduces the mask pattern, the photoresist can produce a high resolution image of the mask.

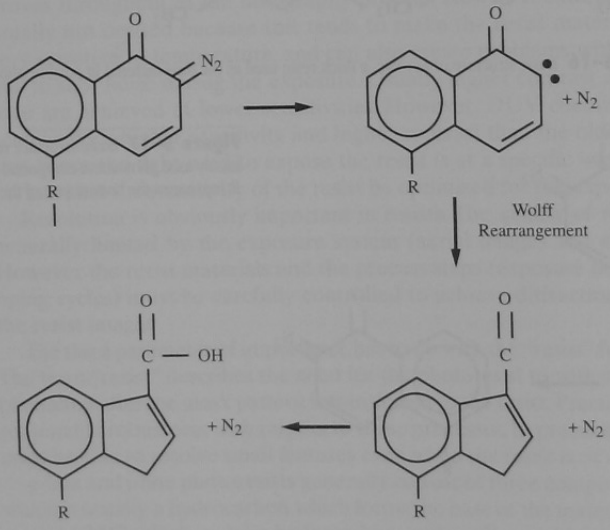


Figure 5–18 Decomposition process occurring in diazoquinones upon exposure to light.

5.2.3.2 Deep Ultraviolet (DUV) Resists

Conventional DNQ resist materials have two significant problems when shorter exposure wavelengths are used. The first is that for wavelengths below i-line (365 nm), these resists strongly absorb the incident photons. Thus the incident radiation cannot penetrate through the full thickness of the resist. This is a significant issue at 248 nm (KrF) which is the wavelength used for the 0.25-μm and 0.18-μm generations. The second problem has to do with resist sensitivity. Some years ago when resists suitable for DUV applications were first being explored, the only viable light source was the Hg arc lamp. These sources work well for g-line and i-line systems, but the intensity of their output in the DUV is much lower than at i-line. Thus it was believed at that time that whatever resist was used for DUV applications, it would have to have improved sensitivity over the standard DNQ resists, in order to maintain manufacturing throughput. Of course bright excimer laser sources are now available and are reliable at 248 nm and so the sensitivity issue is not as crucial. As a result, DNQ-type resists are being revisited today with modified *PAC* compounds that work more effectively at 248 nm.

DUV resists that are in use today, however, are not modified DNQ resists. They are based on a completely new chemistry and make use of chemical amplification (CA resists) [5.7 – 5.9]. Standard DNQ resists achieve quantum efficiencies of about 0.3. This means that about 30% of the incoming photons interact with *PAC* molecules and are effective in exposing the resist. Thus the sensitivity improvement possible with these resists is at most a factor of about 3. CA resists use a different exposure process in which the incoming photons react with a Photo-Acid Generator (*PAG*) molecule, creating an acid molecule. These acid molecules then act as catalysts during a subsequent resist bake to change the resist properties in the exposed regions. Both positive and negative resist versions are possible. In the positive resist case, the *PAG* initiates a chemical reaction that makes the resist soluble in the developer; in a negative resist, the opposite happens. The key point in either case is that the reactions are catalytic; the acid molecule is regenerated after each chemical reaction and may thus participate in tens or hundreds of further reactions. Thus the overall quantum efficiency in a CA resist is the product of the initial efficiency of the light/*PAG* reaction, times the number of subsequent reactions that are catalyzed. This product can be much larger than 1 and is responsible for the improvement in sensitivity of DUV resists compared to DNQ resists (20–40 mJ cm^{-2} compared to 100 mJ cm^{-2}).

Chemical amplification is a very powerful new approach to create resists. Since the number of possible acid catalyzed reactions is very much larger than the number of photochemical reactions, an explosion of new resist possibilities has occurred in recent years. Figure 5–19 illustrates the basic principle behind these resists. Positive resists consist of a *PAG* and a blocked or protected polymer, which is insoluble in the developer because of attached molecules (labeled INSOL in Figure 5–19). A typical example would be a polyhydroxystyrene polymer with attached acid labile groups [5.10, 5.11]. The incident DUV photons react with the *PAG* molecules to create an acid molecule. The spatial pattern of acid molecules in the resist after the exposure is thus a "stored" or latent acid 3D image of the mask pattern. After exposure, the wafer is baked at a temperature on the order of 120°C for a few minutes (Postexposure Bake or PEB). The

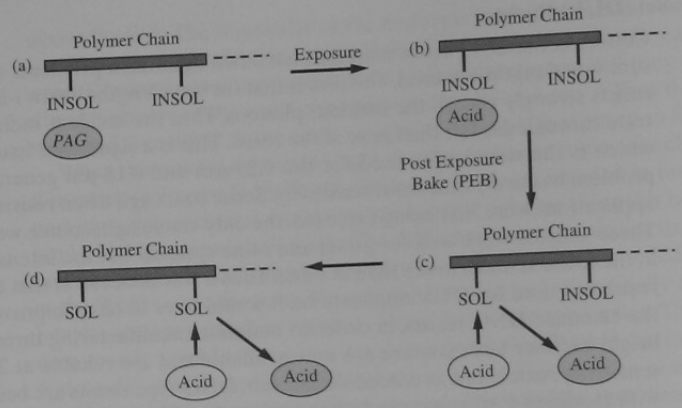


Figure 5–19 Basic operation of a Chemically Amplified (CA) resist. *PAG* is the photo-acid generator; INSOL and SOL are the insoluble and soluble portions of the polymer base. Steps (c) and (d) may repeat tens or hundreds of times during the PEB.

heat provides the energy needed for the reaction between the acid molecules and the insoluble fragments on the polymer chains to take place. It also provides mobility (through diffusion) for the acid molecules to find the insoluble fragments and to react with tens or hundreds of such fragments during the PEB. Thus during the PEB, the insoluble blocked polymer is converted into an unblocked polymer that is soluble in an aqueous alkaline developer. In negative working DUV resists, the *PAG* catalyzes a reaction which crosslinks the polymer chains, making the resist insoluble in the developer. The key mechanism in either process is the catalytic behavior of the acid molecules that are regenerated after each reaction.

While these new resists offer superb sensitivity and resolution more than adequate for today's device structures, they also require very careful manufacturing control. Because there is a time delay between the light exposure and the acid reaction with the polymer chain fragments, poisoning (contamination) of the acid molecules is a concern. Some early DUV resists were very sensitive to very small concentrations of airborne contaminants that reacted with the acid molecules before they could catalyze the resist reactions during the PEB. This resulted in a surface layer on the resist where the acid was neutralized and the resist effectively unexposed. These kinds of problems have been largely solved by adding additional components to DUV resists, by adding protective surface layers, and by careful environmental and manufacturing control. DUV resists also require very careful control of the PEB conditions because that bake is used to drive the chemical reaction that completes the resist exposure. Chemical reactions and the diffusion associated with the acid molecules during the PEB generally depend exponentially on temperature, requiring temperature control on the order of a fraction of a °C during the PEB.

5.2.3.3 Basic Properties and Characterization of Resists

Two basic parameters are often used to describe the properties of photoresists: contrast and the Critical Modulation Transfer Function (*CMTF*). We will define these terms and explain their importance in the following paragraphs. Contrast is really a measure of the resist's ability to distinguish light from dark areas in the aerial image the exposure system produces. As we have seen, diffraction effects and perhaps other imperfections in the exposure system result in an aerial image that does not have abrupt transitions from dark to light. An important question is how the resist responds to the "gray" region at the edges of features in the aerial image.

Contrast is an experimentally determined parameter for each resist and its value is extracted from plots like that shown in Figure 5–20. The data from which these plots can be derived are obtained by exposing resist layers to a variety of exposure doses. Each of the samples is then developed for a fixed period of time and the thickness of the photoresist remaining after developing is measured. For positive resists, samples receiving small exposure doses will not be attacked by the developer to any appreciable extent; those receiving large doses will completely dissolve in the developer. Intermediate doses will result in partial dissolution of the resist. For negative resists, the opposite behavior occurs. Given data like those shown in the figure, the contrast is simply the slope of the steep part of the curve, defined as

$$\gamma = \frac{1}{\log_{10}\frac{Q_f}{Q_O}} \tag{5.13}$$

where Q_O is the dose at which the exposure first begins to have an effect and Q_f is the dose at which exposure is complete.

Typical g-line and i-line resists achieve contrasts of 2–3 and Q_f values of about 100 mJ cm^{-2}. DUV resists achieve significantly better contrast and better sensitivities than this.

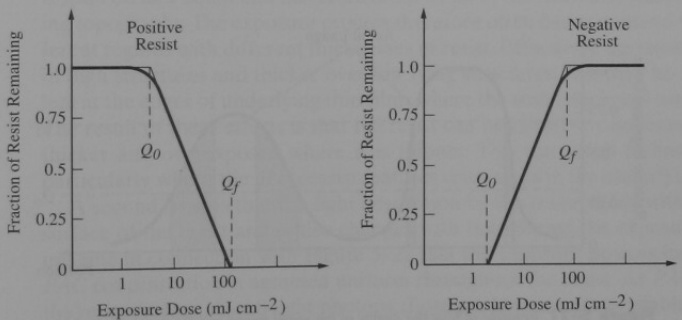


Figure 5–20 Idealized contrast curves for positive and negative resists.

This is basically because the chemical amplification that occurs in DUV resists steepens the transition from the unexposed to the exposed condition. (Once the reaction is begun, the catalytic nature of the process carries it to completion unlike DNQ resists where the *PAC* molecules in the resist must be exposed one by one by incoming photons during the exposure.) Thus DUV resists typically achieve γ values of 5 – 10 and Q_f values of about 20 – 40 mJ cm^{-2}.

It is important to note, however, that γ is not a constant for a particular resist composition. Rather, the experimentally extracted value of γ depends on process parameters like the development chemistry, bake times, temperatures before and after exposure, the wavelength of the exposing light, and the underlying structure of the wafer on which the resist is spun. In general it is desirable to have a resist with a high contrast because this produces better (steeper) edge profiles in the developed resist patterns. (See Figure 5–21.) Intuitively, this arises because a high contrast implies that the resist sharply distinguishes between dark and light areas in the aerial image. Thus resists with high contrast can actually "sharpen up" a poor aerial image.

The Modulation Transfer Function (*MTF*) of the aerial image was defined in Eq. (5.9) and in Figure 5–10. The *MTF* is simply a measure of the "dark" versus "light" intensities in the aerial image produced by the exposure system. It is often useful to define a similar quantity for the resist, in this case called the Critical *MTF* or *CMTF*. The CMTF is roughly the minimum optical transfer function necessary to resolve a pattern in the resist.

$$CMTF_{resist} = \frac{Q_f - Q_O}{Q_f + Q_O} = \frac{10^{1/\gamma} - 1}{10^{1/\gamma} + 1} \tag{5.14}$$

Typical *CMTF* values for g-line and i-line resists are around 0.4. With their higher γ values, chemically amplified DUV resists achieve significantly smaller *CMTF* values ($\approx$ 0.1 – 0.2). The significance of this number is that the *CMTF* must be less than the aerial image *MTF* if the resist is going to resolve the aerial image.

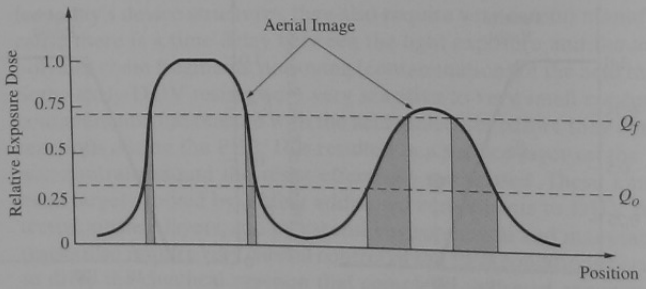


Figure 5–21 Example of how the quality of the aerial image and the resist contrast combine to produce the resist edge profile. The left side shows a sharp aerial image and steep resist edges (gray area). The example on the right shows a poorer aerial image and the resulting gradual edges on the resist profile.

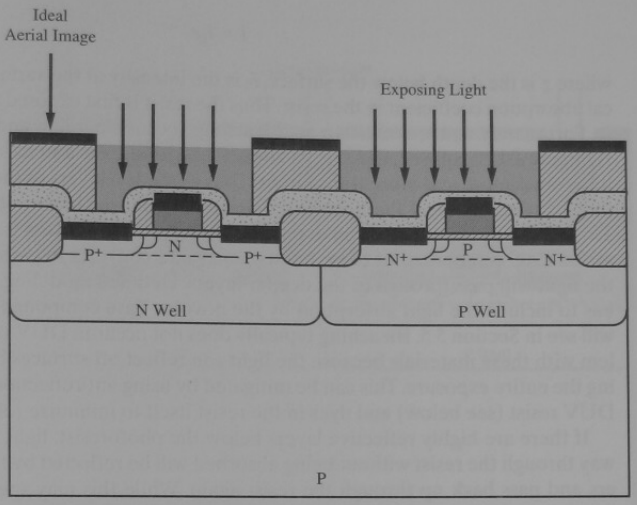


Figure 5–22 Exposure process associated with Figures 2–36 and 2–37 in the CMOS process in Chapter 2. The black bars in the aerial image represent areas on the mask that are opaque and hence ideally no photons strike the resist in these areas.

To this point in our discussion of resists we have effectively treated them as having uniform thickness and we have also treated the exposure process as occurring simultaneously throughout the volume of the resist material. In fact, neither of these is a very good assumption in many cases. Figure 5–22 illustrates some of these issues. Notice in that figure that the resist thickness tends to be nonuniform across the wafer because it is spun on as a liquid and therefore tends to fill in the "hills and valleys" of the underlying topography. The exposure process therefore often has to contend with exposing different regions with different thicknesses of resist. Effectively the resist is thinner on top of high structures and thicker over low lying structures. This may be a particular problem at the edges of underlying thin films where the resist thickness can change abruptly. The result of these effects is that the resist can be effectively underexposed where it is thicker and overexposed where it is thinner. This can result in linewidth variations, particularly where the photoresist features cross steps in the underlying structures.

A second issue is that the light absorption by the resist varies with depth below the surface of the resist and it also changes with time during the exposure process. Again imagine in connection with Figure 5–22 that the exposure process has just begun. The *PAC* concentration is assumed uniform throughout the resist. As *PAC* molecules near the resist surface absorb light photons, those photons are not available for exposing the deeper layers of the resist. Thus the light concentration falls off with depth. We will model this more carefully in Section 5.5, but to first order, the light intensity falls off exponentially with distance into the resist.

$$I = I_0 e^{-\alpha z} \tag{5.15}$$

where z is the depth below the surface, I_0 is the intensity at the surface, and α is the optical absorption coefficient in the resist. Thus the resist is first exposed near the top surface.

Fortunately a process known as "bleaching" occurs in g-line and i-line DNQ resists. As the resist is exposed, the *PAC* is altered, absorbing less and less light, and therefore it becomes more and more transparent. Thus as the top layers of the resist are exposed, they transmit more of the light to the deeper layers which are subsequently exposed. This results in a more uniform exposure. The bleaching process is not surprising since as the *PAC* component of these resists reacts and converts to carboxylic acid, more of the light will pass through to the deeper layers. Detailed modeling of these effects also has to include the light absorption by the novolac base component of the resist, as we will see in Section 5.5. Bleaching typically does not occur in DUV resists. This is a problem with these materials because the light can reflect off surfaces below the resist during the entire exposure. This can be mitigated by using antireflection coatings below the DUV resist (see below) and dyes in the resist itself to minimize reflections.

If there are highly reflective layers below the photoresist, light which passes all the way through the resist without being absorbed will be reflected by these underlying layers and pass back up through the resist again. While this may speed up the exposure process, it also has the potential for setting up standing wave light patterns in the resist because of constructive and destructive interference between the incoming and outgoing waves. Furthermore, if the light is scattered sideways, image resolution can be degraded. Figures 5–23 and 5–24 illustrate these issues. In some cases an Antireflective Coating (ARC) is deposited on the wafer prior to spinning on the resist. This can greatly help in minimizing standing wave effects, but of course increases process complexity. Another approach that is often used in g-line and i-line resists is to add dyes to the resist, which absorb light and minimize reflections.

The "scalloped edges" illustrated in Figure 5–24 are also strongly affected by post-baking of the resist after exposure but before development, since this heat treatment allows diffusion of the *PAC* in g-line and i-line resists or the *PAG* in DUV resists on a limited basis. This smoothes out the standing wave pattern. It is quite possible to model these standing wave effects and to predict the resulting resist edge patterns that result. We will consider these issues more carefully in Section 5.5.

5.2.4 Mask Engineering—Optical Proximity Correction and Phase Shifting

In our discussion to this point, we have considered the mask to be simply a digital device. That is, it has clear areas and dark areas and the patterns of these areas represent exactly the pattern that we wish to print in the resist on the wafer. It is actually possible to do better than this in designing the mask if the objective is to produce the highest quality aerial image. We will briefly discuss two approaches to doing this: Optical Proximity Correction (OPC) and Phase Shift Masks (PSMs). These approaches might be considered "mask engineering." These methods are also sometimes called "wavefront engineering" [5.12].

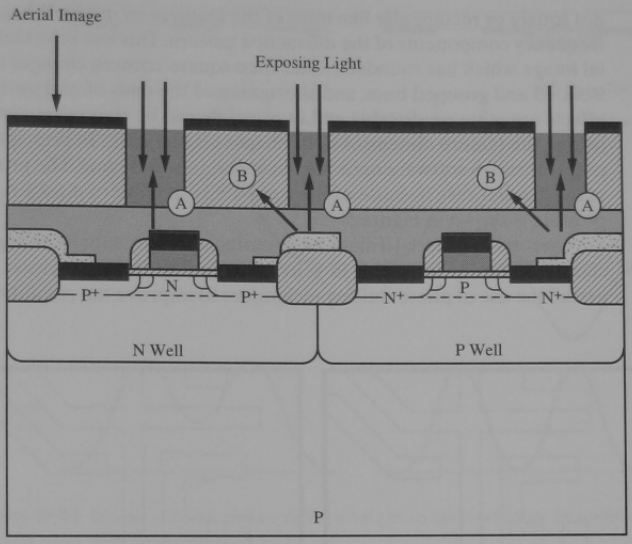


Figure 5–23 Photolithography exposure occurring between Figures 2–39 and 2–40 in the CMOS process described in Chapter 2. The A and B symbols indicate examples of standing wave patterns and sideways light scattering, respectively.

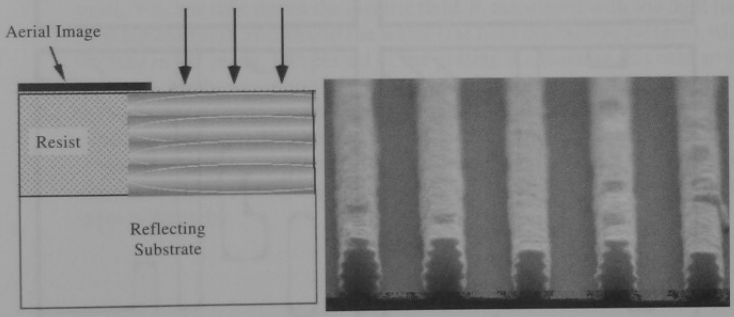


Figure 5–24 The light intensity pattern resulting from a standing wave is sketched on the left. A photoresist image that illustrates these effects after developing is shown on the right. Courtesy of A. Vladar and P. Rissman, Hewlett Packard.

We saw in our discussion of exposure systems that the finite aperture of projection systems results in a loss of some of the light diffracted from the mask features. In addition to their finite size, apertures and lenses in projection systems are generally circular,

not square or rectangular like most of the features on masks. What are lost are the high-frequency components of the diffraction pattern. This lost information results in an aerial image which has rounded rather than square corners, changes in linewidth between isolated and grouped lines, and shortening of the ends of narrow linear features. These effects are quite predictable and in principle can be somewhat compensated for by adjusting the feature dimensions and shapes on the mask. In principle, this could be done purely through software once the mask design is complete. The problem is very difficult in the general case because of the complexity of modern masks. An example of what is possible is shown in Figure 5–25 [5.13].

A second approach to mask engineering involves actually changing the transmission characteristics of the mask in selected areas. In 1982, Levenson and colleagues suggested the use of phase shifting techniques to improve the resolution of the printed aer-

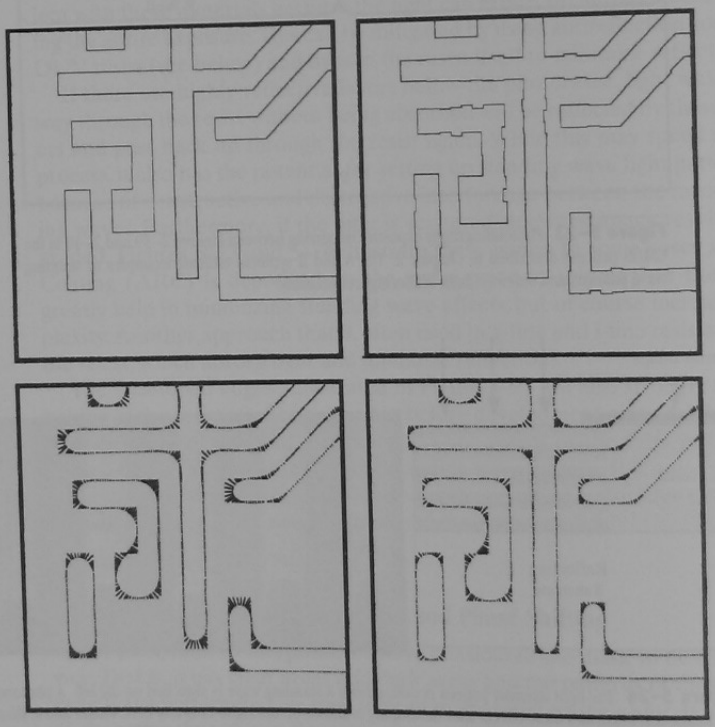


Figure 5–25 Mask patterns with (right) and without (left) OPC are shown on the top. The corresponding aerial images (calculated) are shown on the bottom. Note the improvement in the quality of the aerial image when OPC is used. The dark lines in the bottom patterns indicate the difference between the mask and aerial image in each case [5.13]. Reprinted with permission of SPIE.

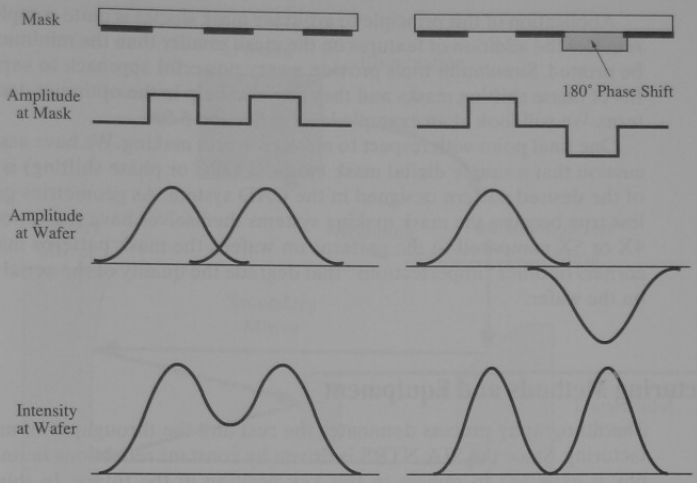


Figure 5–26 Example of the use of phase shifting techniques on masks to improve the resolution of the aerial image. (After Levenson et al. [5.32].)

ial image [5.14]. A simple example of the principle is illustrated in Figure 5–26. In this example a periodic mask with equal lines and spaces (diffraction grating) is used as the mask. The left side shows the electric field, ε, associated with the light just after it passes through the mask and also at the wafer (far field diffraction pattern), without any phase shifting in the mask. The period of the grating is chosen in this case so that the lines on the mask are barely resolved on the wafer. The photoresist responds to the intensity of the light or (ε field)^2 and hence the intensity pattern on the bottom left is barely sufficient to resolve the two lines.

In the example on the right, a material whose thickness and index of refraction are chosen to phase shift the light by exactly 180° is added to the mask. The thickness of this layer is given by

$$d = \frac{\lambda}{2(n-1)} \tag{5.16}$$

where n is the index of refraction of the phase shift material. The period of this added pattern is half the period of the original grating. The corresponding ε fields at the mask and wafer are also shown. Since the light intensity at the aerial image is the square of the ε field intensity, the quality of the resulting aerial image is significantly improved as illustrated. Such phase shifting techniques can either be used to improve the quality of the aerial image or to improve the depth of focus of the exposure system at a constant resolution by using a lower NA system.

Application of this principle to arbitrary mask shapes is quite complex and generally requires the addition of features on the mask smaller than the minimum feature size to be printed. Simulation tools provide a very powerful approach to exploring the benefits of phase shifting masks and they can also help in the optimum design of mask patterns. We will look at an example later in Section 5.5.

One final point with respect to masks is worth making. We have assumed in our discussion that a simple digital mask (without OPC or phase shifting) is an exact replica of the desired pattern designed in the CAD system. As geometries get smaller, this is less true because the mask making systems themselves have resolution limits. Even at 4X or 5X compared to the patterns on wafers, the mask patterns may have rounded corners or other "imperfections" that degrade the quality of the aerial image produced on the wafer.

5.3 Manufacturing Methods and Equipment

The lithography process dominates the cost and the throughput of modern IC manufacturing. Since the SIA NTRS is driven by constant reductions in linewidth, lithography is expected to remain in this key position in the future. In this section we will examine typical industrial processes and equipment for lithography. While contact and proximity printers played a major role in the early days of the silicon industry, they no longer do so because of unacceptable defect levels and/or limited resolution. As a result, we will restrict our discussion in this section to projection aligners. Also, since DNQ positive resists and DUV resists for 248-nm systems dominate industrial use today, we will restrict our discussion to these materials.

5.3.1 Wafer Exposure Systems

Projection aligners have been the dominant exposure tools in the silicon industry for more than 20 years. Perkin-Elmer Corp. pioneered the earliest systems of this type under the name Micralign. These systems were scanning projection aligners, which used 1:1 masks. The principle behind such scanning systems is illustrated in Figure 5–27. The basic idea in these systems is that it is easier to correct optics for aberrations in a small area rather than a large area. Thus the illumination source produces a slit of light which is mechanically scanned across the mask. The wafer is simultaneously scanned so that the mask pattern is printed across the wafer. The system that is illustrated in Figure 5–27 is a purely reflective system, using mirrors. Systems of this general type are also used which have a combination of mirrors and lenses (both reflecting and refracting optics). Such optical systems are known as catadioptic optical systems.

While scanning systems of the type illustrated in Figure 5–27 are very cost effective and have high throughput, they require 1X masks that contain the pattern information for all the chips to be printed on each wafer. Linewidth control on such masks becomes more difficult as geometries shrink. Also, as chips became more complex and as wafers became larger in the late 1970s, it became increasingly difficult to produce perfect full

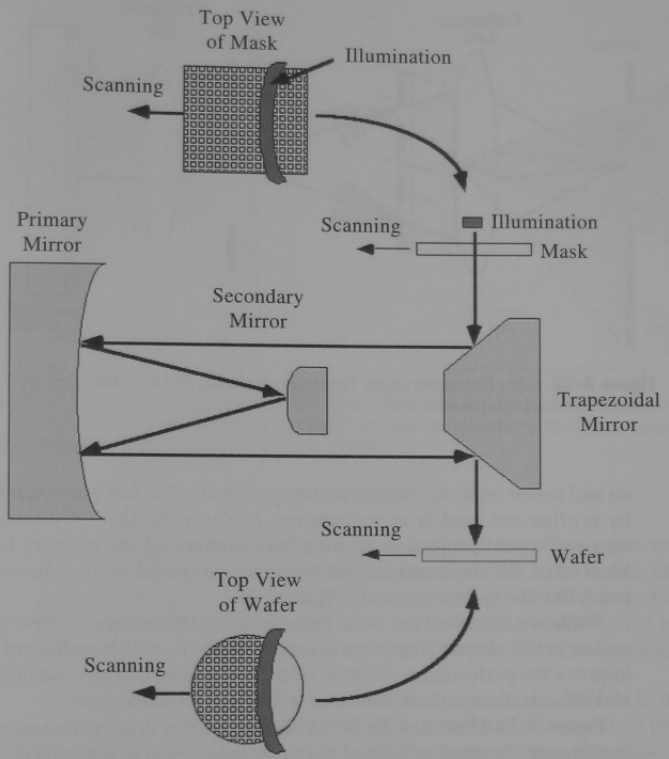


Figure 5–27 Conceptual diagram of a scanning projection printer.

wafer masks and to make optical systems which could scan entire wafers with the required resolution. Scanning systems also generally only allow global alignment of the mask and the wafer patterns, that is, alignment at the wafer level before the scanning begins. With tighter dimensions and the corresponding need for better placement accuracy, local alignment on a die by die or similar scale is often highly desirable. The solution to these issues was to replace the scanning systems with steppers that exposed a limited portion of the wafer (typically a few cm^2) at a time. This eliminates wafer size as a major issue. The mask problem is also addressed in these systems by making the steppers reduce the image by 4X or 5X. Thus the masks use much larger dimensions which makes them more easily repairable and they contain only a few die rather than a full wafer pattern. Steppers also allow alignment on an exposure field by exposure field ba-

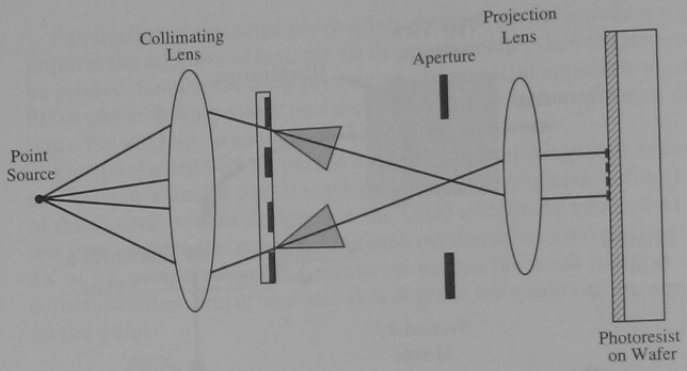


Figure 5–28 Kohler illumination system. The triangles emanating from the mask regions represent the diffracted light pattern.

sis and hence improve overlay accuracy compared to full wafer scanners. (This capability is often not used in manufacturing, however, because of throughput issues.) Such step and repeat projection aligners have dominated the industry for the past decade. Most often, these systems use purely refractive optical systems. In concept they are very much like the system shown in Figure 5–3.

While we described the basic features and performance of these projection steppers earlier in this chapter, there are several key additional ideas that are used in practice to improve the performance of these systems. These include the use of Kohler illumination and off-axis illumination, which we will briefly describe here.

Figure 5–28 illustrates the idea behind a Kohler illumination source. The light passing through the mask is focused at the entrance pupil of the projection lens, rather than having collimated light pass through the mask as illustrated conceptually in Figure 5–3. The reason for using a Kohler source arrangement is so that the projection lens can capture the diffracted light from any of the features on the mask equally well. This is illustrated in Figure 5–28 by the triangles emanating from the mask apertures that represent the angular spread of the diffracted light. If collimated light were used though the mask, it is apparent that much of the diffracted light would be lost from mask features near the outside edge of the mask.

The second "trick" which is often used in modern projection aligners is off-axis illumination. The idea behind this technique can be understood in connection with Figure 5–29. If the light from a coherent source is incident at an angle to the mask, rather than normal to the mask, this illumination is called off-axis illumination. Such illumination obviously changes the angle of the light passing through the mask and changes the angle of the diffracted light as well. While some of the diffracted light will be lost through the use of off-axis illumination because it will be outside the projection lens aperture, some of the higher-order diffracted light which would be lost through nor-

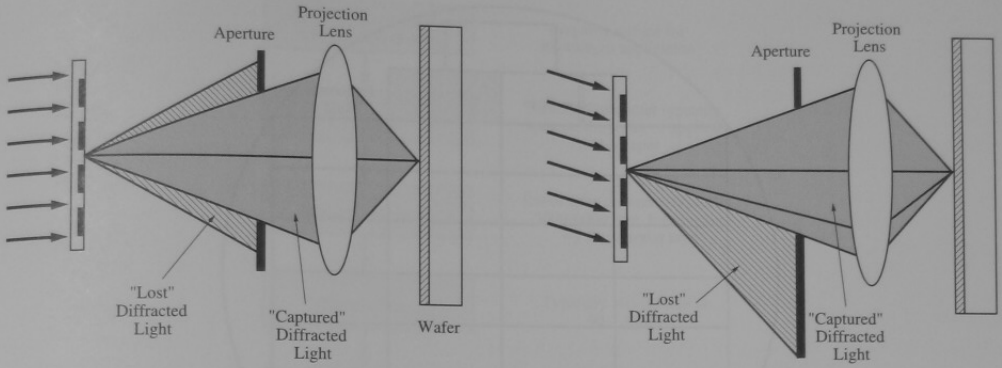


Figure 5–29 Illustration of the idea of on-axis illumination (left) and off-axis illumination (right). The shaded areas emanating from the mask regions represent the diffracted light pattern. Practical implementations generally use a light source producing several angles of incident light (+ and −). Some of the higher order diffracted information is captured in the off-axis scheme.

mally incident radiation will also be captured in the off-axis system. The net result is that the resolution can be somewhat improved. The basic idea is similar to the arguments presented earlier for using partially coherent illumination rather than coherent illumination. It is apparent then, that considerable attention must be paid to the illumination part of the optical system as performance in projection aligners is pushed toward physical limits.

Within the past few years, it has become increasingly difficult to build steppers that have the resolution required by the SIA NTRS and also the field of view required by this roadmap. Because of this, a hybrid type of system which uses both stepping and scanning has recently come into use. Figure 5–30 illustrates the principle. Stepping is used to move the wafer between major exposure fields. Within each exposure field, the mask pattern is scanned across the wafer. This type of system has some of the advantages of scanners (the optical system needs to be "perfect" in a smaller area). It also has some of the advantages of the stepper systems (4X or 5X reduction can be used to simplify mask fabrication, and the total field printed at each exposure is much smaller than the wafer, eliminating wafer size as a major issue and simplifying lens design). However, the mechanical system does need to scan the mask 4X or 5X times faster than the wafer and keep them synchronized. These hybrid systems also allow local alignment of the mask pattern with the wafer at each exposure field, which is one of the inherent advantages of stepper systems. At the same time, however, these systems are very complex and very expensive because they have some of the disadvantages of each type of system (the synchronized mechanical motion of scanners, and the costly refractive lens design of steppers). Nevertheless, it is likely that these hybrid systems will find widespread use through the next several technology generations.

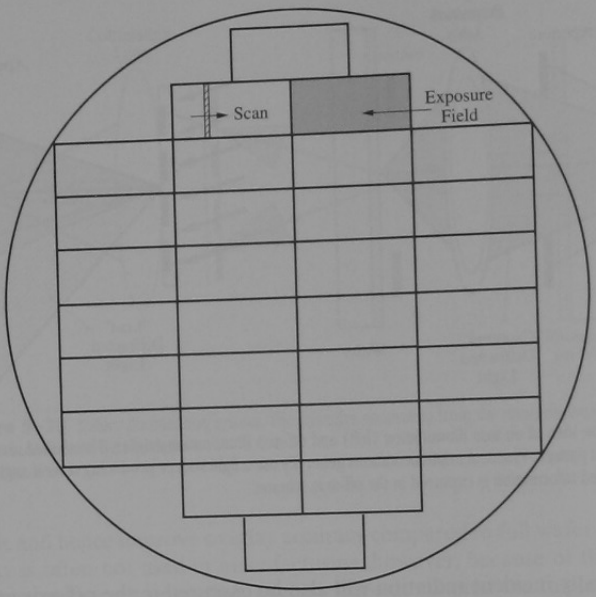


Figure 5–30 Step and scan system. Stepping accomplishes the major moves from one exposure field to another. Within each exposure field, the mask pattern is scanned across the field.

5.3.2 Photoresists

The actual details associated with photoresist exposure and processing are, not surprisingly, considerably more complicated than the qualitative description given earlier might indicate. The additional details that we will consider here are generally designed to make the whole process more manufacturable. Figure 5–31 illustrates a typical process flow associated with photolithography. The specific numbers refer to typical conditions for DNQ g-line or i-line resists. DUV resists generally utilize a similar process flow with some changes in the process details.

The first step in the lithography process is ensuring that the resist will adhere well to the wafer. Depending on the stage in the process flow, this may involve one or more operations. Normally the wafer is clean just before resist application since resist is commonly deposited on wafers just after thin film deposition or just after some other high-temperature process step. If this is not the case, the wafer may need to be chemically cleaned, using the procedures described in Chapter 4. It may also be necessary to heat the wafer to several hundred °C to drive off any water vapor on the surface. These steps are represented by the first box in Figure 5–31.

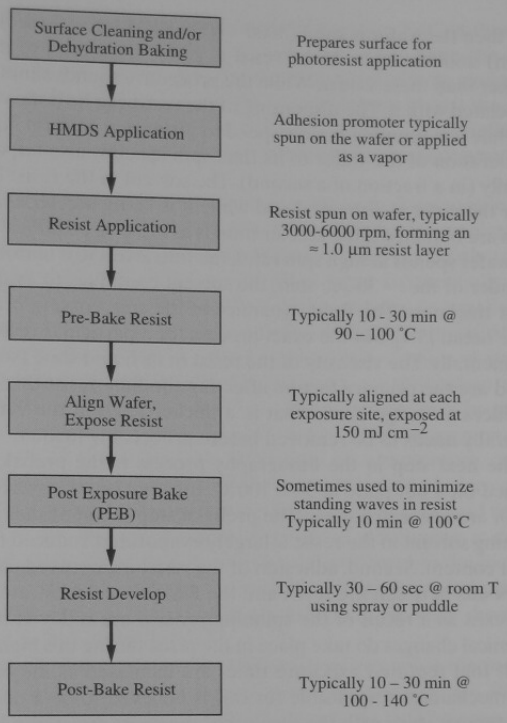


Figure 5–31 Typical photoresist process flow for DNQ g-line and i-line positive resists.

Even with a clean, dry surface, adhesion of resists to silicon ICs may not be as good as desired and so it is common practice to use an adhesion promoter. Hexamethyldisilane (HMDS) is the most common substance used for this purpose. HMDS can be applied to the wafer in liquid form at room temperature by spinning under conditions similar to resist spinning (3000 – 6000 RPM for ≈ 30 sec). However, in most cases it is applied in vapor form with the HMDS being introduced as a vapor into a chamber containing the wafers since this more easily produces the desired single monolayer on the wafers. In either case, one end of the HMDS molecule bonds readily with SiO_2 surfaces and the other bonds with the resist. Thus the HMDS acts an adhesion promoter. For surfaces other than SiO_2, other adhesion treatments may be used in addition to or instead of HMDS (Chapter 6 of [5.4]).

The next step is spinning the resist itself onto the wafer and this should normally be done immediately after the HMDS application. The resist is dispensed onto the wafer

and then the wafer is spun (3000 – 6000 RPM for ≈ 30 sec) to produce a thin (≈ 0.6 – 1-μm) uniform layer. If the resist is going to be used to mask an implant, it may be thicker than these values. While the procedure sounds simple, there are some subtleties associated with it. The dispensing of the resist can be done while the wafer is stationary or while it is spinning at slow speed to produce a uniform liquid layer on the wafer. The acceleration of the wafer to its final spin speed is also important and is typically done rapidly (in a fraction of a second). The solvent in the resist begins to evaporate rapidly after the resist is dispensed and while it is being accelerated. Generally more uniform films are obtained if the acceleration is as rapid as possible. During the first few seconds the wafer spends at high spin rates, the film levels to a uniform thickness. During the remainder of the ≈ 30-sec spin, the solvent continues to evaporate to produce the final resist thickness. The fluid dynamics of the spin coating process have been studied in some detail [5.4], but the exact process for a particular resist is usually determined experimentally. The viscosity of the resist in its liquid state (solvent content) and the spin speed are the primary factors affecting the final resist thickness. The spinning process produces an "edge bead," that is, a thicker resist at the very edge of the wafer, which generally needs to be removed before proceeding further.

The next step in the lithography process is the prebake. This is normally accomplished on a hot plate at 90 – 100°C. Infrared or microwave heating can also be used, which are faster processes. The prebake step accomplishes several things. First, the remaining solvent in the resist is largely evaporated, reduced from ≈ 25% to ≈ 5% of the resist content. Second, adhesion of the resist is improved since the heating strengthens the bonds between the resist and the HMDS and substrate. Finally, stresses present in the resist as a result of the spinning process are relieved through thermal relaxation. Chemical changes do take place in the resist during this high temperature bake with the result that required exposure times are increased as the bake temperature increases. The mechanism responsible for this is believed to be a decomposition of the *PAC* at high temperature with the result that the resist sensitivity is degraded. We will discuss these effects more quantitatively in Section 5.5.

The exposure process creates a latent image in the resist that can be developed later. Most resists exhibit reciprocity, which means that light intensity and exposure time can be directly traded off with each other. Thus increases in the exposure system aerial intensity directly reduce exposure times. The required exposure time is also affected by the prebake thermal cycle as was just discussed and by the thickness of the resist. All of these parameters must therefore be carefully controlled. Exposure doses are normally designed to be well above Q_f (Figure 5–20) since this produces latent images in the resist that have the sharpest edges (limited of course by the quality of the aerial image). For typical DNQ resists, this corresponds to a dose > 100 mJ cm^{-2}. For DUV resists, which use chemical amplification, the dose is typically 20 – 40 mJ cm^{-2}. Detailed models of the exposure process have been developed, including the effects of the quality of the aerial image, standing waves in the resist, and so forth. We will consider these in Section 5.5.

The next step illustrated in Figure 5–31 is a Postexposure Bake (PEB). In g-line and i-line resists, this step is sometimes performed before development of the resist latent image, in order to minimize standing wave effects in the resist. The mechanism is

straightforward. At elevated temperatures, the *PAC* in these resists can diffuse. If the temperature and time are well controlled in this bake, the *PAC* molecules can diffuse far enough to "smear out" the standing wave effects along the edge of the resist features, but not diffuse far enough to significantly distort the image features themselves. A typical bake cycle might be $\approx$ 10 min at 100°C. If antireflective coatings are used under the resist, this postexposure bake may not be necessary because the standing wave problem is less severe.

In DUV resists, the PEB is a necessary and critical step in the process. This is the step in which the *PAG* reacts with the polymer chain to complete the exposure process. The time and especially the temperature must be very tightly controlled because chemical reaction rates and diffusion typically depend exponentially on temperature.

DNQ resists are developed in basic solutions (normally TMAH—tetramethyl ammonium hydroxide—solutions diluted with H_2O, but NaOH or KOH solutions can also be used). The developer can be applied in a number of ways. The wafers may be immersed in the developer, the developer may be sprayed on a batch of wafers or on a single wafer at a time, or a puddle of the developer may be placed on the wafer. In each case, rinsing the wafers with H_2O stops the developing process. The rate at which developing proceeds is highly dependent on temperature, on developer concentration, and on all the exposure and bake procedures used before developing. In DNQ resists, the developing proceeds by dissolving the carboxylic acid that results from the exposed *PAC* in the alkaline solution. The rate of developing is dependent on the local carboxylic acid concentration, which is proportional to the local exposure intensity in the resist. In positive working DUV resists, developing takes place in a similar fashion since the unblocked polymer chains are soluble in the basic developer. We will consider detailed models in Section 5.5.

The final step in the photolithography process is the postbake. This step is done at higher temperature than the earlier bakes (typically 10 – 30 min at 100 – 140°C), and is designed to harden the resist and improve its etch resistance. This bake process causes the resist to flow slightly and hence may also modify edge profiles. Any remaining solvents in the resist are driven off by this bake and the adhesion of the resist to the underlying substrate is also improved. In some cases when the resist must withstand a particularly harsh etching process or a high current ion implantation which may raise its temperature considerably, a further hardening and crosslinking of the resist with a blanket deep UV exposure at elevated temperatures can improve the robustness of the resist layer.

In modern lithography processes, many of the steps we have described above are accomplished in a single integrated machine known as a wafer track system. These machines cost > $1 million and are usually tightly integrated with the exposure tool so that easy wafer transfer can occur between the two systems.

5.4 Measurement Methods

Measurement issues associated with lithography can be divided into several areas. The first is the mask itself. Here we need to know the dimensions of the features on the

mask and to verify that they correspond to the intended design. It is generally also important to ensure that there are no defects on the mask which are large enough to print on the wafer during the exposure process.

As we have seen in our earlier discussion, it is convenient to divide the exposure process into the optical system that produces the aerial image, and the resist processing which transfers the aerial image into a resist profile. The aerial image itself is not normally directly measurable, so that measurement issues associated with lithography normally focus on the resist pattern after it is developed. This is the second major area where measurement methods are needed. In many process steps, of course, the resist pattern is transferred into underlying thin films or into the silicon substrate itself by etching and in these cases, there is also a need to measure the etched pattern.

The final issue associated with the lithographic process is alignment of patterns with respect to underlying features on the wafer (earlier mask layers). As we saw earlier in this chapter, generally layers must be aligned with a tolerance on the order of one-fifth to one-third of the minimum feature size. We will discuss a number of these issues in this section. An excellent reference on many of these topics is [5.15].

5.4.1 Measurement of Mask Features and Defects

Virtually all exposure systems used today in high-volume manufacturing are reducing steppers or reducing step and scan projection systems. This means that the mask contains the features for only a few die and this pattern is exposed multiple times on each wafer. If there is a defect on such a mask, it will affect a large fraction of all chips on every wafer the mask is used to print. This would result in intolerable yield losses and so it is absolutely crucial that the mask be "perfect." What this really means is that there are no printable defects on the mask. Since the mask has 4X to 5X larger feature sizes than the wafer patterns, mask defects below some critical size will not print on the wafer. Nevertheless, the two issues with respect to the mask are the following: Does the mask pattern correspond to the design for that mask level? And are there any defects large enough to print on the wafer?

While mask inspection was accomplished many years ago simply by using a microscope, such an approach is unworkable today because of the complexity of today's chips. As a result, highly automated systems have been developed to accomplish this inspection. The principle of operation is illustrated in Figure 5–32. Light is passed through the mask and collected by an imaging system. Generally a solid-state image sensor is used to detect the pattern of the transmitted light. This information can be compared either with the design database used to generate the mask, or with a second identical mask pattern, if the mask contains the patterns for more than one chip. Scanning the inspection system over the mask thus provides full information on defects due either to errors in transmitting data from the database or due to random defects on the mask. If the mask incorporates Optical Proximity Correction (OPC) or advanced techniques like phase shifting, the inspection process is often more difficult because these techniques introduce structures on the mask that may be smaller than the minimum feature size and therefore might be interpreted as "defects" by the inspection system.

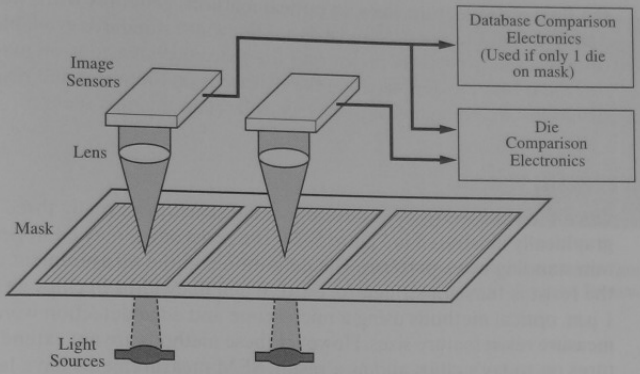


Figure 5–32 Mask inspection system. Such systems operate by comparing the feature information on the mask either with the original design database or with an identical feature on an adjacent mask site. In this example three identical die are shown on the mask.

If defects are detected by this approach, they can often be corrected. Opaque defects, in which chrome is present on the mask in areas where it should not be, can be corrected with lasers or ion beams which are focused on the defect area and which evaporate the excess chrome. Clear defects, in which chrome is not present where it should be, are often more difficult to correct because they require a deposition of an opaque material to cover the clear area. Laser-assisted deposition from a chromium bearing gas which is introduced above the mask can be used, as can ion beam deposition.

The final crucial issue with respect to the mask is the actual size of the features on the mask. There are a number of issues associated with e-beam mask making that can lead to disparities in feature size between the design database and the mask. The spot size of the e-beam is finite (usually 0.125 to 0.5 μm) and this can be important in writing very small features. Also, when small features are being written, "proximity effects" can be an issue when such features are closely spaced. Proximity effects arise from electron backscattering in the resist on the mask and can result in pattern distortion.

Many masks today are made with laser-based systems instead of e-beam systems because the laser-based systems are less expensive. However, laser-based systems have the same kinds of resolution limits as do the steppers used in projecting images onto the wafer. Thus the features printed on the mask itself can suffer from rounded corners and other diffraction effects just as is the case in aerial images projected onto the wafer.

The feature sizes on the mask can be measured with a system very similar to that shown in Figure 5–32. Alternatively, a video camera and microscope can be used on the top side of the mask. In either case, edge detection techniques are used on the signal that is received to determine the feature dimension. This is a relatively straightforward procedure on masks because the thin chrome layer on masks provides a high contrast image with good edge definition. In addition, the linewidths are 4X to 5X larger than

the final wafer feature sizes, so optical methods generally work well for mask features. The National Bureau of Standards has linewidth standards available that can be used as references in accurately measuring the feature dimensions on masks. Linewidth measurement is not as simple when the feature is on the wafer and resist or other thin film structures are being measured, as we will see in the next section.

5.4.2 Measurement of Resist Patterns

Once the exposed resist is developed, the structure is really three dimensional, as was graphically illustrated in Figure 5–24. The resist edges may be sloped and they may contain standing wave patterns. The definition and the measurement of the "linewidth" in the resist is thus not straightforward. When linewidths in chips were larger than about 1 μm, optical methods using a microscope and edge detection were commonly used to measure resist feature sizes. However these methods are not extendible to the small features on today's chips and as a result, SEM measurements have largely replaced optical methods for accurate linewidth measurements.

We discussed the basic operation of electron microscopes in Chapters 3 and 4 (Sections 3.4 and 4.4). SEM methods have the potential for very high resolution. The electron wavelength is < 1 nm at typical accelerating voltages, so diffraction effects are negligible. SEMs designed for inspection and metrology applications usually have to operate nondestructively because the wafers that are used for the inspections are product wafers. Thus ideally the SEM operates in an in-line fashion as part of the manufacturing line. Such tools are widely available today. They generally operate with a small beam size (≈ 10 nm) and at fairly low voltage (a few keV) to minimize damage to the wafer structures, to minimize charging problems, and to allow imaging of the resist surface. They generally are designed to handle full 6″ or 8″ wafers. Low-voltage operation also minimizes the depth to which the electrons penetrate in the resist or other thin films and thus confines the interactions of the electrons to the near surface region. This improves the ability of the SEM to image surface features and topography.

Figure 5–33 shows examples of SEM images of a resist pattern after development. The resolution capability of the technique is apparent, but these photos and Figure 5–24 also illustrate some of the difficulties in "measuring" linewidths in such structures. The possibility of sloped edges, standing wave patterns, and variations in the linewidth along photoresist lines all create difficulties in defining exactly what is meant by the linewidth. Generally, the linewidth is defined as the width of the photoresist material at a specific height above the resist-substrate interface. Often sophisticated algorithms are used to translate an SEM line scan into a "number" representing the linewidth.

5.4.3 Measurement of Etched Features

Photoresist patterns are often transferred into underlying thin films by etching, a process described in detail in Chapter 10. The quality of this pattern transfer process is obviously just as important as the quality of the photoresist pattern itself. The same

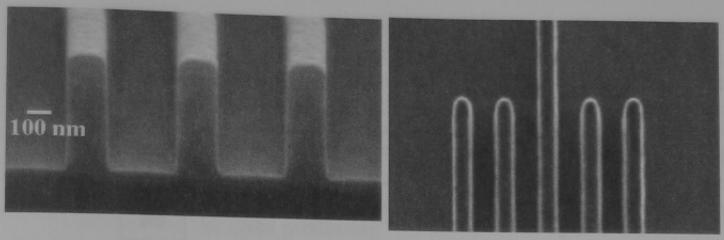


Figure 5–33 Cross section (left) and top (right) view SEM images of developed photoresist features showing well-developed lines with sub 0.25-μm lines and spaces. Courtesy of A. Vladar and P. Rissman, Hewlett Packard.

linewidth measuring methods, primarily based on SEMs, are used to measure the etched patterns in oxide, polysilicon, Al, or other materials.

Unlike the photoresist patterns, which are removed from the wafer after they are used, the etched thin film patterns remain as part of the final chip. Because of this, it is possible in many cases to make measurements on these thin film layers after wafer processing is complete. In addition to SEM methods, it is often convenient to make such measurements electrically during wafer testing. A significant amount of work has been done in recent years to develop electrical test structures which measure parameters like linewidth and alignment accuracy, in addition to the more usual device electrical parameters [5.15 – 5.19]. Complete test chips have been developed which provide detailed information about the process. These chips are very useful during the development stages of new processes. Specific test structures are also sometimes incorporated onto product chips once a process is in manufacturing in order to provide ongoing information about manufacturing tolerances. Usually these test structures are placed in the scribe lines between chips so they do not consume chip area.

Figure 5–34 illustrates the basic idea behind many of these test structures. The overall structure is assumed to be a conducting material (polysilicon, silicide, aluminum, etc.). The right-hand part of this structure (pads 3–6) is a van der Pauw structure designed to extract the sheet resistance of the material making up the test structure. The geometry is chosen to define one square of the material (labeled with the ρ_s symbol in the figure). If a current I_{5-6} is forced between terminals 5 and 6, and the voltage V_{3-4} is measured between terminals 3 and 4, then the sheet resistance is given simply by [see Eq. (3.10) in Chapter 3].

$$\rho_S = \frac{\pi}{\ln(2)} \frac{V_{3-4}}{I_{5-6}} \tag{5.17}$$

Once the sheet resistance is measured, then the linewidth W can be extracted from the rest of the structure through an additional electrical measurement. In this case, a current I_{1-5} is forced between terminals 1 and 5, and a voltage V_{2-3} is measured between terminals 2 and 3. The linewidth W is then given by

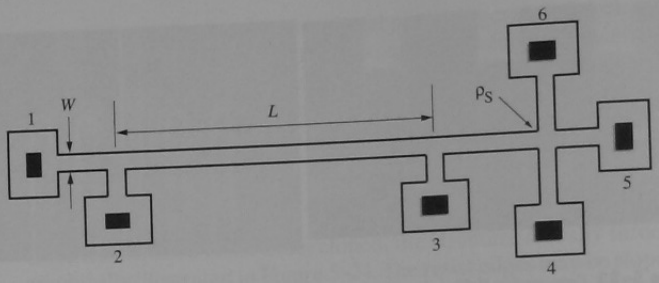


Figure 5–34 Test structure designed to electrically extract sheet resistance and linewidth *W*.

$$W = \rho_S L \frac{I_{1-5}}{V_{2-3}} \tag{5.18}$$

It should be noted that the linewidth extracted from test structures such as this may be different than the linewidth measured optically or with an SEM. This is because the electrical linewidth really measures the effective cross sectional area through which current flows in the material. If the edges of the material are sloped, or if the surfaces affect the conductivity, the electrical width may be slightly different than the physical width extracted from an image of the line.

Finally, modifications to the basic test structure shown in Figure 5–34 can allow similar structures to extract information about alignment between levels on a chip. In this case, parts of the structure are fabricated in different levels of the circuit and measurements of resistances in the completed test structure provide the alignment information [5.17, 5.18].

5.5 Models and Simulation

Lithography simulation relies on two principal areas of science. The first of these is optics, which provides a mathematical description of the behavior of light in the exposure systems used in modern lithographic tools. The second is chemistry which provides the tools to treat exposure, baking, and developing of the resists that are used to convert the aerial image created by the exposure system into a three-dimensional replica of the mask patterns. Many of the basic ideas and models that optics provides about light propagation and diffraction effects have their origins in work done more than 100 years ago by Maxwell, Kirchoff, and others. And of course the basic ideas from chemistry regarding the rates of chemical reactions, catalysts, and the like are equally old ideas.

Applying these concepts to lithography, however, is a relatively recent venture, dating back only to the mid 1970s. The pioneering work in this area began at IBM and resulted in a series of widely referenced papers written by Dill and his co-workers [5.20

– 5.23]. Following that work, a number of other groups began work on lithography simulation [5.24, 5.25]. Within the past 10 years, a number of simulation tools for optical lithography have become commercially available. Representative examples include PROLITH (Finle Technologies) [5.3], DEPICT (Avant!) [5.26], and ATHENA (Silvaco) [5.27]. The basic models used in each of these simulation tools are similar. In the following sections we will describe these models in a general way and show specific examples of the use of these simulators in optical lithography. As we did earlier in this chapter, we will divide the discussion into the formation of the aerial image (exposure system) and the resist processing.

5.5.1 Wafer Exposure Systems

Since projection exposure systems dominate industrial practice today, most of the simulation tools that are currently available apply to these kinds of systems. They thus model far field or Fraunhofer diffraction. Some simulators do have the ability to simulate contact or proximity printing (PROLITH, for example), but because these systems are not used for VLSI manufacturing, we will not consider near field (Fresnel diffraction) simulation tools in this book. The interested reader is referred to [5.25] for issues related to simulation of proximity or contact lithography simulation.

For the purposes of this discussion, we will consider a generic projection lithography system as illustrated in Figure 5–35. Mathematically modeling such a system requires a mathematical description of how light behaves. We begin with a basic description of light waves and their propagation.

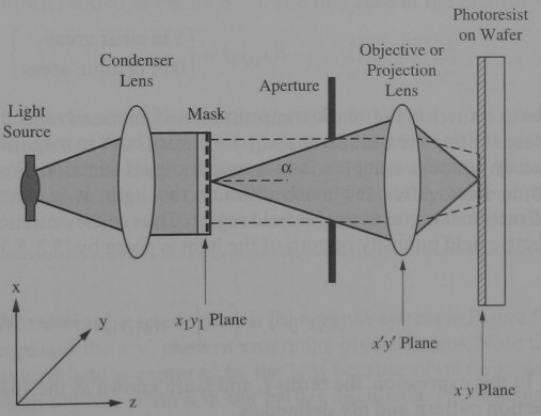


Figure 5–35 Generic projection lithography system. The shaded area to the right of the mask illustrates the portion of the diffraction pattern that is collected by the projection lens and used to form the aerial image of the mask feature.

Light travels as an electromagnetic wave. The electric and magnetic fields are perpendicular to each other and both are perpendicular to the direction of wave propagation. However, the materials that photons interact with in lithography are generally nonmagnetic, so we can treat the light waves in such systems simply in terms of a traveling electric field ε. Thus a general description of a propagating monochromatic light wave at a point W in space is simply

$$\varepsilon(W,t) = C(W)\cos(\omega t + \phi(t)) \tag{5.19}$$

where C is the amplitude, ω is the frequency and ϕ is the phase. We can also write this in the form of a complex exponential function.

$$\varepsilon(W,t) = Re\{U(W)e^{-j\omega t}\} \text{ where } U(W) = C(W)e^{-j\phi(W)} \tag{5.20}$$

The generic projection lithography system in Figure 5–35 consists of several components, each of which we have discussed in earlier sections of this chapter. The light source and condenser lens form the illumination system which needs to deliver light to the mask with the specified intensity, uniformity, spectral characteristics, and spatial coherency. Once the light passes through the mask, diffraction effects come into play. The objective or projection lens is required to collect as much of the diffracted light as possible and focus it onto the resist layer on the wafer. Generally demagnification also takes place through the objective lens.

We will consider the mask to be transparent in some areas and completely opaque in others. For the chrome masks that are typically used today, this is a very good approximation since the mask contrast is very high. Thus in the x_1y_1 plane, we can represent the transmission of the mask with a simple digital function.

$$t(x_1,y_1) = \begin{Bmatrix} 1 \text{ in clear areas} \\ 0 \text{ in opaque areas} \end{Bmatrix} \tag{5.21}$$

where $t(x_1,y_1)$ is the mask transmittance. If more advanced mask techniques such as phase shifting are used, then $t(x_1,y_1)$ will vary both in magnitude and phase and will obviously be more complex than a simple digital function. We will consider such an example later. After the mask diffracts the light, it is described by the Fraunhofer diffraction integral in the far field region. Thus at the entrance to the objective lens, the electric field intensity pattern of the light is given by [5.2, 5.3]

$$\varepsilon(x',y') = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} t(x_1,y_1)e^{-2\pi j(f_x x + f_y y)}dxdy \tag{5.22}$$

In this expression, the terms f_x and f_y are known as the spatial frequencies of the diffraction pattern and are defined as

$$f_x = \frac{x'}{z\lambda} \text{ and } f_y = \frac{y'}{z\lambda} \tag{5.23}$$

where z is the distance from the mask to the lens. Eq. (5.22) is simply the Fourier transform of the mask pattern. The light intensity pattern entering the objective lens can thus be calculated using well-known methods from the field of Fourier optics. In the shorthand of the Fourier transform, we can rewrite Eq. (5.22) as

$$\varepsilon(f_x,f_y) = F\{t(x_1,y_1)\} \tag{5.24}$$

where F represents the Fourier transform. f_x and f_y are simply scaled spatial coordinates in the x', y' plane. The intensity distribution of the light is the square of the magnitude of the electric field, so that

$$I(f_x,f_y) = |\varepsilon(f_x,f_y)|^2 = |F\{t(x_1,y_1)\}|^2 \tag{5.25}$$

Example Consider a mask pattern, which is simply a long rectangular slit of width w (Figure 5–36). Find the resulting optical intensity pattern in the far field (Fraunhofer) region.

Answer The Fourier transform of the function $t(x)$ is found in standard textbooks [5.2] and is given by

$$\varepsilon(x') = F\{t(x)\} = \frac{\sin(\pi w f_x)}{\pi f_x} \tag{5.26}$$

The resulting optical intensity distribution is the square of this function and is also shown plotted in Figure 5–36. The first zero of the intensity occurs at

$$f_x = \frac{1}{w} \text{ or } x' = \frac{z\lambda}{w} \tag{5.27}$$

Thus the image of the slit spreads out the further away we move from the mask (z increases). The image also spreads out more when longer wavelengths of light are used. Both of these results are to be expected based on our previous discussion of diffraction. Perhaps counterintuitively, the image spreads out more the smaller the slit is (as w decreases). This was discussed in connection with Figure 5–5 earlier.

We now return to our generic lithography system in Figure 5–35. When the diffracted light passes the $x'y'$ plane, it enters the objective lens. Note that only a portion of the diffracted light is captured by the lens because of its finite size. We characterized this through the $NA = \sin\alpha$ in our earlier discussion. The function of the objective lens is to reconstruct the diffraction pattern and to focus it on the wafer. Since the intensity pattern of the light at the $x'y'$ plane is simply the Fourier transform of the mask pattern, the objective lens effectively needs to perform an inverse Fourier transform. This is in

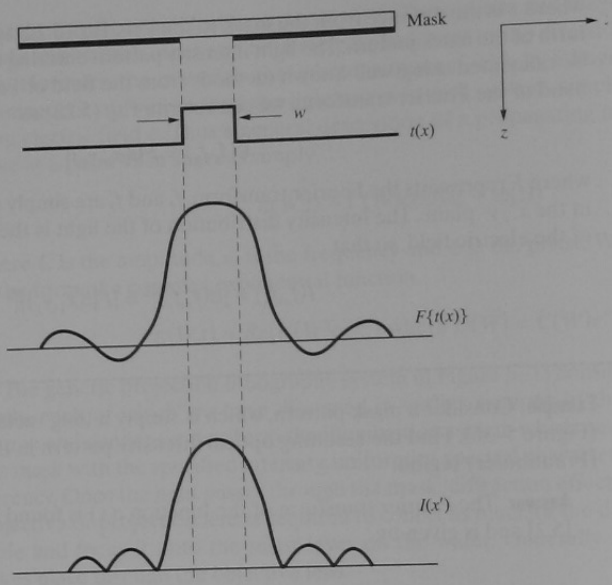


Figure 5–36 Rectangular aperture in a mask and the resulting far field Fraunhofer diffraction pattern and optical intensity pattern.

fact exactly what spherical lenses do. The lens can perform this operation only on the portion of the diffracted light that enters the lens and in this sense we speak of the image that the lens forms as being "diffraction limited."

We can characterize the portion of the diffracted light that the objective lens captures with a "pupil function P" defined as

$$P(f_x f_y) = \begin{Bmatrix} 1 \text{ if } \sqrt{f_x^2 + f_y^2} < NA/\lambda \\ 0 \text{ if } \sqrt{f_x^2 + f_y^2} > NA/\lambda \end{Bmatrix} \tag{5.28}$$

since for small angles the spatial frequencies f_x and f_y are given by,

$$f_x = \frac{x'}{z\lambda} \cong \frac{\sin\alpha}{\lambda} = \frac{NA}{\lambda} \tag{5.29}$$

P may be limited by the physical size of the objective lens, or by an aperture placed in front of the lens. Either way, the information passing through the objective lens is effectively low pass filtered by the pupil function P. Spatial frequencies higher than NA/λ are simply not passed on to the image on the wafer. The analogy in electric circuits, a sharp pulse passed through a low-pass filter which then ends up with rounded edges,

may help to explain the optical degradation in the aerial image caused by P. A sharply defined object on the mask is imaged on the wafer with "smeared" edges, and sharp corners become rounded.

The objective lens now performs the inverse Fourier transform function on the portion of the diffracted light that it captures. Thus at the image plane on the wafer (resist surface), the electric field associated with the light pattern is given by

$$\varepsilon(x,y) = F^{-1}\{\varepsilon(f_x,f_y)P(f_x,f_y)\} = F^{-1}\{F\{t(x_1,y_1)\}P(f_x,f_y)\} \tag{5.30}$$

Finally, the intensity of the light distribution at the resist surface $I_i(x,y)$ is simply given by

$$I_i(x,y) = |\varepsilon(x,y)|^2 \tag{5.31}$$

Eq. (5.31) now represents the aerial image produced by the exposure system. Figure 5–37 summarizes the mathematical models that are used to describe the exposure system.

In real projection lithography systems, there are some additional issues that must be considered in calculating the aerial image. We discussed these issues qualitatively earlier in this chapter, but we now need to quantify them to enable simulation in real systems. The issues which are often important in real imaging systems include off-axis illumination, the use of a partially coherent light source, aberrations in the lens system, depth of focus issues, and the finite thickness of the photoresist, which means that the aerial image must be determined over some volume, not just at the resist surface. We will describe in the following paragraphs the modifications we need to make to Eq. (5.31) in order to deal with these issues.

The first issues are off-axis illumination and partially coherent light sources. In both cases, the light that enters the mask does not have a direction normal to the mask. The result of this (Figure 5–30) is that the position of the diffraction pattern is shifted in the x' y' plane at the entrance to the objective lens. If the light enters the mask at an angle

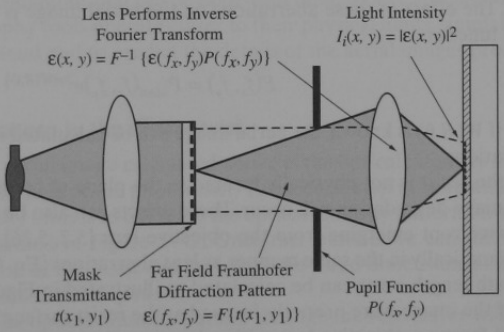


Figure 5–37 Mathematical models used to describe the projection exposure system at each point in space.

θ, then in terms of the spatial frequencies f_x and f_y, the diffraction pattern is shifted by an amount sinθ/λ. The result then is that the electric field associated with the light pattern in the image plane is expressed as [5.3]

$$\varepsilon(x,y,f'_x,f'_y) = F^{-1}\{\varepsilon(f_x - f'_x, f_y - f'_y)P(f_x,f_y)\} \tag{5.32}$$

and

$$I_i(x,y,f'_x,f'_y) = |\varepsilon(x,y,f'_x,f'_y)|^2 \tag{5.33}$$

In the case of a partially coherent source, light enters the mask from a variety of angles. This is usually treated by superposition [5.27]. The source is divided up into individual point sources and the intensities of the light at the image plane due to each point source are added. Expressed mathematically, if $S(f'_x, f'_y)$ is the function describing the source intensity versus position or angle, then [5.3]

$$I_{total}(x,y) = \frac{\iint_{source} I(x,y,f'_x,f'_y)S(f'_x,f'_y)df'_x,df'_y}{\iint_{source} S(f'_x,f'_y)df'_x,df'_y} \tag{5.34}$$

where I_{total} is the total light flux falling on the resist surface. Eq. (5.34) thus allows the calculation of the aerial image for an arbitrary source illumination.

Aberrations or defects are always present in practical lens systems. They arise from manufacturing tolerances, mechanical changes in a system during use (lenses out of alignment with each other), and from the fact that it is impossible to design a perfect lens which is completely free of aberrations. While such issues could be modeled in a number of ways, the most common method is to treat the wavefront emerging from the lens as having an error in phase or path difference, compared to an ideal aberration free lens. The error in a practical lens will be a function of position along the wavefront and is usually expressed in polar coordinates since lenses are generally circular. The phase error is often expressed as a Zernike polynomial $W(R,\theta)$, which may have many terms [5.3]. $W(R,\theta)$ is the path length difference between an ideal lens and one with aberrations. The effect of these aberrations on the aerial image is included by modifying the pupil function

$$P(f_x,f_y) = P_{ideal}(f_x,f_y)e^{2\pi jW(R,\Theta)} \tag{5.35}$$

Thus if $W(R,\theta)$ is known, the aerial image can be calculated including the effects of lens aberrations.

If the resist is not physically located in the plane of focus of the objective lens, further image degradation will occur. These effects can also be treated as a phase error in the wavefront emerging from the objective lens [5.2, 5.26]. Thus they can be treated mathematically in the same manner as lens aberrations [Eq. (5.35)]. In the defocus case, the path length error can be calculated as illustrated in Figure 5–38. Here we imagine that if the image were properly focused at the resist surface, the light would have originated from an objective lens with a larger radius of curvature (dashed lines) rather than from the actual lens. The Optical Path Difference (OPD) is zero at the center of the lens and maximum at the outside radius of the lens. If $R >> \delta$, then

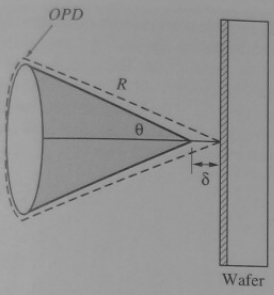


Figure 5–38 Representation of defocusing (δ) as an error in optical path length (optical path difference = *OPD*).

$$OPD \equiv \delta(1 - \cos\Theta) \tag{5.36}$$

This expression can then be used in Eq. (5.35) to calculate the effect on the aerial image due to defocusing. Notice that the *OPD* is greatest at the edges of the lens where the higher order diffracted light passes through. This is the light that contains the detailed information about the mask shape, so defocusing will clearly degrade the quality of the aerial image.

There are a number of commercially available simulation tools that implement the models we have just described. Generally the mask is described either as a combination of rectangles or as an arbitrary geometrical structure. In some cases, mask information directly from a design database can be downloaded into the lithography simulation tool. However it is accomplished, a transmittance function $t(x_1, y_1)$ must be specified by the user. The exposure tool is described in terms of its *NA*, the illumination wavelength λ the degree of coherence of the source, any defocusing present, and any aberrations in the lens. The simulation tools then calculate the aerial image using the equations from Fourier optics that we have described. Figures 5–39 to 5–41 show examples of the kinds of simulations that can be done with these tools. Especially as linewidths shrink and lithography tools operate closer to their physical limits, simulators become essential to understand and to predict the details of the aerial images produced during exposure.

5.5.2 Optical Intensity Pattern in the Photoresist

The aerial image calculated above is the optical intensity in the air immediately above the surface of the photoresist. The second step in simulating the performance of lithography tools is to calculate the optical intensity pattern through the photoresist layer as illustrated in Figure 5–42. Once this is done, we can then calculate the chemical response of the resist and therefore the three-dimensional structure of the resist after its pattern is developed.

The optical intensity pattern through the resist layer will be different than the aerial image for a number of reasons. First, the resist has a finite thickness and hence the image produced by the optical system cannot be in perfect focus everywhere in the resist. Another potential issue is the presence of standing waves in the resist, caused by light

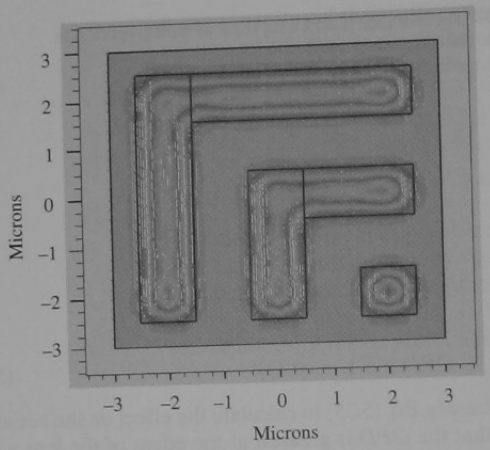


Figure 5–39 Example of aerial image calculation using the ATHENA simulator by Silvaco. The shading corresponds to optical intensity in the aerial image. The black borders correspond to the mask image that is being printed. The exposure system simulated had a $NA = 0.43$, partially coherent g-line illumination ($\lambda = 436$ nm) and no other aberrations or defocusing. The minimum feature size is 1 μm.

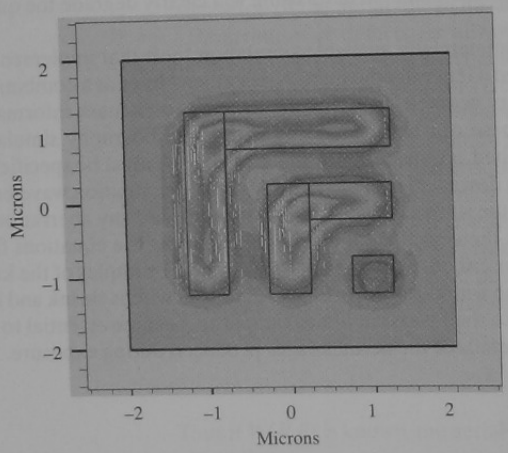


Figure 5–40 Same example as in Figure 5–39 except that the feature size has been reduced to 0.5 μm. Note the much poorer image quality.

reflections from underlying structures. We illustrated this problem in Figures 5–23 and 5–24. In addition, as we also discussed in connection with Figure 5–22, how absorption of light in typical DNQ resists changes with exposure time due to bleaching. This results in a light intensity that varies with depth and with time in the resist. Finally, in high *NA* exposure systems, the light striking the resist is not all vertically incident. This complicates the calculation of any standing wave patterns that may be present in the resist and generally requires that an integration over the total angle of the incoming light be performed. This procedure can also include the effects of partial coherency in the light source because this also results in light striking the resist at nonnormal angles.

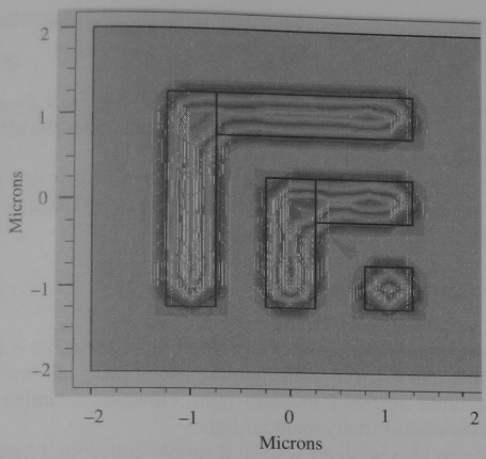


Figure 5–41 Same example as in Figure 5–40 except that the illumination wavelength has now been changed to i-line illumination ($\lambda = 365$ nm) and the *NA* has been increased to 0.5. Note the improvement in image quality resulting from the shorter exposure wavelength and higher *NA* system.

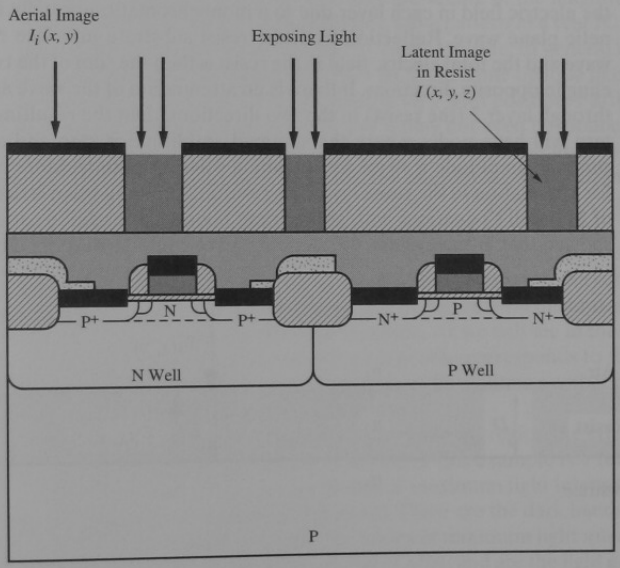


Figure 5–42 Translation of the aerial image into light intensity through the resist layer. This is the latent image produced in the photoresist by the exposing photons. Standing waves, defocusing, and bleaching effects all contribute to make $I(x,y,z)$ different from $I_i(x,y)$.

The simplest way to deal with many of these issues is to modify Eq. (5.31) as follows:

$$I(x,y,z) = I_i(x,y)I_r(x,y,z) \tag{5.37}$$

$I_i(x,y)$ is again the light intensity pattern incident on the top surface of the resist [aerial image given by Eq. (5.31)] and $I_r(x,y,z)$ is a correction factor which includes the effects of defocus, standing waves, and resist bleaching.

Defocusing of the image through the resist layer can be dealt with through Eqs. (5.35) and (5.36). If the aerial image is in focus at the top surface of the resist, then the defocusing that occurs through the resist thickness is given by

$$\delta(z) = \frac{z}{n} \tag{5.38}$$

where z is the depth into the resist and n is the index of refraction of the resist material. This factor can then be included in Eq. (5.37) through the use of the pupil function [Eq. (5.35)]. Eq. (5.38) does assume normally incident light. However a similar approach can be used to account for nonnormally incident light.

Standing waves in the resist can also be dealt with in a simple conceptual fashion. Consider Figure 5–43 in which n is the index of refraction of each of the layers, and ε is the electric field in each layer due to a monochromatic, normally incident electromagnetic plane wave. Reflection from the resist substrate interface results in a reflected wave and the total electric field in the resist is then the sum of the two plane waves traveling in opposite directions. If there is no attenuation of the wave amplitude as it passes through layer 2 (the resist) in the two directions, then the resulting standing wave pattern will have nodes where the summed amplitude is zero and antinodes where the summed amplitude is twice the incoming wave amplitude. More generally, accounting for attenuation in the resist layer, the resulting standing wave is described by a solution of Maxwell's equations with appropriate boundary conditions at each interface. Mack showed that in this system, the standing wave could be described by [5.25]

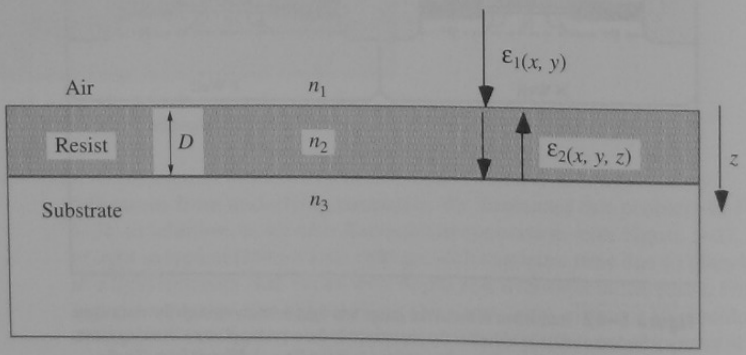


Figure 5–43 Simple example of the formation of standing waves in a resist layer.

$$\varepsilon_2(x,y,z) = \varepsilon_1(x,y)\frac{\tau_{12}(e^{-2\pi j n_2 z/\lambda} + \rho_{23}\tau_D^2 e^{2\pi j n_2 z/\lambda})}{1 + \rho_{12}\rho_{23}\tau_D^2} \tag{5.39}$$

where $\varepsilon_1(x,y)$ is the incident electromagnetic plane wave, $\varepsilon_2(x,y)$ is the plane wave in the resist, n_j is the complex index of refraction of layer j, λ is the wavelength of the incident light, and

$$\rho_{ij} = \frac{n_i - n_j}{n_i + n_j}, \text{ the reflection coefficient at interface } ij$$

$$\tau_{ij} = \frac{2n_i}{n_i + n_j}, \text{ the transmission coefficient across interface } ij$$

$$\tau_D = \exp(-jk_2 D), \text{ the transmittance of the resist film}$$

$$k_j = \frac{2\pi n_j}{\lambda}, \text{ the propagation constant of layer } j$$

The light intensity in the resist is simply the square of the magnitude of $\varepsilon_2(x,y,z)$.

The actual calculation of the standing wave pattern in a real IC structure is considerably more complicated than Eq. (5.39) might indicate. In general there will be multiple layers (some optically transparent like SiO_2 and some opaque like Al) below the resist, and the resist thickness itself will usually vary with position. Furthermore, the incident light is often not normally incident and it may not even be monochromatic in some exposure systems. Nevertheless, basic optical theory can be used to calculate $\varepsilon_2(x,y,z)$ at each point in the resist even with all these additional complications [5.3, 5.26, 5.27].

An example of such a simulation is shown in Figure 5–44. A relatively simple structure is defined with a silicon substrate etched to form a sloped step. The photoresist layer is deposited so it has a flat upper surface and hence its thickness varies over the structure. A mask is defined which allows light to enter in the center part of the structure ($x = 1$ to $x = 2$ μm). The lower part of Figure 5–44 illustrates the calculated DNQ resist *PAC* concentration after the exposure. As we will see in the next section on resist exposure kinetics, this 2D concentration profile corresponds to the integrated light intensity profile through the exposure process. There are a number of interesting features that can be observed from this simulation.

First, notice at the left-hand side of the exposed region the effects of standing waves in the light intensity. The silicon substrate in this example is a reflecting substrate and so standing waves will be set up. Nodes of maximum light intensity correspond to minimums in the *PAC* (maximum exposure). These are the dark bands on the left side. They occur at $\lambda/2n_2$ distance intervals. The nodes of minimum light intensity (maximum *PAC* concentration) also occur with a period of $\lambda/2n_2$ and are the light gray bands. Notice also the effect of the sloped sidewall in the center part of the structure. Here the incoming light waves are reflected to the right, setting up a standing wave pattern parallel to the

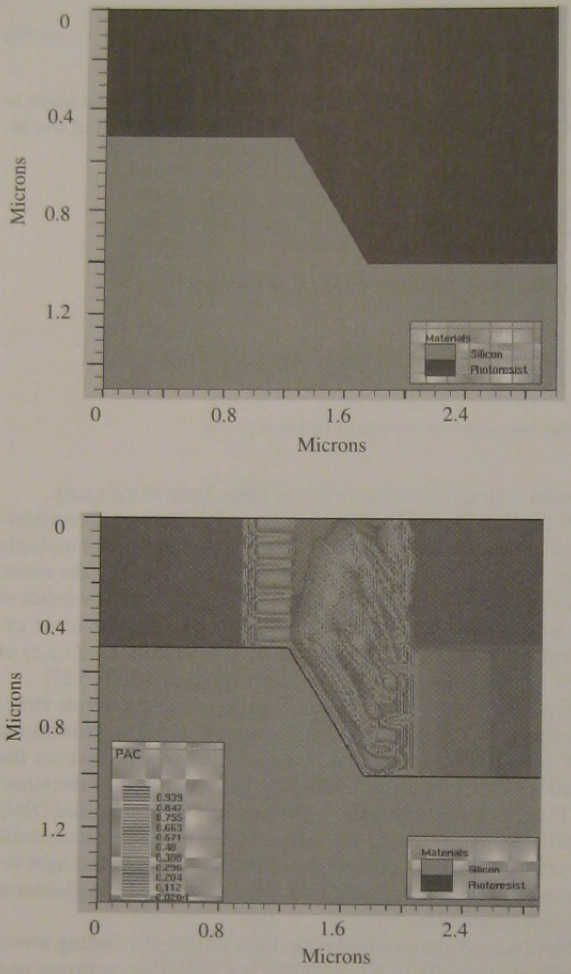


Figure 5–44 Example of calculation of light intensity distribution in a photoresist layer during exposure using the ATHENA simulator by Silvaco. A simple structure is defined at the top with a photoresist layer covering a silicon substrate, which has two flat regions and a sloped sidewall. The exposure takes place between $x = 1$ and $x = 2$ μm. The bottom simulation shows the calculated DNQ resist *PAC* concentration after an exposure of 200 mJ cm^{-2}. Lower *PAC* values correspond to more exposure. The gray-scale contours thus correspond to the integrated light intensity from the exposure.

sloped sidewall. In addition to this standing wave pattern, the reflected light waves will cause the exposed region to extend into the masked region on the right. This will result in image degradation. We will return to this example later to see the effects of post-baking and resist development on the final resist structure. For now, it should be apparent that numerical simulation tools are required to deal with the complexity of structures encountered in modern ICs.

The final issue with respect to the light intensity distribution through the photoresist is resist bleaching. As we saw earlier, this occurs in DNQ resists because the *PAC* in

these resists initially absorbs photons while it is being exposed, and then becomes more transmissive as the *PAC* component is converted to carboxylic acid. Thus the top layers of the resist initially absorb the incoming photons and later in the exposure allow more light to penetrate to the deeper layers of the resist. This results in a time-dependent light intensity throughout the resist. Because this issue is directly related to the resist exposure chemistry, we will discuss models for it in the next section on resist exposure. Qualitatively, the effects of this bleaching process can be observed in the bottom of Figure 5–44 which shows, on average, a gradual increase in the *PAC* concentration with depth into the resist.

5.5.3 Photoresist Exposure

In this section until its end, we will neglect any standing waves that might be present in the resist. These effects can easily be included in the exposure model using Eq. (5.39) and we will include such effects in the simulation examples at the end of the section. However, the principle ideas of photoresist exposure can be understood without the complexity of standing waves, and that is the approach we will initially take here. We will also, for the moment, simply consider a one-dimensional structure with a uniform resist layer being exposed.

5.5.3.1 g-line and i-line DNQ Resists

The light incident on the photoresist is primarily absorbed by the *PAC* component of the resist. The *PAC* is assumed to be uniformly distributed throughout the resist. As we have seen, bleaching occurs from the top layers of the resist downward as the *PAC* is exposed. The basic equation describing the light absorption is given by

$$\frac{dI}{dz} = -\alpha I \tag{5.40}$$

where z is the direction into the resist. This equation basically says that the probability that a photon will be absorbed is proportional to the light intensity with alpha, the absorption coefficient, as the proportionality constant. If alpha is a constant, integration of this equation results in Eq. (5.15), which we presented earlier as a simple description of how the light intensity falls off in the photoresist with depth.

Alpha will turn out to be related to the concentration of the *PAC*, which is unexposed in the resist at a particular time and location. Thus alpha is not a constant but rather is a function of position and time. At any time during the exposure, the light intensity profile in the resist will be given by

$$I(z) = I_0 \exp\left(\int_0^z \alpha(z')dz'\right) \tag{5.41}$$

where the integral in the exponent is the absorbance of the resist down to a depth z.

We now need to relate α to the material properties of the resist. If the *PAC* component of the resist is dilute, then the absorption coefficient of the resist is simply

$$\alpha_{resist} = \alpha_{PAC}[PAC] \tag{5.42}$$

where $[PAC]$ is the concentration of the *PAC* constituent in the resist and α_{PAC} is the absorption coefficient of the *PAC* material. Eq. (5.42) simply says that the resist absorption is determined by the amount of *PAC* in the resist. The dilute approximation required for Eq. (5.42) has been shown to hold for practical resist formulations [5.29].

DNQ resists consist of several constituents as we saw earlier in this chapter. In general, such resists will have a resin R (usually novolac), a photoactive component *PAC*, a solvent S, and as the exposure proceeds, exposure products P (principally the carboxylic acid produced by the *PAC*). More generally, then, Eq. (5.42) can be written as

$$\alpha_{resist} = \alpha_{PAC}[PAC] + \alpha_R[R] + \alpha_S[S] + \alpha_P[P] \tag{5.43}$$

Since the exposure products result from the *PAC*, we may write that

$$P = [PAC]_0 - [PAC] \tag{5.44}$$

Following the approach of [5.21], we rewrite Eq. (5.43) in the following form

$$\alpha_{resist} = Am + B \tag{5.45}$$

where

$$A = (\alpha_{PAC} - \alpha_P)[PAC]_0$$

$$B = \alpha_P[PAC]_0 + \alpha_R[R] + \alpha_S[S]$$

$$m = \frac{[PAC]}{[PAC]_0}$$

A and B in Eq. (5.45) are experimentally measurable parameters for a given photoresist and are known as the first two Dill parameters [5.21]. A is the absorption coefficient of the bleachable components of the resist and B is the absorption coefficient of the nonbleachable components of the resist. We will describe how these parameters can be measured for a given resist shortly. Note that when the resist is unexposed, $m = 1$ and $\alpha_{resist} = A + B$. When the resist is fully exposed, $m = 0$ and $\alpha_{resist} = B$.

We can now substitute Eq. (5.45) into Eq. (5.40) and express the light intensity through the resist layer as

$$\frac{dI}{dz} = -(Am + B)I \tag{5.46}$$

Unfortunately, solving this equation is complicated by the fact that m is a function of time and position so that Eq. (5.46) really should be written

$$\frac{dI}{dz} = -[Am(z,t) + B]I \tag{5.47}$$

m is the fraction of the *PAC* component of the resist that is unexposed at time t. If we again assume first order reaction kinetics, then

$$\frac{dm}{dt} = -CIm \tag{5.48}$$

This equation states that the exposure rate of the *PAC* is proportional to the remaining unexposed *PAC* concentration and to the light intensity. The constant of proportionality is C, which is the third Dill resist parameter.

Eqs. (5.47) and (5.48) are coupled equations that must be solved simultaneously. The third equation that must be included is Eq. (5.39), which accounts for any standing wave patterns in the resist. Simulators which are available today [5.3, 5.26, 5.27] solve this system of equations to calculate the time evolution of the exposure of the resist. An iterative procedure is used which basically operates as follows.

At $t = 0, m = 1$ and the light intensity pattern can be calculated in the resist using Eq. (5.39). This intensity distribution is then used in Eq. (5.48) to calculate the exposure versus position. The result is $m(x,y,z,t_0)$ in the general three-dimensional case. The result is then used in Eqs. (5.47) and (5.39) to calculate the light intensity at $t = 0 + \Delta t$. This result is again used in Eq. (5.48) to calculate the new light intensity after the next time step. Through this iterative procedure, the time evolution of the exposure process is simulated. The overall result is that the aerial image produced by the optical system is converted into a latent 3D image in the photoresist. Figure 5–44 showed an example of such a simulation in which the output shown in the bottom of the figure is the [*PAC*] as a function of position at the end of the exposure.

Calculation of the latent image in the resist requires knowledge of the Dill resist parameters A, B, and C. These parameters can be measured experimentally for a particular resist material as illustrated in Figure 5–45. In their series of classic papers in 1975, Dill and his co-workers showed that a simple measurement of the time dependent light transmission through an exposing resist film could be used to extract the A, B, and C parameters. As illustrated in Figure 5–45, light at the exposing wavelength is used to illuminate a resist coated substrate. In the simplest case, a transparent substrate is chosen to have the same index of refraction as the resist, so that there are no reflections at the resist/substrate interface. An antireflection coating is used on the backside of the substrate to avoid reflections at that interface. A typical measurement result is shown schematically in Figure 5–46. For this simple case, Dill and colleagues showed that [5.21]

$$A = \frac{1}{D} \ln\left(\frac{T_\infty}{T_0}\right) \tag{5.49}$$

$$B = -\frac{1}{D} \ln(T_\infty) \tag{5.50}$$

$$C = \frac{A + B}{AT_0(1 - T_0)T_{12}} \left.\frac{dT}{dE}\right|_{E=0} \tag{5.51}$$

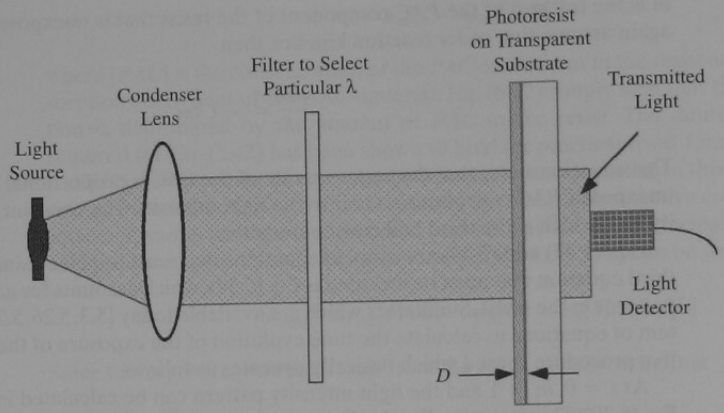


Figure 5–45 Conceptual experimental setup for measuring the Dill parameters characterizing a particular resist.

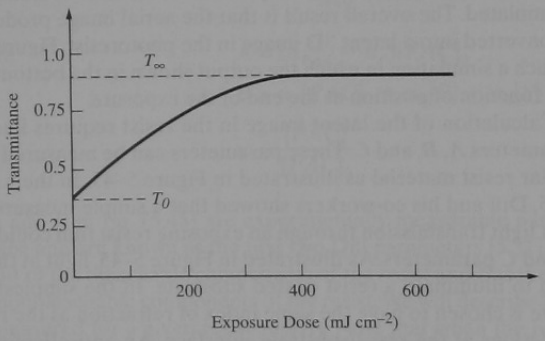


Figure 5–46 Typical experimental result from a measurement like that shown in Figure 5–48.

where

$$T_{12} = 1 - \left(\frac{n_{resist} - 1}{n_{resist} + 1}\right)^2 \tag{5.52}$$

T_0 is the transmitted light intensity at the start of the exposure, T_∞ is the transmitted light intensity at the end of the exposure, D is the resist thickness, and T_{12} is the transmittance through the air/resist interface, with n_{resist} the index of refraction of the resist. More sophisticated parameter extraction methods can be used if the experimental situation is not quite as ideal as Figure 5–45 [5.3, 5.21].

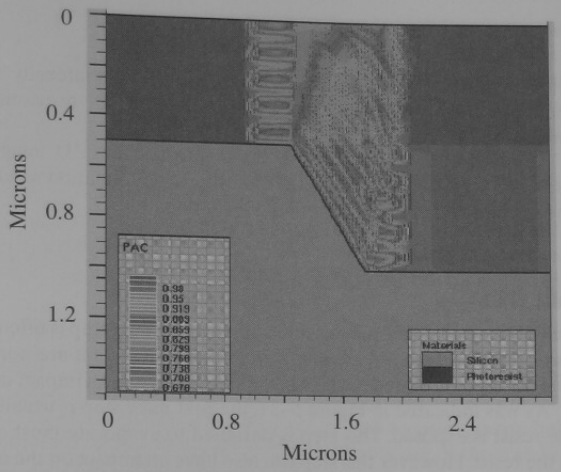


Figure 5–47 Same simulation example as in Figure 5–44 except that the exposure dose has been reduced to 10% of the value used in Figure 5–44. The deeper regions of the photoresist are not as exposed as the near surface regions because of the effects of bleaching.

Figure 5–47 shows an additional simulation with the same structure as was used in Figure 5–44. The exposure dose in Figure 5–47 has been reduced to 10% of the previous value to illustrate the effects of bleaching during resist exposure. With only 10% of the exposure dose, it is apparent that the upper layers of the resist are still exposed but the lower regions have not received sufficient photon flux to fully expose the resist.

5.5.3.2 DUV Resists

Chemically amplified DUV resists consist of a polymer resin (blocked to make it insoluble in the developer in the case of a positive resist), a *PAG*, and other additives. The kinetics of the exposure reaction are assumed to be first order, that is,

$$\frac{d[PAG]}{dt} = -CI[PAG] \tag{5.53}$$

where $[PAG]$ is the time-dependent *PAG* concentration, I is the exposure intensity, and C is the exposure rate. Note the similarity to Eq. (5.48). If I is constant during the exposure,

$$[PAG] = [PAG]_0 e^{-CIt} \tag{5.54}$$

Since the acid concentration $[H]$ results directly from the *PAG* exposure, $[H]$ is given by

$$[H] = [PAG]_0(1 - e^{-CIt}) \tag{5.55}$$

where in general, $[H]$ and I will be $f(x,y,z)$. If the light intensity is also a function of time, then an iterative technique like that described in the previous section can be used to calculate $H(x,y,z,t)$.

The aerial image is thus converted into a latent 3D image in the resist which is "stored" as $H(x,y,z)$ at the end of the exposure. Simulation results would look much like the examples shown earlier in Figures 5–44 and 5–47.

5.5.4 Postexposure Bake (PEB)

Once the exposure of the resist is complete, the next step is often a postexposure thermal bake (PEB). There are actually several bake steps that are normally employed in photoresist processing and each of these steps can have an impact on the final resist pattern.

As was indicated in Figure 5–31, the first bake step is usually a prebake step before the resist is exposed. This step is designed to evaporate most of the remaining solvent in the resist. However this step can also have an impact on the photoresist chemistry. At the prebake temperature (typically 90 – 100°C), the *PAC* in DNQ resists begins to decompose into a nonphotosensitive product [5.30]. The qualitative result of this is that the concentration of the *PAC* at the start of the exposure, $[PAC]_0$, is smaller than the original concentration in the resist. Standard prebake conditions usually result in 10% or less of the *PAC* being decomposed. There has been some work on quantitative modeling of these effects [5.3, 5.31], but in many cases the resist parameters needed for such models are not well known and the chemical processes taking place are not fully understood. Most lithography simulators that are available today do not model the effects of resist prebake at all, so we will not deal with this particular bake step further.

5.5.4.1 g-line and i-line DNQ Resists

The postexposure bake is sometimes incorporated into DNQ resist process flows in order to reduce the effects of standing waves. We saw examples of these effects in the simulations in the last section. If nothing is done to mitigate the standing wave effects, the developed resist image will have distinctly scalloped edges (recall Figure 5–24). This is generally undesirable because it degrades linewidth control and resolution. Fortunately, it has been found that a simple thermal bake between exposure and development can have a dramatic effect on the standing wave effects [5.32]. Typically this bake is 10–30 minutes at about 100°C (similar to the resist prebake step).

The effect of the postexposure bake on DNQ resist chemistry is usually modeled as a simple diffusion process. The exposed *PAC* (*P* in our terminology), is assumed to be able to diffuse in the resist with some diffusivity given by

$$D_P = D_0 \, e^{-E_A/kT} \tag{5.56}$$

This is of the same form as dopant diffusion coefficients that we will discuss in Chapter 7. The effect of the *PAC* diffusion is simply to "smear out" the standing wave effects.

Such a diffusion process can be modeled in lithography simulators. However, this is more complicated than might appear at first glance, first because D_P has not been measured for most resists, and second, because it is likely that the diffusivity is not as simple as Eq. (5.56) might suggest. For example, it is believed that D_P is concentration dependent.

In some lithography simulators, default or user specified values for D_P are used along with Eq. (5.56). The postexposure bake is then modeled as a diffusion process. In other simulators, a simpler approach is taken because of the lack of D_P data. These simulators model the postexposure bake with a simple diffusion length model. The distance over which the *PAC* "smearing" occurs during the bake is simply given by

$$\sigma = \sqrt{2 D_P t_{bake}} \tag{5.57}$$

where t_{bake} is the bake time. σ has units of distance and is generally a user-supplied parameter in simulators. A typical value might be ≈ 0.05 μm (50 nm) or about $^1/_4$ λ.

Figure 5–48 illustrates the effect of such a postexposure bake on the simulation previously shown in Figure 5–44. The smearing out of the standing wave pattern is apparent in the calculated [*PAC*].

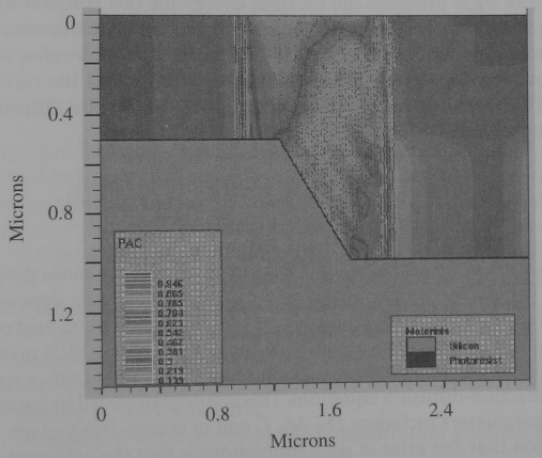


Figure 5–48 Same simulation example as in Figure 5–44 except that a postexposure bake of 45 minutes at 115°C has now been included. The gray-scale contours again correspond to the [*PAC*] after exposure. Note that the standing wave effects apparent in Figure 5–46 have been "smeared out" by this bake, producing a more uniform [*PAC*] distribution.

5.5.4.2 DUV Resists

In DUV resists, the PEB is a critical step because it completes the exposure process by driving the reaction between the *PAG* molecules and the polymer chains. In positive DUV resists this reaction unblocks the polymer chain, making it soluble in the developer. In negative resists, the *PAG* drives a crosslinking reaction that makes the resist insoluble in the developer. In either case, the acid molecules are not consumed by the reactions and hence to first order, $[H]$, given by Eq. (5.55), remains constant during the PEB.

If $[M]$ represents the concentration of the reactive sites on the resist polymer chains, then

$$\frac{d[M]}{dt} = -C[M][H] \tag{5.58}$$

where C is the reaction rate constant. This expression assumes that the reaction between the acid H and the reactive sites M is first order. During the PEB, M begins with a starting concentration $[M_0]$ and decreases with time.

$$[M_0] - [M] = [M_0](1 - e^{-C[H]t}) \tag{5.59}$$

In general $[H]$ and therefore $[M]$ are $f(x,y,z,t)$.

The PEB provides the thermal energy for this reaction to take place because C in Eq. (5.58) is generally exponentially dependent on temperature. The second process driven by temperature is the diffusion of the acid molecules, since they must physically move from reactive site to reactive site to complete the resist exposure (recall Figure 5–19). This diffusion process is described by a standard diffusion equation, which in one dimension is

$$\frac{\partial[H]}{\partial t} = \frac{\partial}{\partial x}\left(D_H \frac{\partial[H]}{\partial x}\right) \tag{5.60}$$

We will discuss diffusion and this equation in much more detail in Chapter 7.

The diffusivity D_H of the acid molecules generally depends exponentially on temperature as in Eq. (5.56). But it may also depend on the acid concentration, making the modeling problem more difficult. Simulation of the PEB process in a DUV resist thus depends on simultaneously solving Eqs. (5.58) and (5.60), a system of equations known as reaction-diffusion equations. This is straightforward to implement although a number of parameters including D_H, M_0, C and of course $H(x,y,z)$ are required. A final complication that can arise in such simulations is that the acid concentration $[H]$ may not be time independent because of various loss mechanisms that can occur during the PEB. These mechanisms include poisoning by atmospheric base contaminants, evaporation of the acid molecules, or simply trapping of the acid molecules at sites in the photoresist. All of these mechanisms can be modeled and included in simulations if specific models and parameters are available. Generally such effects are not included in simulators today. Simulation of the PEB effects on DUV resists would look similar to Figure 5–48.

5.5.5 Photoresist Developing

A number of models have been used to describe photoresist developing. The earliest models came from Dill and his co-workers and were empirically based [5.21]. Some more recent models have attempted to use a more physical basis [5.33]. In all cases, the process is basically described as a surface controlled etching process. The developer solution is assumed to etch the surface of the resist isotropically at a point (x,y,z) at a rate that is determined by the resist inhibitor concentration at that point. The photoactive compound (PAC) in DNQ resists or the blocked polymers in DUV resists are insoluble in the developer whereas when the PAC is converted to carboxylic acid (P) in DNQ resists, or when the PAG unblocks the polymer chains in DUV resists in the exposure process, the resist becomes quite soluble in the developer. We will describe the developing process below in terms of DNQ resist parameters. However, the same models apply to DUV resists.

The development or etching rate at point (x,y,z), thus depends on the $[P]$ or $[PAC]$ at that point. This is exactly what is calculated by the exposure models we previously described, and what is shown in the examples in Figures 5–44, 5–47 and 5–48. What is needed then, is a mathematical relationship between $[P]$ or $[PAC]$ and the development rate.

The first such relationship was described by Dill and colleagues [5.21] and is implemented in simulators today typically in a form like the following [5.26]:

$$R(x,y,z) = \begin{cases} 0.006 \exp(E_1 + E_2 m + E_3 m^2) & \text{if } m > -0.5\dfrac{E_2}{E_3} \\ 0.006 \exp\left(E_1 + \dfrac{E_2}{E_3}(E_2 - 1)\right) & \text{otherwise} \end{cases} \tag{5.61}$$

$R(x,y,z)$ is the local development or etching rate in μm min^{-1}, m [actually $m(x,y,z)$] was defined in Eq. (5.45) and is the local $[PAC]$ after exposure, and E_1, E_2 and E_3 are empirical parameters obtained from fitting Eq. (5.61) to experimental data.

A more physically based model for resist development has been proposed by Mack [5.33]. This model has many similarities to the Deal-Grove oxidation model that we will discuss in Chapter 6. Mack's model is illustrated in Figure 5–49. In this model, three fluxes are used to describe the development process. F_1 represents the diffusion of developer from the bulk of the liquid to the resist surface. F_2 represents the reaction of the developer with the resist, and F_3 represents the diffusion of the reaction products back into the developer solution. F_3 is assumed to be very fast (not rate limiting) and hence is neglected in Mack's analysis.

F_1 is assumed to be driven by the concentration gradient between the liquid developer and the resist surface and is given by

$$F_1 = k_D(C_D - C_S) \tag{5.62}$$

where k_D is the mass transfer coefficient associated with developer diffusion (cm sec^{-1}), C_D is the bulk developer concentration and C_S is the developer concentration at the re-

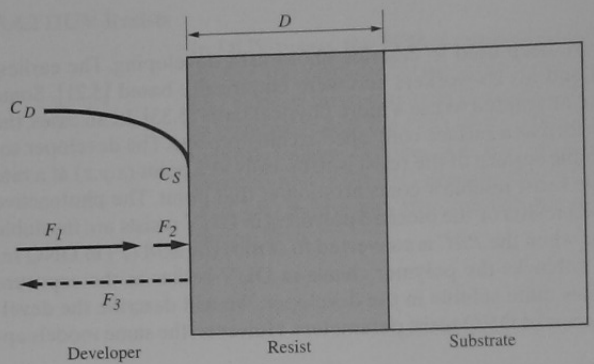


Figure 5–49 Mack's model for photoresist development.

sist surface. F_2 is assumed to depend on the concentrations of the two reacting species at the resist surface and is given by

$$F_2 = k_R C_S [P]^n \tag{5.63}$$

where k_R is the rate constant associated with the developer-resist reaction (cm $\sec^{-1}$) and $[P]$ is the local concentration of the reacted *PAC* as defined earlier. Here it is assumed that the reaction is proportional to some power of $[P]$, that is, that some number of molecules n of product P react with the developer to dissolve a resin molecule in the resist.

F_1 and F_2 are in series and so we let $F_1 = F_2$ in steady state. Thus

$$F_1 = F_2 = \frac{k_D k_R C_D [P]^n}{k_D + k_R [P]^n} \tag{5.64}$$

Recalling the definition of $[P]$ and m in Eqs. (5.44) and (5.45), we have finally that the development rate $r = F_1 = F_2$ is

$$r = \frac{k_D C_D (1 - m)^n}{\frac{k_D}{k_R [PAC]_0^n} + (1 - m)^n} \tag{5.65}$$

When $m = 1$, the resist is unexposed and $r \to 0$. This is not physically correct because conventional DNQ resists do dissolve in the developer at a nonzero rate even if they are unexposed. Eq. (5.65) can be corrected to account for this as follows:

$$r = \frac{k_D C_D (1 - m)^n}{\frac{k_D}{k_R [PAC]_0^n} + (1 - m)^n} + r_{min} \tag{5.66}$$

where r_{min} is the dissolving rate in unexposed resist. When $m = 0$, the resist is fully exposed and $r \rightarrow r_{max}$ where

$$r_{max} = \frac{k_D C_D}{\frac{k_D}{k_R [PAC]_0^n} + 1} + r_{min} \tag{5.67}$$

We can consider two limiting cases of this model. If the diffusion process in the liquid developer is fast, the overall reaction will be limited by the reaction at the surface (surface reaction rate limited). On the other hand, if the surface reaction rate is the faster of the two processes, the overall reaction will be "diffusion limited."

$$r \approx \begin{cases} k_R C_D [P]^n + r_{min} & \text{Surface Reaction Rate Limited} \\ k_D C_D + r_{min} & \text{Diffusion Rate Limited} \end{cases} \tag{5.68}$$

Eq. (5.66) is often written in the following form:

$$r = r_{max} \frac{(a+1)(1-m)^n}{a + (1-m)^n} + r_{min} \tag{5.69}$$

where

$$a = \frac{k_D}{k_R [PAC]_0^n} \tag{5.70}$$

The parameter a is really a measure of the relative rates of diffusion and the surface reaction.

To use this model, there are four parameters that must be determined experimentally, a, m, r_{max} and r_{min} although r_{min} is often negligible so there may be only three parameters. Dill's original model also contained three parameters. However Mack's model attempts to use parameters which have some physical basis as opposed to the purely empirical parameters in Dill's model.

Commercially available simulators generally implement a number of models for resist development, including the two described here. The time evolution of the developing resist profile is calculated by setting up a two- or three-dimensional grid in which each grid point in the resist has a specific value of $m(x,y,z)$. The developing pattern is then allowed to evolve over time with the developer moving into the resist at a local rate determined by the development model.

Figure 5–50 shows an example of a simulation of the development process. In this example the simplest development model (the Dill model) was used. Normal parameters were used in the simulation for exposure time and development time, and a post-exposure bake was included to minimize standing wave effects. The resulting developed resist pattern shows good resolution and fairly sharp sidewalls. The simulation on the bottom left shows the resist profile partway through the development. Note how the shape of the remaining resist at this point mirrors the *PAC* concentration in the resist pattern.

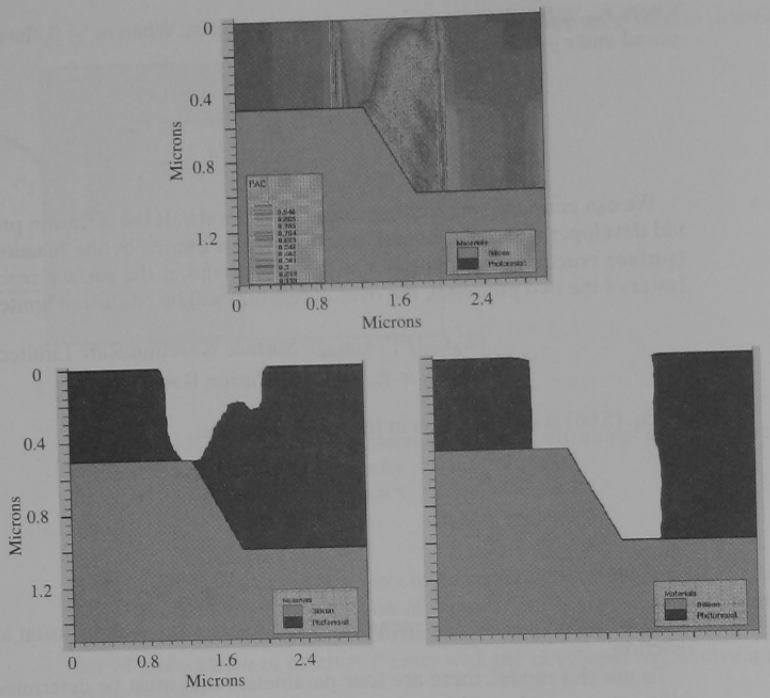


Figure 5–50 Example of the calculation of a developed photoresist layer using the ATHENA simulator by Silvaco. The resist was exposed with a dose of 200 mJ cm^{-2}, a postexposure bake of 45 min at 115°C was used, and the pattern was developed for 60 seconds, all normal parameters. The Dill development model was used. The top image shows the *PAC* concentration after exposure and postbake. The bottom simulations show the developed resist pattern partway through the development (left) and after complete development (right). The resist pattern is well defined in this simulation.

5.5.6 Photoresist Postbake

As shown in Figure 5–31, the final step in a typical photoresist process is a postbake, typically done at 100 – 140°C for 10 – 30 minutes. This step hardens the resist by evaporating any remaining solvents and also improves adhesion to underlying materials. Since the resist is generally used to mask a plasma etching step, or to block an ion implantation step, this hardening process is important in improving the robustness of the resist. The resist will generally flow somewhat during this postbake, rounding the edges of the profile. Obviously the degree of flowing depends on the time and temperature of the postbake, as well as on the resist material properties.

Most lithography simulators do not model this postbake process. However some of these tools have implemented a flow simulator using models which are similar to the

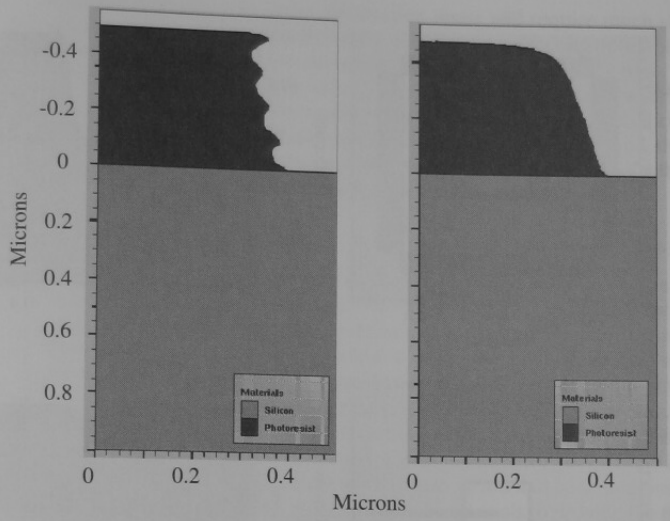


Figure 5–51 Example of the simulation of the reflow process using the ATHENA simulator by Silvaco. The initial developed resist pattern is shown on the left. A 10-min 130°C postbake results in the structure on the right. Note the standing wave effects present in the developed resist profile, which disappear after the reflow.

glass reflow models discussed in Chapter 11. A detailed discussion of such models is included in Chapter 11, so it will not be repeated here. Figure 5–51 shows an example of the simulated effects of the postbake on the resist profile.

5.5.7 Advanced Mask Engineering

In Section 5.2.4, we described some of the advanced mask engineering concepts that are being applied to masks to improve the resolution of optical projection systems. These methods include Optical Proximity Correction (OPC) and phase shift masks. These techniques basically change the description of the mask transmission function [Eq. (5.21)]. Modern simulators include these possibilities simply by changing the mask description in the simulator input. An example of such a simulation is shown in Figure 5–52 in which a small contact hole is printed in resist using a mask with and without phase shifting. The higher quality of the developed image in the resist is apparent in the case with a phase shifted mask.

Simulation tools will probably play an increasingly important role as more mask designs incorporate these advanced features. The design of such mask patterns is not straightforward and the use of computers to both help in the design and to "check" the design using simulation tools is likely to become very common in the future.

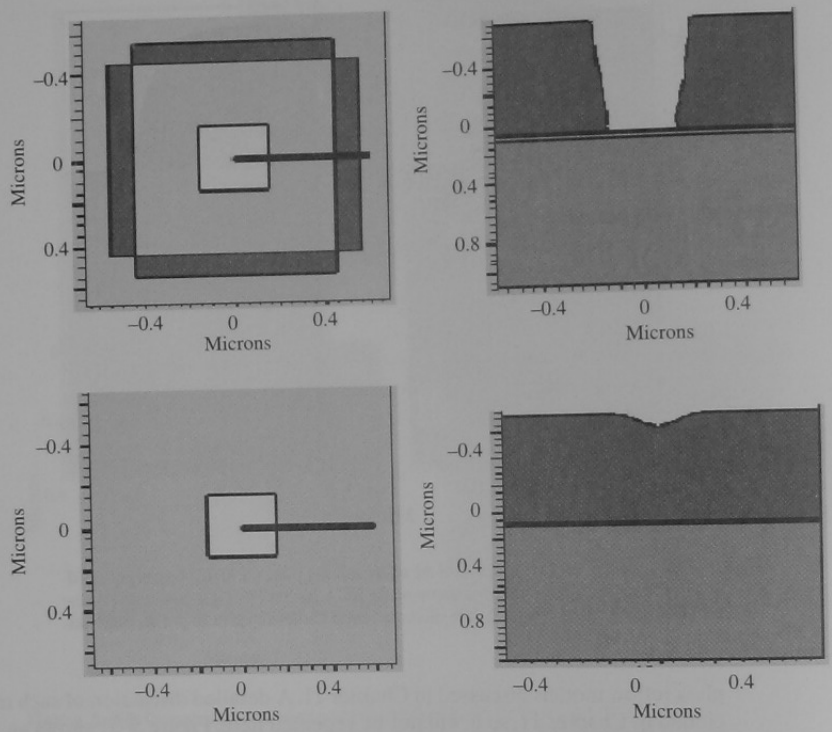


Figure 5–52 Example of the simulation of an developed photoresist image with a phase shifting mask using the ATHENA simulator by Silvaco. In the top example, the mask design on the left is a 0.3-μm contact hole (central area) with 0.1-μm phase shifting features outside the border of the contact hole ("outriggers"). These features are absent in the bottom example. The difference in the developed contact hole features is apparent. In this case an i-line stepper with a 0.54 *NA* was used in the simulation along with normal postexposure bake and development cycles.

5.6 Limits and Future Trends in Technologies and Models

The demise of optical lithography has been predicted many times over the past decade. Yet continued innovations in optical exposure systems have kept pace with manufacturing needs. The SIA NTRS described at the beginning of this chapter projects the need for 0.13-μm features in 2003, and 0.1-μm features in 2006. It now appears likely that optical exposure systems can provide the tools needed for the 0.13-μm generation and perhaps even the 0.1-μm generation, using 193 nm excimer laser sources and step and scan systems. There is even some work today on a 157-nm (F_2) excimer laser source. Somewhere beyond the 0.1-μm generation, however, the actual demise of optical lith-

ography probably will occur. Diffraction limits will require much shorter wavelength sources, and exposure systems likely will not use refracting optics because of the difficulty in finding transparent materials at these shorter wavelengths.

There is no consensus on lithography approaches for manufacturing < 0.1-μm integrated circuits. There is obviously a significant amount of research currently underway to evaluate various alternatives, because if a solution is not found, this problem will be a show stopper for the continued evolution of ICs. We will briefly mention here some of the significant alternatives for lithography in the future.

5.6.1 Electron Beam Lithography

Electron beam systems have been used for many years to make the masks needed for optical lithography, as was illustrated in Figure 5–1. There is no conceptual reason why such a system cannot be used to directly write patterns in photoresist on wafers rather than in photoresist on mask substrates. In fact for many years, this technique has been the method of choice for advanced device research because of the superb resolution offered by e-beam tools. While electrons do have wavelike properties, the wavelength of the electrons used in these exposure tools is less than 0.1 nm, so diffraction effects and the limits they impose on optical exposure systems are not an issue in e-beam systems. Even 20 years ago, feature sizes smaller than 10 nm had been demonstrated [5.34]. Thus there is no question about the ability of e-beam tools to provide the resolution required for feature sizes through the end of the current SIA NTRS.

The major drawback of e-beam tools is the very small throughput compared to current optical steppers. The simplest e-beam tools essentially expose the resist pixel by pixel, a serial process. Optical steppers on the other hand expose the pixels on an area of the wafer in parallel. This is a major difference and results in e-beam lithography tools having a throughput on the order of one wafer per hour compared with >50 wafers per hour for modern optical steppers. The manufacturing cost differences associated with the two technologies are such that e-beam tools have not found application to date in mainstream manufacturing plants. e-beam tools are used today for fabricating small quantities of experimental devices and circuits and for small-scale manufacturing of special-purpose chips, which require small features or fast turnaround. (Since no masks have to be made, the time required to go from design to a fabricated chip can be shorter using e-beam lithography, although only for small quantities of chips.)

The throughput limitation on e-beam systems arises from a number of considerations. First, the electron source can provide only a finite current density. Since a minimum number of electrons are required to expose each pixel in the resist, this sets an upper bound on how fast the beam can be scanned on the wafer. If the resist sensitivity is S, then the minimum number of electrons needed to expose each pixel is simply

$$N = \frac{SI_P^2}{q} \tag{5.71}$$

where I_P is the pixel dimension. Second, at high-beam intensities, resolution can be degraded because of defocusing caused by the coulomb repulsion of the charged electrons in the beam. Approaches to overcoming these limitations include using multiple-electron

sources so that parallel writing can be accomplished, shaping the beam into rectangles large enough to expose full features on the wafer rather than individual pixels, improving the sensitivity of e-beam resists, and using e-beam projection lithography in which a stencil mask containing complex patterns is focused on the wafer. This latter approach is most useful when the IC contains repetitive patterns, such as in memory chips.

Proximity effects are an additional significant issue associated with e-beam direct writing. The electrons striking the resist typically have an energy of 10 – 20 keV and travel distances greater than a micron before coming to rest. They give up their energy to the resist and substrate materials through a variety of interactions including elastic scattering off nuclei and inelastic interactions like those illustrated in Figure 4–18. In fact the overall stopping process is not unlike the ion implantation process described in detail in Chapter 8. Because the electrons have such a long range and because they scatter so readily, exposure of regions adjacent to the desired exposure areas does occur. This is known as the proximity effect. It is possible to model this effect since the overall mask pattern is known and to correct the exposure process for this effect. The procedure basically is to calculate how much exposure takes place in adjacent regions because of scattering and then to adjust the electron dose spatially to compensate for such effects. Sophisticated algorithms have been developed to accomplish this [5.35].

Perhaps the most promising approach to high-throughput e-beam lithography is the SCALPEL® (SCattering with Angular Limitation Projection Electron-beam Lithography), invented at Bell Laboratories in 1989 [5.36]. This invention was motivated by two observations about e-beam projection systems. The first observation was that existing stencil mask systems which basically absorb the electron beam in the parts of the mask intended to be opaque are subject to heating problems because of the energy absorbed from the e-beam. This limits the accelerating voltage that can be used in these systems. The second observation was that full field electron optics systems require small numerical apertures as die sizes increase and feature sizes decrease. As a result, the beam current densities needed for reasonable throughput in these systems result in significant space-charge effects that destroy the system resolution.

The solution to these problems proposed in the SCALPEL® system basically involves a new mask design. The principle is illustrated in Figure 5–53. The mask consists of a low atomic number membrane (typically Si_3N_4) and a thin high atomic number pattern (typically a few tens of nm of Cr or W). Both regions are essentially transparent to the collimated, incoherent 100-KeV electron beam that is used in the system. However the electrons interact differently with the two mask regions. In the membrane areas, the electrons are only weakly scattered to small angles. In the Cr/W regions, the electrons are strongly scattered to large angles. An electron lens system focuses the scattered electrons on the wafer. However, an aperture in the back focal plane of the projection system blocks the strongly scattered electrons. Thus a high-contrast, high-resolution image is produced on the wafer. This concept has been demonstrated experimentally and has both the resolution and the potential throughput to meet the SIA NTRS requirements through all the technology generations currently in the roadmap [5.37]. The system reduces the mask image typically by 4:1, requires no exotic mask patterning (OPC or phase shift masks), utilizes a step and scan approach very similar to modern optical exposure systems, achieves a large depth of focus, and works with the same chemically amplified DUV resists currently used for 248-nm lithography. This approach has not yet

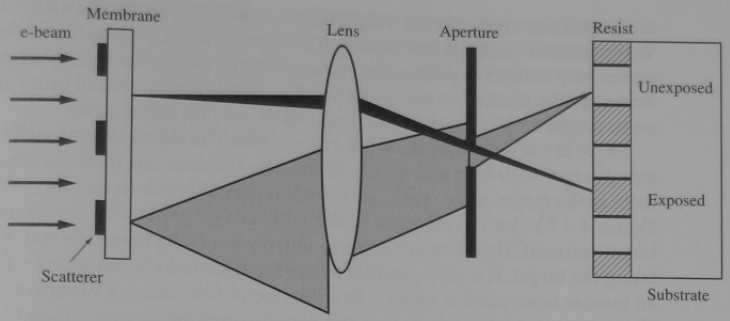


Figure 5–53 SCALPEL® e-beam projection lithography system. The resist is "exposed" in areas where the e-beam has not been widely scattered (mask areas where there is no scatterer). When the e-beam is widely scattered, the intensity reaching the resist is insufficient for exposure.

been demonstrated in a full-scale manufacturing environment, but it appears to have significant promise for future lithography needs.

5.6.2 X-ray Lithography

X-ray lithography systems use X-rays (photons) with energies in the 1 – 10 keV region, corresponding to wavelengths on the order of 1 nm. Thus diffraction effects are negligible as was the case with e-beam lithography. Focusing of X-rays is a very difficult problem so X-ray systems today are generally proximity printing systems. They have the basic configuration illustrated in Figure 5–3 except that an X-ray source is used rather than an optical source. The system is normally operated in a vacuum environment, although the wafer itself is usually in an air ambient. The X-rays emerge from the exposure tool through a thin Be window.

The discussion of near field (Fresnel) diffraction earlier in the Chapter (Section 5.2.2.3), would suggest that very good resolution could be achieved with a system operating with an exposure wavelength of about 1 nm. This is indeed the case with X-ray systems. A number of such systems are in use today in research laboratories and they have achieved < 0.1-μm feature sizes. These systems also have throughputs comparable to optical exposure systems because they expose a significant area on the wafer during each exposure. The systems often use a step and repeat system approach like optical steppers because of the limitations in making large area 1:1 X-ray masks.

The principal issues limiting the adoption of X-ray systems in manufacturing have to do with the masks and the X-ray source. The "clear" areas on the mask must be transparent to 1-nm X-rays and the "dark" areas must stop them. No materials transmit X-rays easily, so the clear areas need to be made from thin layers of low mass materials. Si, Si_3N_4, SiC, and BN a few microns thick are common choices. The dark areas need to be made from materials like Au or W which absorb X-rays, with W likely being the dominant choice in the future because of concerns with Au contamination in fabrication

facilities. Thus a typical mask structure is a thin Si_3N_4 membrane with a deposited Au or W film etched to define the mask pattern. Such a mask is a mechanically fragile and requires very careful manufacturing to control stresses in the thin membrane. The actual writing of the pattern on the X-ray mask is usually accomplished using photoresist spun on to the Au or W absorber and direct write e-beam lithography.

There are several options for X-ray sources. The simplest approach is to bombard a metal target with high energy electrons. The incident electrons excite core level electrons in the target, which emit X-rays when they fall back to their normal states. (Recall Figure 4–17.) The wavelengths of the emitted X-ray photons are characteristic of the target material. The intensity of these sources can be increased by water cooling and rotating the target. It is also possible to use laser-heated plasma sources. In these sources, an intense laser pulse is applied to a thin metal film, causing vaporization of the metal. The energy is high enough that the superheated metal vapor radiates X-rays with a wavelength of 0.5 – 2 nm. The most intense X-ray sources are synchrotron storage rings. In these systems, electrons are circulated around a ring at energies of $10^6 - 10^9$ eV. At each bending magnet around such a ring, an intense X-ray beam is emitted, each of which is suitable as the source for an X-ray exposure system. Thus one storage ring can support many aligners. Several such rings have been built in recent years for lithography research and ICs have been fabricated with them. The major drawback to synchrotron storage rings is their very high cost. However, any lithography approach for $<$ 0.1-μm feature sizes is likely going to be very expensive and synchrotron sources are the likely choice if X-ray systems are widely used in the future. Figure 5–54 illustrates a modern proximity X-ray system with a synchrotron source.

Another area of current research involves lenses and mirrors for reflecting and focusing X-rays. Most such lenses use reflection from glancing incidence mirrors or multilayer mirrors with layer thicknesses designed to make use of constructive interference

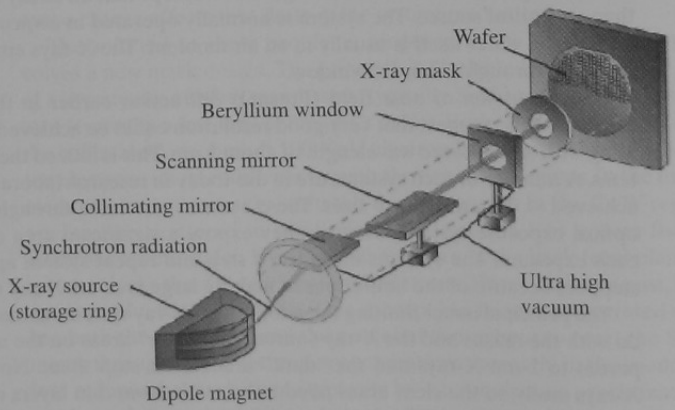


Figure 5–54 Proximity X-ray exposure system. This figure was taken from the Sematech Web site at http://www.sematech.org/public/.

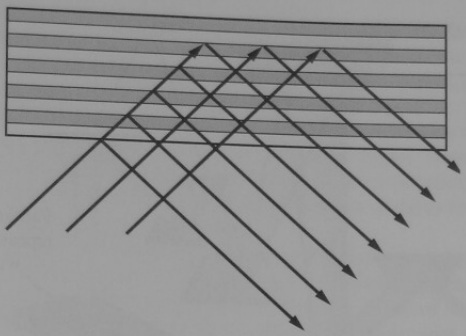


Figure 5–55 Multilayer mirror suitable for X-ray reflection.

effects. If such lenses could be sufficiently developed, they hold out the promise of making X-ray masks on reflecting substrates, a development that would have a significant impact on the potential for X-ray exposure systems.

The most promising approach for X-ray mirrors uses the concept illustrated in Figure 5–55. These mirrors use alternating layers of two materials with widely different electron concentrations (often Mo and Si). The high-mass layer acts as a scatterer and the low-mass layer acts as a spacer. The thicknesses are chosen so that constructive interference occurs between the partially reflected waves at each layer. These mirrors were originally developed for X-ray astronomy [5.38] and in recent years have achieved reflectivities above 60%. They require extraordinary manufacturing tolerances on the layer thicknesses (typically 0.1 nm) because of the short wavelengths the mirrors are designed to reflect. If such mirrors can be developed for lithography applications, designs have been proposed for complete exposure systems that use them both for masks and for reducing optics [5.39]. In the latter case, the mirrors would have to be constructed on curved surfaces. Such systems have been termed EUVs or Extreme Ultraviolet lithography systems. Figure 5–56 conceptually illustrates what an EUV lithography system might look like.

The throughput capability and the potential resolution of X-ray lithography systems will likely drive continued research on these systems. If the mask problems can be solved for either proximity or EUV systems, such systems may provide a viable manufacturing option for < 0.1-μm features.

5.6.3 Advanced Mask Engineering

Efforts to continue to improve the resolution of lithography systems must focus on the entire system, not just the exposure tool. Powerful approaches to designing masks with a more sophisticated pattern than simply a binary, exact representation of the desired pattern are beginning to be used. These methods include optical proximity correction

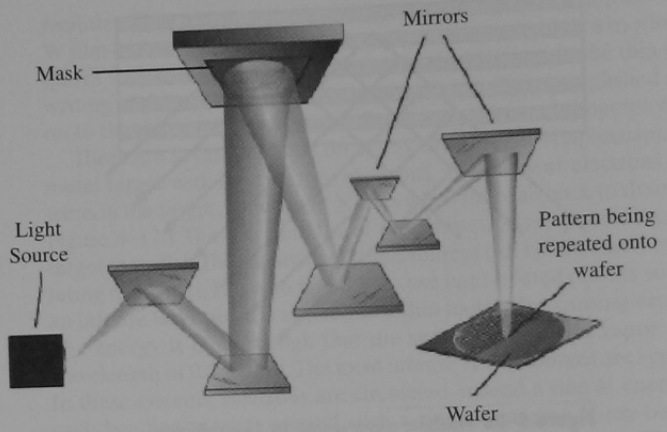


Figure 5–56 Conceptual drawing of an EUV imaging system. The mirrors reduce the projected image by about 4X. This figure was taken directly from the Sematech Web site at http://www.sematech.org/public/.

and phase shift masks. It is likely that these methods will be more widely used in the future. The major barriers to more widespread use include the sophisticated software tools that are needed to help in designing such masks, and of course, the higher cost of the mask fabrication itself. Since many of these techniques involve placing features on the mask that are smaller than the minimum feature size, mask inspection and defect detection are also more difficult with such masks. Nevertheless, these methods do provide a significant improvement in resolution and hence the motivation for using them is strong.

Many of the proposed replacements for optical lithography require innovation or invention in the masks used with the exposure tool. It is likely that these issues will have a significant impact on which of the postoptical exposure systems actually reach manufacturing maturity.

5.6.4 New Resists

DNQ resists that have been the mainstay of the semiconductor industry for many years are now being replaced with new materials and new chemistries more suitable for DUV exposure systems. The basic resin (Novolac) in DNQ resists begins to strongly absorb UV radiation below 250-nm wavelengths. Thus these resists cannot be used for 248-nm KrF or 193-nm ArF excimer laser-based exposure systems.

The concept of chemical amplification has proven to be a powerful idea for new resists and many new materials have been developed in recent years. Resists of this type are now used in manufacturing in 0.25-μm and 0.18-μm technology generations and will likely dominate the industry for some years to come. Many alternative material systems have been and are being explored, including inorganic resists (Ag-doped Se-Ge) [5.40], organosilanes which form SiO_2 when exposed [5.41], and polysilanes [5.42]. Standard

e-beam resists like PMMA are also available, of course. While it is not clear at this point which of these approaches will ultimately be successful, it does seem likely that suitable resists will be found with adequate sensitivity, resolution, defect density, and process robustness for technology generations beyond 0.1 μm.

Before concluding this section, there are several other topics we will briefly consider. These include planarization and multilayer resists. These improvements to the basic lithography process are aimed at improving the resolution of the image produced in the resist or improving the resist sensitivity. Some of these techniques are currently beginning to find use in manufacturing.

The need for planarization in resist processing was briefly discussed earlier in connection with Figure 5–22. The issue is that surface topography under the resist will cause thickness variations in the spun on resist. The result of this will be changes in the exposure process in thin versus thick regions. The thick resist regions will be underexposed, or the thin regions will be overexposed and therefore there will be linewidth variations in the latent image produced in the resist. From a practical perspective, these issues become most important at steps in the underlying topography because that is where there are abrupt changes in resist thickness. Generally what is observed is a narrowing of the resist image right at the step and a broadening of the image just off the step. These effects become very important when the height of the step is on the order of the feature size being printed, a criterion which is increasingly being met in silicon technology as linewidths shrink.

A solution to this problem is to planarize the wafer surface before the resist is applied. In back-end processing, it is common today to use Chemical Mechanical Polishing (CMP) to do this, as discussed in Chapters 2 and 11. CMP is also beginning to find application in front-end processing as well. (See Figure 2–9 in the CMOS process flow in Chapter 2.) Alternatives include the use of multilayer resists, which we will discuss below.

Figure 5–57 illustrates the basic idea behind multilayer resists. A two- or three-layer structure is used, whose purpose is to first planarize the underlying topography and then to form a relatively thin photoresist layer on top, which is the imaging layer. The intermediate layer may be needed as an etch stop in some particular cases. While there are a number of variations to this structure, the basic idea is as follows. First a relatively thick layer is spun onto the wafer to planarize the surface. Often this is simply a photoresist layer that is not sensitive to the particular wavelength of light being used in the exposure system. For example, PMMA, which is an electron beam resist that is not sensitive in the near UV, might be used in connection with a g-line or i-line exposure system. In the simplest mutlilayer resist systems, this layer would then be baked and a second thin layer of resist spun directly on top of it. This upper resist layer is sensitive to the exposure wavelength and after appropriate prebaking, this resist layer would then be exposed. Because the imaging resist layer is thin and uniform in thickness, a high-resolution latent image can be generated. The upper imaging layer is then developed.

Using this layer now as a mask, the lower planarizing layer is now flood exposed with a DUV source. Here it is important that the upper imaging layer not be transparent to the DUV, so that the lower planarizing layer is exposed only in the regions where the upper layer has been developed away. With a PMMA lower-layer DUV exposure would work since this resist is sensitive to DUV wavelengths. In essence, the thin imaging layer is being used as a contact mask for the exposure of the lower layer. As we saw earlier,

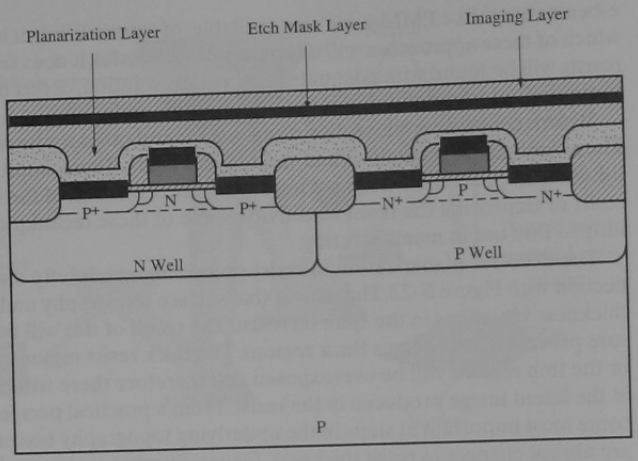


Figure 5–57 Multilayer resist structure designed to create a thin, uniform imaging layer suitable for high-resolution imaging.

contact printing can produce high-resolution images because it minimizes diffraction effects. The lower resist latent image is now developed, producing a patterned resist layer that can then be used to etch underlying layers or as an ion implantation mask.

An alternative to the above process uses the intermediate etch stop layer also shown in Figure 5–57. Here this additional layer is deposited between the planarizing and imaging layers. Typically this might be a layer of SiO_2 deposited from a spun on glass source, or it could be deposited by PECVD at a low temperature that could be tolerated by the planarizing layer. The exposure and developing of the upper imaging layer proceeds as described previously. However, etching is now used to transfer the pattern into the bottom planarizing layer. First Reactive Ion Etching (RIE) is used to etch the intermediate etch stop layer using the imaging resist as a mask. Since the imaging resist layer is thin, it likely would not be possible to use this layer as a mask during RIE etching all the way through the underlying thick planarization layer. Thus the etch stop layer is used as an intermediate pattern transfer layer. Once the pattern is etched in the etch stop layer, RIE etching of the bottom planarization layer then transfers the pattern to that layer. If PMMA or another organic material is used for the bottom layer, an O_2 RIE process can be used to etch this layer. This process has high selectivity to the intermediate layer (SiO_2, for example) and hence a relatively thin etch stop layer can be used. Whether a two- or three-layer process flow is used, the key ideas are planarization of the underlying topography followed by exposure of a thin high-resolution imaging resist layer.

Multilayer resists have not found widespread use in manufacturing at this point in time, primarily because the more complex process has not been economically viable because of high defect densities. However as minimum feature sizes continue to shrink, the concept of a thin and uniform imaging layer becomes increasingly attractive and this approach may be used in the future [5.43].

Another variation in these mutlilayer resist structures is to make the lower planarizing layer act as an Antireflection Coating (ARC) as well as a planarizing layer. In this case, the material chosen for the planarizing layer is highly absorbing at the exposure wavelength and does not bleach appreciably. Thus photons that pass through the upper thin imaging resist layer will be absorbed by the planarizing layer and will not reflect off underlying substrate features. Hence the term antireflection coating that is applied to such layers. Once the latent image is formed in the upper imaging layer, it can be transferred through the lower planarizing layer (ARC) by dry etching following development. ARC layers are useful in DNQ resists to minimize standing wave effects (Figure 5–24). They are essential in DUV resists because these resists do not bleach and hence reflections are more of an issue throughout the exposure.

5.7 Summary of Key Ideas

Lithography is one of the fundamental technologies on which modern ICs are based. The two key elements of modern optical lithography systems are the exposure tool and the photoresist. Exposure tools today generally use projection optics with diffraction-limited refracting lenses. These systems can print an area on the order of a few cm^2 on the wafer so that a step and repeat or step and scan approach must be used to print an entire wafer. g-line and i-line resists today are largely based on DNQ materials, which are used for manufacturing down to the 0.35-μm generation. DUV resists that use chemical amplification are used in the 0.25-μm and 0.18-μm technology generations and should be suitable for several future generations. Lithography simulation tools are based on two scientific fields: Fourier optics to describe the performance of exposure tools and resist chemistry to describe the formation of the mask pattern in the resist. These tools are very useful today in understanding and optimizing the performance of lithographic systems. A major change in the way in which patterns are printed on wafers will likely have to occur within the next 10 years. It does not seem possible to extend today's optical lithography techniques to the end of the SIA NTRS. Options do exist to replace optical lithography, but none are clearly viable today. If history is any guide, however, solutions will be found to the lithography problem, allowing continued shrinkage in semiconductor feature sizes.

5.8 References

[5.1]. "National Technology Roadmap for Semiconductors," SIA, 1997.

[5.2]. J. W. Goodman, *Introduction to Fourier Optics*, McGraw-Hill, 1968.

[5.3]. C. A. Mack, *Inside Prolith, a Comprehensive Guide to Optical Lithography Simulation*, Finle Technologies, Austin TX, 1997.

[5.4]. W. M. Moreau, *Semiconductor Lithography Principles, Practices and Materials*, Plenum Press, 1988.

[5.5]. W. Waldo, "Techniques and Tools for Optical Lithography" in *Handbook of VLSI Microlithography Principles, Technology and Applications*, edited by W. B. Glendinning and J. N. Helbert, Noyes Publications, 1991.

[5.6]. B. J. Lin, "Electromagnetic Near-Field Diffraction Pattern of a Medium Slit," *J. Opt. Soc. of Amer.*, vol. 62, p. 976, 1972.

[5.7]. H. Ito and C. G. Willson, "Polymers in Electronics," T. Davidson, ed., *Symposium Series 242*, American Chemical Society, Washington DC, p. 11, 1984.

[5.8]. D. Seeger, "Chemically Amplified Resists for Advanced Lithography: Road to Success or Detour?" *Solid State Tech.*, p. 115, June 1997.

[5.9]. D. Wallraff et al. "Single Layer Chemically Amplified Photoresists for 193 nm Lithography," *J. Vac. Sci. Tech.*, vol. B11, p. 2783, 1993.

[5.10]. H. Ito, "Deep-UV Resists: Evolution and Status," *Solid State Tech.*, p. 164, July 1996.

[5.11]. H. Ito, "Chemical Amplification Resists: History and Development Within IBM," *IBM J. Res. and Dev.*, vol. 41, p. 69, 1997.

[5.12]. M. D. Levenson, "Extending the Lifetime of Optical Lithography Technologies with Wavefront Engineering," *Jpn. J. Appl. Phys.*, vol. 33, p. 6765, 1994.

[5.13]. J. P. Stirniman and M. L. Rieger, *Proc. SPIE*, vol. 2197, 294, 1994.

[5.14]. M. D. Levenson, N. S. Visnwathen, and R. A. Simpson, "Improving Resolution in Photolithography with a Phase Shifting Mask," *IEEE Trans. Elec. Dev.*, vol. ED-29, p. 1828, 1982.

[5.15]. R. Larrabee, L. Linholm, and M. T. Postek, "Microlithography Metrology" in *Handbook of VLSI Microlithography Principles, Technology and Applications*, edited by W. B. Glendinning and J. N. Helbert, Noyes Publications, 1991.

[5.16]. D. Nyyssonen and R. D. Larrabee, "Submicrometer Linewidth Metrology in the Optical Microscope," *J. Res. Natl. Bur. Stand.*, vol. 92, p. 187, 1987.

[5.17]. M. G. Buehler, "Microelectronic Test Chips for VLSI Electronics," in *VLSI Electronics: Microstructure Science*, vol. 6, p. 529, Academic Press, 1983.

[5.18]. T. F. Hasan and D. S. Perloff, "Automated Electrical Measurement Techniques to Control VLSI Linewidth, Resistivity and Registration," *Test and Measurement World*, vol. 5, p. 78, 1985.

[5.19]. D. K. Schroder, *Semiconductor Material and Device Characterization*, John Wiley & Sons, 1990.

[5.20]. F. H. Dill, "Optical Lithography," *IEEE Trans. Elec. Dev.*, vol. ED-22, p. 440, 1975.

[5.21]. F. H. Dill, W. P. Hornberger, P. S. Hauge, and J. M. Shaw, "Characterization of Positive Photoresist," *IEEE Trans. Elec. Dev.*, vol. ED-22, p. 445, 1975.

[5.22]. K. L. Konnerth and F. H.Dill, "In-Situ Measurement of Dielectric Thickness During Etching of Developing Processes," *IEEE Trans. Elec. Dev.*, vol. ED-22, p. 452, 1975.

[5.23]. F. H.Dill, A. R. Neureuther, J. A. Tuttle, and E. J. Walker, "Modeling Projection Printing of Positive Photoresists," *IEEE Trans. Elec. Dev.*, vol. ED-22, p. 456, 1975.

[5.24]. W. G. Oldham, S. N. Nandgaonkar, A. R. Neureuther, and M. O'Toole, "A General Simulator for VLSI Lithography and Etching Processes: Part 1—Application to Projection Lithography," *IEEE Trans. Elec. Dev.*, vol. ED-26, p. 717, 1979.

[5.25]. C. A. Mack, "PROLITH: A Comprehensive Optical Lithography Model," *Optical Microlithography IV, Proc. SPIE*, vol. 538, p. 207, 1985.

[5.26]. DEPICT User's manual, Dec. 1996, Avant! Inc.

[5.27]. ATHENA User's Manual, Oct. 1996, Silvaco Inc.

[5.28]. C. A. Mack, "Analytical Expression for the Standing Wave Intensity in Photoresist," *Appl. Optics*, vol. 25, p. 1958, 1986.

[5.29]. C. A. Mack, "Absorption and Exposure in Positive Photoresists," *Appl. Optics*, vol. 27, p. 4913, 1988.

[5.30]. F. H. Dill and J. M. Shaw, "Thermal Effects on the Photoresist AZ1350J," *IBM Journal Res. and Dev.*, vol. 21, p. 210, 1977.

[5.31]. C. A. Mack and R. T. Carback, "Modeling the Effects of Prebake on Positive Resist Processing," *Proc. of Kodak Microelectronics Seminar*, p. 155, 1985.

[5.32]. E. J. Walker, "Reduction of Photoresist Standing-Wave Effects by Postexposure Bake," *IEEE Trans. Elec. Dev.*, vol. ED-22, p. 464, 1975.

[5.33]. C. A. Mack, "New Kinetic Model for Resist Dissolution," *J. Electrochem. Soc.*, vol. 139, p. L35, 1992.

[5.34]. A. N. Broers, W. W. Molzen, J. J. Cuomo, and N. D. Wittels, "Electron Beam Fabrication of 80 Å Metal Structures," *Appl. Phys. Lett.*, vol. 29, 1976.

[5.35]. C. Y. Chang, G. Owen, R. F. Pease, and T. Kailath, "A Computational Method for the Correction of Proximity Effects in Electron-Beam Lithography, Electron-Beam, X-ray and Ion Beam Submicron Lithographies," *Proc. SPIE*, vol. 1671, p. 208, 1992.

[5.36]. S. D. Berger and J. M. Gibson, "New Approach to Projection-Electron Lithography with Demonstrated 0.1 μm Linewidth," *Appl. Phys. Lett.*, vol. 57, p. 153, 1990.

[5.37]. L. R. Harriott, "Preliminary Results From a Prototype Projection Electron-Beam Stepper —SCALPEL Proof-of-Concept System," *J. Vac. Sci. Tech.*, vol. B14, p. 3825, 1996.

[5.38]. E. Spiller, "Reflective Multilayer Coatings for the Far UV Region," *Appl. Optics*, vol. 15, p. 2333, 1975.

[5.39]. A. M. Hawryluk, N. M. Ceglio, and D. A. Markle, "EUV Lithography," *Solid State Tech.*, p. 151, 1997.

[5.40]. Y. Yoshikawa, O.Ochi, H. Nagai, and Y. Mizushima, "A Novel Inorganic Photoresist Utilizing Ag Photodoping in Se-Ge Glass Films," *Appl. Phys. Lett.*, vol. 29, p. 677, 1977.

[5.41]. D. C. Hofer, R. D. Miller, and C. G. Willson, "Polysilane Bilayer UV Lithography," *SPIE Proc.*, vol. 469, p. 16, 1984.

[5.42]. R. R. Kunz, P. A. Bianconi, M. W. Horn, R. R. Paladuga, D. C. Shaver, D. A. Smith, and C. A. Freed, "Polsilyne Resists for 193 nm Excimer Laser Lithography," *Advances in Resist Technology and Processing VIII, SPIE Proc.*, vol. 1446, p. 218, 1991.

[5.43]. D. E. Seeger, D. C. La Turlipe Jr., R. R. Kunz, C. M. Garza, and M. A. Hanratty, "Thin-Film Imaging: Past, Present, Prognosis," *IBM J. Res. and Dev.*, vol. 41, p. 105, 1997.

5.9 Problems

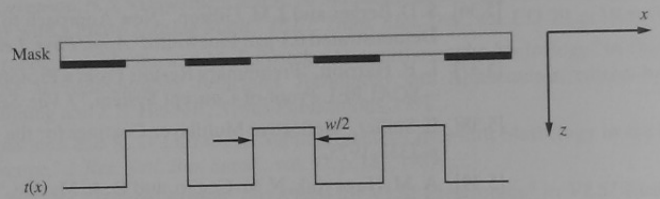
5.1. Calculate and plot versus exposure wavelength the theoretical resolution and depth of focus for a projection exposure system with an *NA* of 0.6 (about the best that can be done today). Assume $k_1 = 0.6$ and $k_2 = 0.5$ (both typical values). Consider wavelengths between 100 nm and 1000 nm (DUV and visible light). Indicate the common exposure wavelengths being used or considered today on your plot (g-line, i-line, KrF, and ArF). Will an ArF source be adequate for the 0.13-μm and 0.1-μm technology generations according to these simple calculations?

5.2. In a particular positive resist process it is sometimes noticed that there is difficulty developing away the last few hundred angstroms of resist in exposed areas. This sometimes causes etching problems because the resist remains in the areas to be etched. Suggest a possible cause of this problem and therefore propose a solution.

5.3. An X-ray exposure system uses photons with an energy of 1 keV. If the separation between the mask and wafer is 20 μm, estimate the diffraction-limited resolution that is achievable by this system.

5.4. Estimate the exposure wavelength that was used in the simulation example in Figure 5–44 in the text. Assume that the index of refraction of the photoresist is 1.68 (typical value).

5.5. In this chapter, we considered an example of an isolated space on a mask and saw that the Fourier transform of this pattern is the $\sin(x)/x$ function. This function describes the spatial variation of the light intensity that the objective lens must capture. An even more illustrative example of the effects of diffraction is given by a mask pattern consisting of a periodic pattern of lines and spaces as shown below.



Find the Fourier transform for this mask function and plot it as was done in the example in the chapter. The objective lens must capture at least the first diffraction order if it is going to resolve features with a pitch p. Use this criterion to derive an equation like Eq. (5.2) for the resolution of such a system.

5.6. Assume that Figure 5–46 in the text was experimentally determined for a 0.6-μm-thick resist with an index of refraction of 1.68. Estimate the Dill resist parameters from the data in this figure.

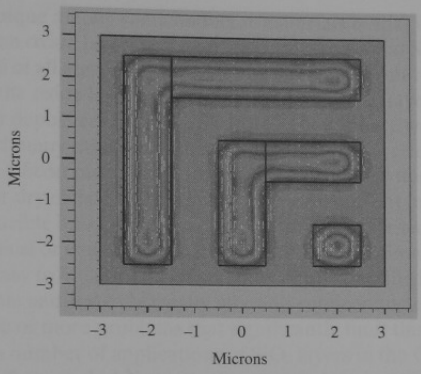
5.7. Lithography often has to be done over underlying topography on a silicon chip. This can result in variations in the resist thickness as the underlying topography goes up and down. This can sometimes cause some parts of the photoresist image to be underexposed and/or other regions to be overexposed. Explain in terms of the chemistry of the resist exposure process why these underexposure and overexposure problems occur.

5.8. As described in this chapter, there are no clear choices for lithography systems beyond optical projection tools based on 193-nm ArF eximer lasers. One possibility is an optical projection system using a 157-nm F_2 excimer laser.

a. Assuming a numerical aperture of 0.8 and $k_1 = 0.75$, what is the expected resolution of such a system using a first order estimate of resolution?

b. Actual projections for such systems suggest that they might be capable of resolving features suitable for the 2009 0.07-μm generation. Suggest three approaches to actually achieving this resolution with these systems.

5.9. Current optical projection lithography tools produce diffraction-limited aerial images. A typical aerial image produced by such a system is shown in the simulation below where a square and rectangular mask regions produce the image shown. (The mask features are the black outlines, the calculated aerial image is the gray scale inside the black rectangles.) The major feature of the aerial image is its rounded corners compared to the sharp square corners of the desired pattern. Explain physically why these features look the way they do, using diffraction theory and the physical properties of modern projection optical lithography tools.



5.10. Future optical lithography systems will likely use shorter exposure wavelengths to achieve higher resolution and they will also likely use planarization techniques to provide "flat" substrates on which to expose the resist layers. Explain why "flat" substrates will be more important in the future than they have been in the past.