Silicon VLSI Technology

Fundamentals, Practice and Modeling

James D. Plummer
Michael Deal
Peter B. Griffin
Department of Electrical Engineering
Stanford University

Prentice Hall
Upper Saddle River, NJ 07458

Library of Congress Cataloging-in-Publication Data
Silicon VLSI technology
p. cm.
ISBN 0-13-085037-3
1. Integrated circuits—Very large scale integration—Design and construction. 2. Silicon. 3. Silicon oxide films. 4. Metal oxide semiconductors. 5. Silicon technology.
TK7874.75.S54 2000
621.39'5—dc21 99-42745
CIP

Publisher: Tom Robbins
Associate Editor: Alice Dworkin
Editorial/Production Supervision: Rose Kernan
Vice President and Editorial Director, ECS: Marcia Horton
Vice President of Production and Manufacturing: David W. Riccardi
Executive Managing Editor: Vince O'Brien
Marketing Manager: Danny Hoyt
Managing Editor: David A. George
Manufacturing Buyer: Pat Brown
Manufacturing Manager: Trudy Pisciotti
Art Director: Jayne Conte
Cover Design: Bruce Kenselaar
Editorial Assistant: Jesse Power
Copy Editor: Martha Williams
Composition: D&G Limited, LLC

Printed in the United States of America

10 9 8 7 6 5 4 3 2

ISBN 0-13-085037-3

Prentice Hall International (UK) Limited, *London*
Prentice Hall of Australia Pty. Limited, *Sydney*
Prentice Hall Canada Inc., *Toronto*
Prentice Hall Hispanoamericana, S.A., *Mexico*
Prentice Hall of India Private Limited, *New Delhi*
Prentice Hall of Japan, Inc., *Tokyo*
Pearson Education Pte., Ltd., *Singapore*
Editora Prentice† Hall do Brasil, Ltda., *Rio de Janeiro*

Contents

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/m²)	100	80	60	50	40	30
@ Defect Size (nm)	@ 80	@ 60	@ 40	@ 30	@ 20	@ 15
DRAM Chip Size (mm²)	280	400	560	790	1120	1580
MPU Chip Size (mm²)	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

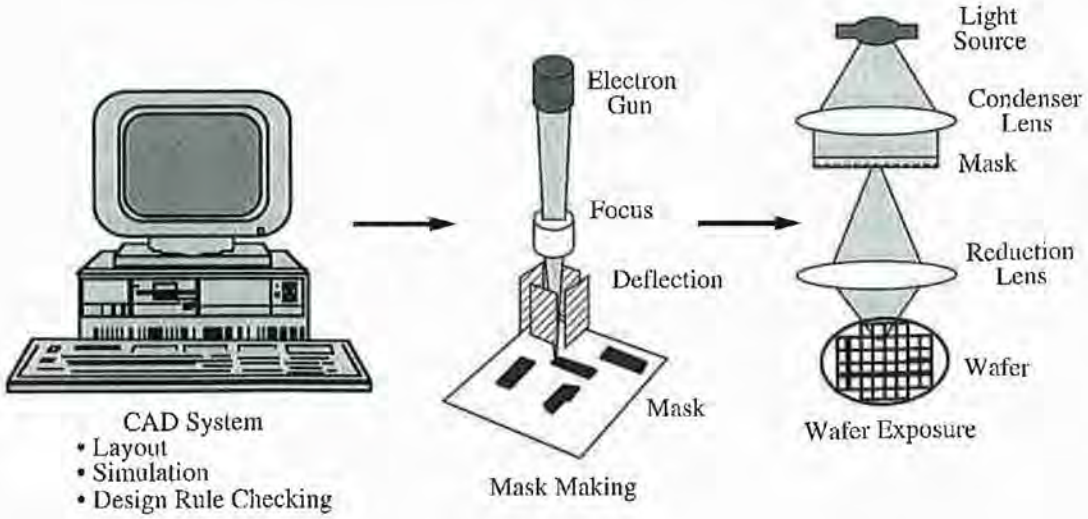

Figure 5–1 Lithography process from mask design to wafer printing.

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

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].

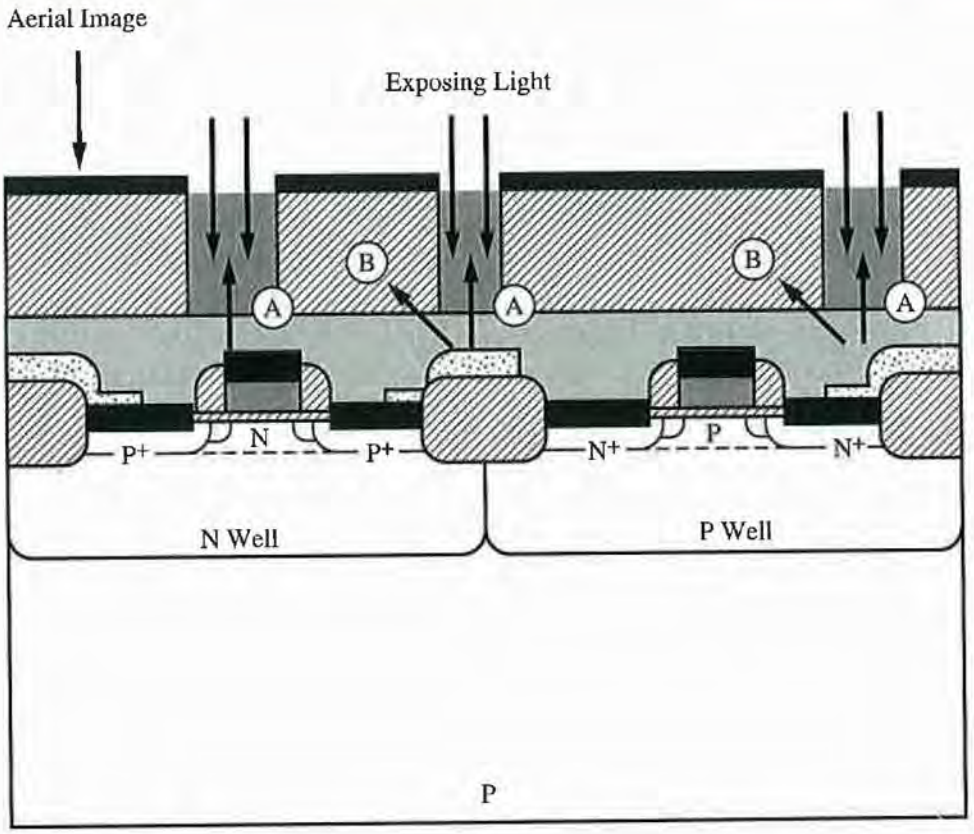

Figure 5–23 Photolithography exposure occurring between Figures 2–39 and 2–40 in the CMOS process described in Chapter 2. The A and B symbols indicate examples of standing wave patterns and sideways light scattering, respectively.

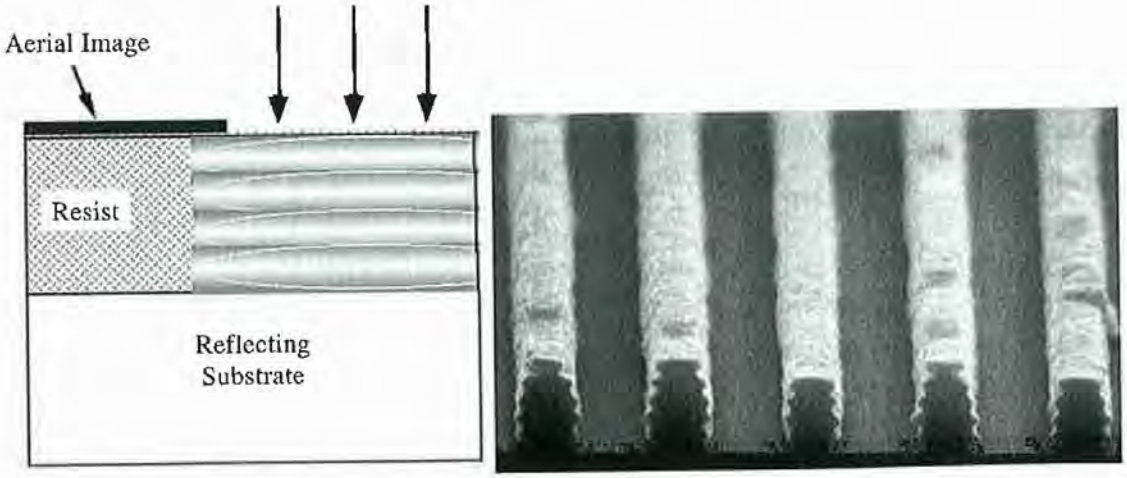

Figure 5–24 The light intensity pattern resulting from a standing wave is sketched on the left. A photoresist image that illustrates these effects after developing is shown on the right. Courtesy of A. Vladar and P. Rissman, Hewlett Packard.

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-

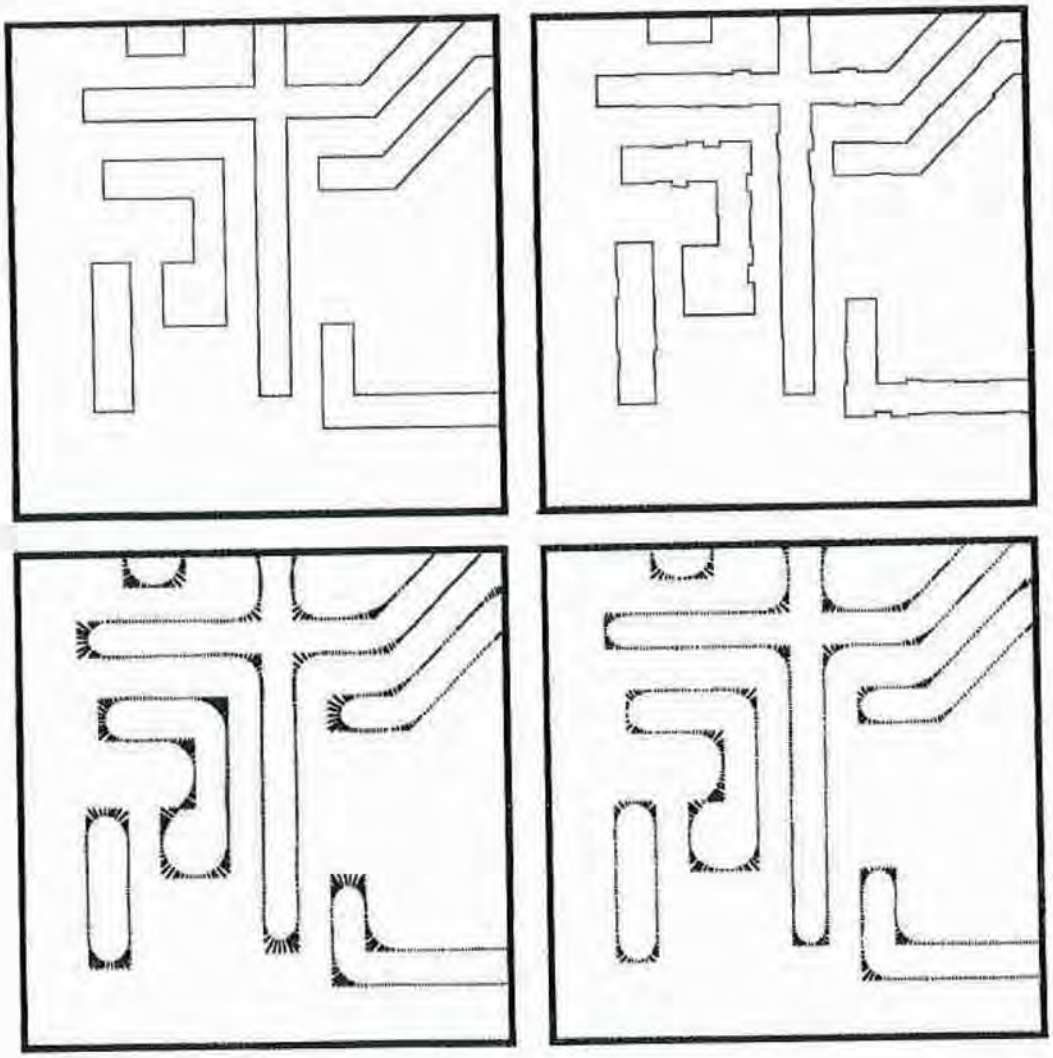

Figure 5–25 Mask patterns with (right) and without (left) OPC are shown on the top. The corresponding aerial images (calculated) are shown on the bottom. Note the improvement in the quality of the aerial image when OPC is used. The dark lines in the bottom patterns indicate the difference between the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

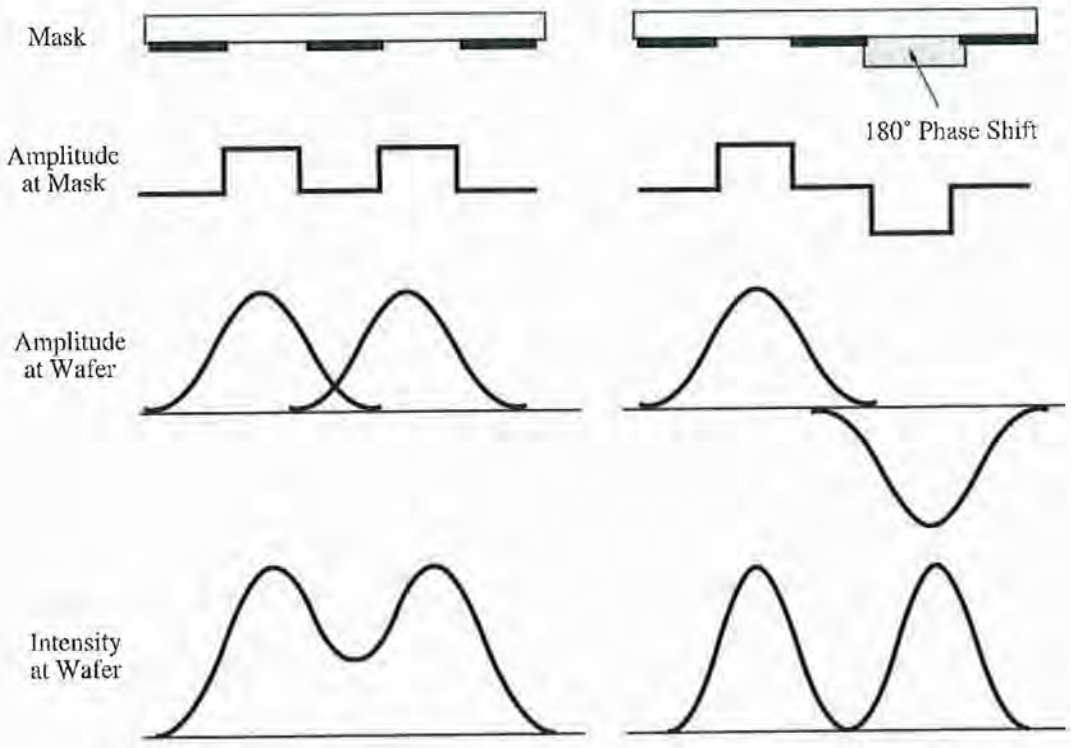

Figure 5–26 Example of the use of phase shifting techniques on masks to improve the resolution of the aerial image. (After Levenson et al. [5.32].)

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

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]:

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

$R(x,y,z)$ is the local development or etching rate in $\mu m\ min^{-1}$, 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

$$F_1 = k_D(C_D - C_S) \quad (5.62)$$

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-

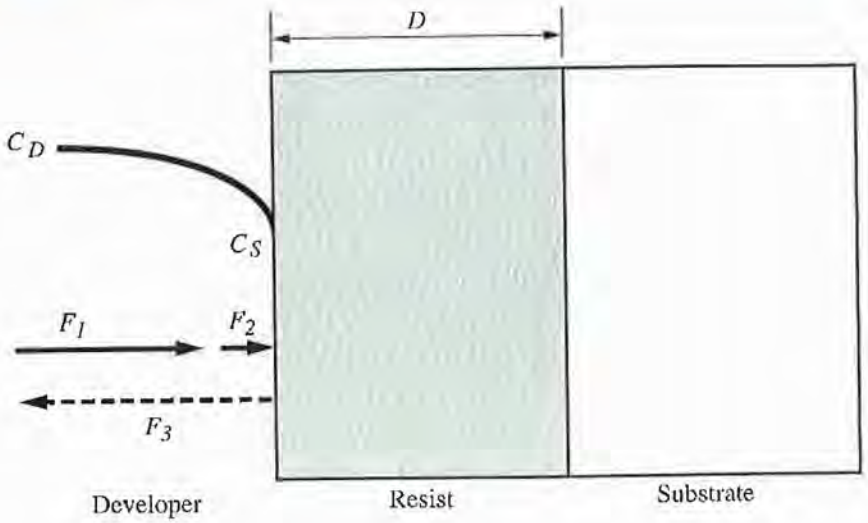

Figure 5–49 Mack's model for photoresist development.

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

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

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 postexposure 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.

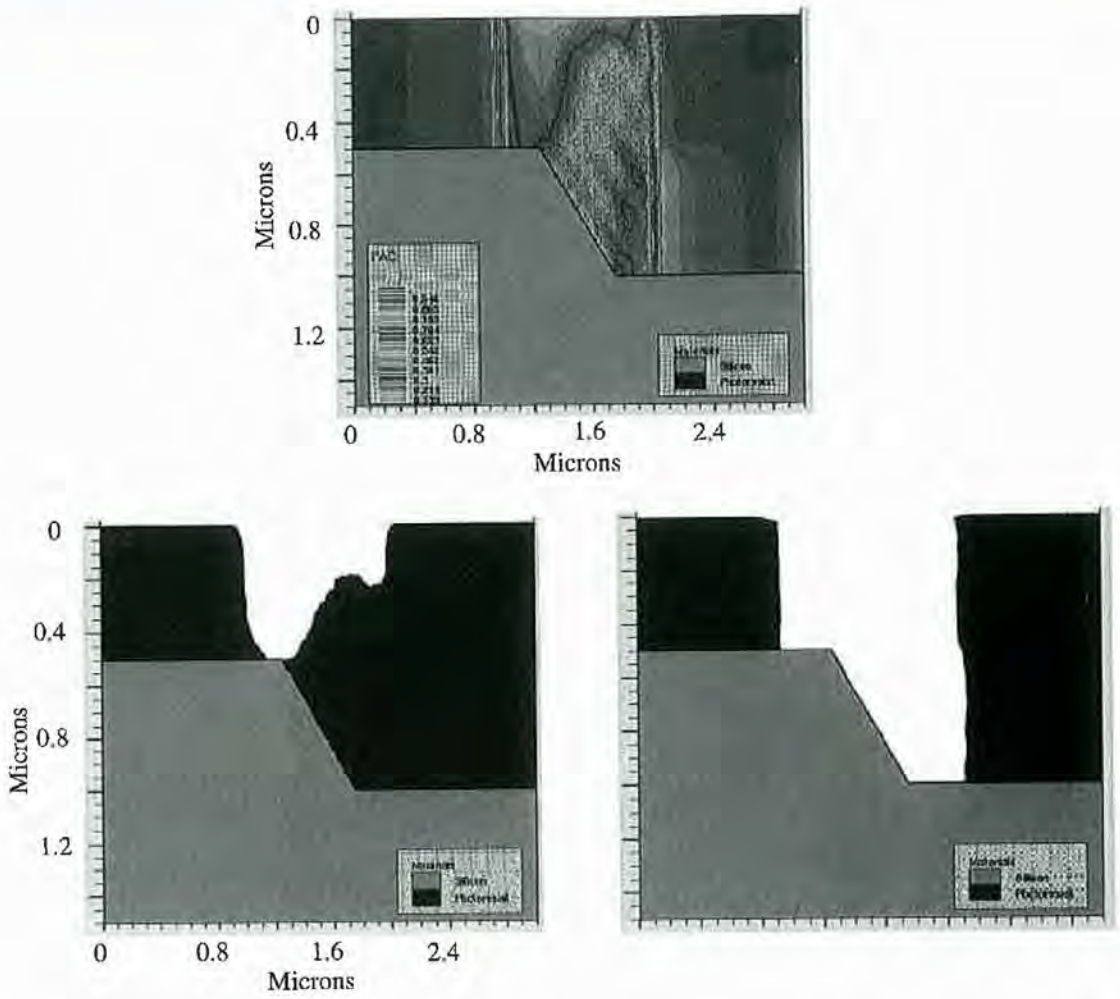

Figure 5–50 Example of the calculation of a developed photoresist layer using the ATHENA simulator by Silvaco. The resist was exposed with a dose of 200 mJ cm^{-2}, a postexposure bake of 45 min at 115°C was used, and the pattern was developed for 60 seconds, all normal parameters. The Dill development 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

Etching

10

10.1 Introduction

After thin films are deposited on the wafer surface, they are usually selectively removed by etching to leave the desired pattern of the film on the wafer surface. This process was used many times in the CMOS process flow in Chapter 2 and is summarized in Figure 10–1. In addition to deposited films, parts of the silicon substrate itself may be etched, such as in creating trenches in isolation structures. The masking layer may be photoresist, or it may be another thin film such as silicon dioxide or silicon nitride. Oxide or nitride masks stand up better to etching conditions than photoresist and are often called hard masks. But they themselves must be selectively etched, usually using lithographically defined photoresist as the masking layer. The etching of a thin film is usually done until a different layer is reached underneath. Etching only partway through a layer without reaching an underlying material is sometimes done but can be difficult to control uniformly over the whole wafer. Multilayer structures can be etched sequentially using the same masking layer.

Etching can be done in either a "wet" or "dry" environment. Wet etching involves the use of liquid etchants. The wafers are usually immersed in the etchant solution and the exposed material is etched mostly by chemical processes. While such processes are not common in modern VLSI processing for reasons we shall see, they are sometimes used, as in Figure 2–34 in Chapter 2. Dry etching involves the use of gas-phase etchants in a plasma. Here the etching usually takes place by a combination of chemical and physical processes. Because a plasma is involved, dry etching is often called "plasma etching." This is normally the etching method today and was used many times in Chapter 2 (Figure 2–4 for example).

The ideal etch profile in Figure 10–1 has perfectly straight sidewalls exactly under the edge of the mask. What can happen in reality is shown in Figure 10–2. In Figure 10–2(a) the etching is both vertical and lateral. This leads to undercutting of the photoresist mask and nonvertical sidewalls in the film. In Figure 10–2(b) the photoresist is not perfectly rectangular, with rounded top corners and sloped sides. In addition, the etch rate of the photoresist is nonzero, and the mask itself is etched, or eroded, to some degree. This occurs both on its top and on its sides. This leads to even more lateral etching of the film since the photoresist mask gets narrower as the etching continues. If the substrate is attacked by the etchant as well, then its profile will change. Figure 10–2(b) also illustrates this issue, which is related to the selectivity of the etching process.

Etch selectivity is the ratio of the etch rates of the different materials in an etch process. When the etch rates of the mask and the underlying substrate are near zero,

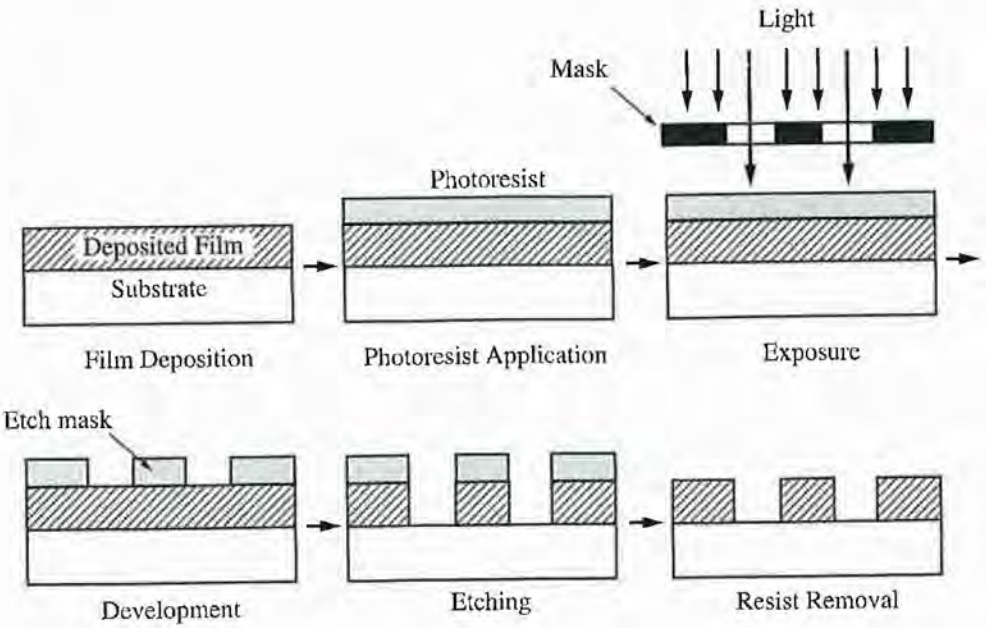

Figure 10–1 General process used in integrated circuit fabrication to define patterned films.

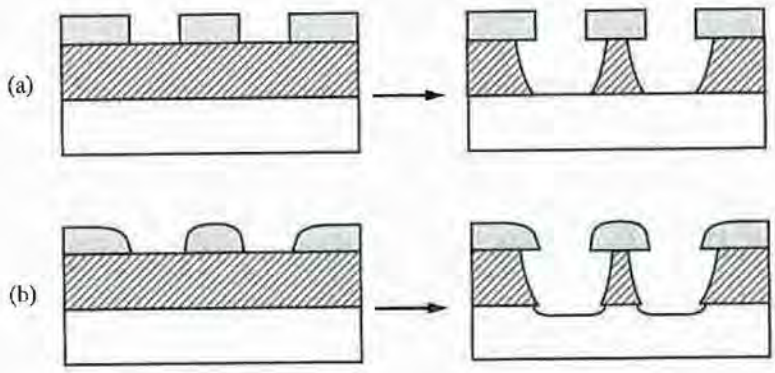

Figure 10–2 Actual etch profiles that can occur. (a) Lateral etching under mask. (b) Rounded photoresist which is further eroded during etching, leading to even more lateral etching. (b) also illustrates etch selectivity.

while the etch rate of the film is appreciable, then the etch selectivity of the film with respect to both the mask and the substrate is high. This is normally a desired situation. If the etch rate of the mask or substrate is significant, then the selectivity is poor. Selectivities, or etch rate ratios, in the range of 25–50 are usually considered reasonable. Different materials have different etch rates usually because of the chemical effects involved with etching, as opposed to the physical effects. Chemical reactions can be very selective whereas physical processes like sputtering tend to be similar for most materials.

Etch directionality is a measure of the relative etch rates in different directions, usually vertical versus lateral, as illustrated in Figure 10–3. Isotropic etching occurs when

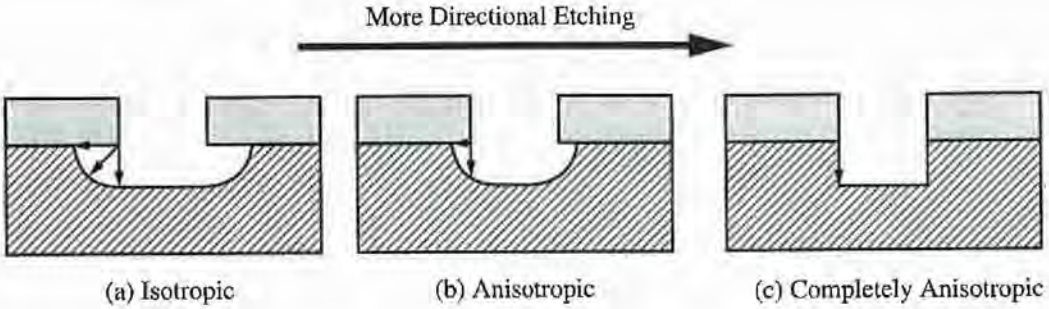

Figure 10–3 Etch profiles for different degrees of anisotropic, or directional, etching: (a) purely isotropic etching; (b) anisotropic etching; (c) completely anisotropic etching.

etch rates are the same in all directions, as in Figure 10–3(a). The etch distance laterally under each side of the mask is the same as vertically under the mask opening. As the opening under the mask approaches zero, the etch profile would approach a perfect semicircle. Figure 10–3(b) shows anisotropic etching, where less etching occurs laterally. Figure 10–3(c) shows completely anisotropic etching, with no etching at all in the lateral direction. Etch directionality is usually related to physical effects in etching such as ion bombardment and sputtering. As we shall see, the more physical an etch process is, the more directional and the less selective it generally is. On the other hand, the more chemical an etch process is, the more selective and the less directional it generally is.

Directionality is often desired in an etch process. Vertical etching maintains lithographically defined feature sizes as they are transferred by etching into underlying films. Vertical etching is also utilized in special applications such as sidewall spacer formation. (Recall Figure 2–30 in the CMOS process flow in Chapter 2.) However, vertical structures that result from vertical etching may also cause problems in subsequent processing. Depositing another layer on top of a structure with vertical steps may cause step coverage problems, as we discussed in Chapter 9. Less steep steps, obtained with less directional etching, would make it easier to obtain good step coverage and would allow easier filling in later deposited layers. Because of shrinking feature sizes, however, modern fabrication methods tend to employ directional etching and vertical steps with little undercutting, and use deposition techniques that achieve good filling and step coverage even for near-vertical structures.

Selectivity is also a requirement. This means that the etch rate of the layer to be etched should be fast compared to the etch rate of the mask and etch rate of the substrate or layer below. Achieving good selectivity and directional etching at the same time is sometimes quite difficult. Other requirements for an etch process are that the gases and liquids used for the etching must be able to be transported to the wafer surface, and the byproducts of the etch process must be transported away easily. The latter means that soluble byproducts are produced in wet etching, and that gaseous or volatile byproducts are produced in dry etching. The etch process should be uniform, safe to use, and should cause minimal damage to structures on the chip. Finally, the etching process should be clean, with low particulate production and little film contamination, and should be cost effective.

10.2 Historical Development and Basic Concepts

There are two main methods of etching used in the semiconductor industry: wet etching and dry, or plasma, etching. In the early days of the industry, wet etching was used exclusively. It was a well-established, simple, and inexpensive technology. Wet etching can also be very selective. Eventually, however, the need for smaller linewidths and more vertical structures required new techniques. Plasma etching methods were developed for integrated circuit fabrication and are used for most etching steps today.

In this section we will describe the basic concepts involved in both of these techniques. We will start with wet etching, which we will use to introduce concepts such as selectivity, overetch, etch bias, and anisotropy that are applicable to both wet and dry etching.

10.2.1 Wet Etching

The first etchants used in the integrated circuit industry were simple wet chemical etchants. By immersing the wafers into baths of liquid chemicals, the exposed films could be etched away, leaving unetched those regions of the film that were masked with photoresist or other films. Wet etches were developed for all steps in the fabrication process. A common wet etch for SiO_2 is hydrofluoric acid, HF. The overall reaction is

$$SiO_2 + 6HF \rightarrow H_2SiF_6 + 2H_2O \tag{10.1}$$

Wet etchants work through chemical processes rather than physical processes. In general, wet etchants work by chemically reacting with the film to form water-soluble byproducts or gases, such as the H_2SiF_6 byproduct which is a water-soluble complex in the above example.

In some cases the etching occurs by first oxidizing the surface of the film or material and then dissolving the oxide. For example, a common etchant for silicon is a mixture of nitric acid (HNO_3) and HF. The nitric acid partially decomposes to nitrogen dioxide, NO_2, which then oxidizes the surface of the silicon by the reaction

$$Si + 2NO_2 + 2H_2O \rightarrow SiO_2 + H_2 + 2HNO_2 \tag{10.2}$$

The HF then dissolves the SiO_2 by the reaction given in Eq. (10.1). The overall reaction is

$$Si + HNO_3 + 6HF \rightarrow H_2SiF_6 + HNO_2 + H_2O + H_2 \tag{10.3}$$

Buffering agents are often added to the etch solutions to keep the etchants at maximum strength over use and time. For example, ammonium fluoride (NH_4F) is added to HF to help prevent depletion of the fluoride ions in the oxide etch. This is called "Buffered HF" or BHF, or more commonly BOE for "Buffered Oxide Etch." The addition of ammonium fluoride also decreases the etch rate of photoresist and helps to minimize lifting of the resist during oxide etching. Acetic acid, CH_3COOH, is often added to the nitric acid/hydrofluoric acid silicon etch to limit the dissociation of the nitric acid.

Wet etches can be very selective because they depend on chemistry. The limiting steps in most etch processes are usually the chemical reactions in which the etch species react

with the film, forming the soluble byproducts. The selectivity, S, of an etch process between two materials, 1 and 2, is simply the ratio of their etch rates, r, in that etchant, or

$$S = \frac{r_1}{r_2} \tag{10.4}$$

Material 1 is usually the film being etched, and material 2 is either the masking material (photoresist or perhaps silicon dioxide) or the material below the film. We often speak of an etch process for material 1 having a certain selectivity over (or "with respect to," or just "to") material 2. For a large selectivity ($S>>1$) we say that the etch process for material 1 has good selectivity over material 2.

Selectivity is important for a number of reasons. First, the selectivity for the material to be etched with respect to the mask determines how thick the mask material must be so that the mask is not completely removed during the etching process. Second, all etch processes generally employ some amount of overetching in order to ensure that the desired etching process has gone to completion. Thus underlying films are normally subjected to the etchant for some period of time toward the end of the etch period. The selectivity for the material being etched with respect to the underlying material determines how much etching of the underlying materials occurs.

Example 1 μm of SiO_2 is being etched on top of a Si substrate. The etch rate of the oxide is 0.40 μm min^{-1}, and the etch selectivity of the oxide with respect to the silicon is 25-to-1. If the etch is done for three minutes, how much of the underlying Si is etched?

Answer From Eq. (10.4), the etch rate of the Si can be calculated:

$$S = \frac{r_1}{r_2} \Rightarrow 25 = \frac{0.40\ \mu m\ min^{-1}}{r_{Si}} \Rightarrow r_{Si} = 0.016\ \mu m\ min^{-1}$$

The Si is only exposed to the etch after the overlying oxide is completely etched off. The time required to etch away the oxide equals 1 μm/0.40 μm min^{-1}, which equals 2.5 minutes. Therefore, the Si is etched for 0.5 minutes. Hence, the amount of Si etched is 0.016 μm min^{-1} * 0.5 min, which equals 0.008 μm, or 8 nm.

Most wet chemical etchants etch isotropically. The exceptions to this are those which are sensitive to crystallographic orientation. For example, some etchants etch much slower in the <111> direction in crystalline silicon as compared to the other directions. But most commonly used wet etches etch equally in all directions. In Figure 10–3(a) we saw how isotropic etching results in etching underneath the edge of the mask. The amount of undercutting is called the "etch bias." This is shown in Figure 10–4(a). Here we assume that the selectivity of the etch with respect to both the mask and the substrate is infinity (i.e., no etching at all of those materials). The etch thickness or depth is

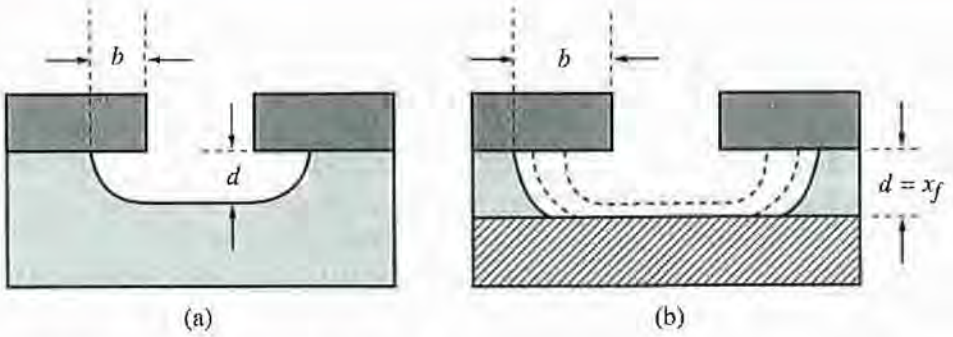

Figure 10–4 Illustration of etch bias and overetch. In (a) the etch bias, b, is shown for a given etch depth, d. In (b), overetching is illustrated where etching is continued even after the etch depth, d, equals the film thickness, x_f, with the result that the etch bias increases.

d and the bias is b. For perfectly isotropic etching b would equal d, but only if etching is stopped before the bottom of the film is reached.

As mentioned earlier, some overetching is normally part of every etch process. If each layer were of uniform thickness, and the etch were perfectly uniform through time and space, then overetching would theoretically not be needed. The etch time required would be based on the previously determined etch rate or on other endpoint detection methods, to be discussed later. It would not even matter what the etch selectivity over the layer beneath is in that case. However, the thickness of films is not perfectly uniform, and they do not always etch exactly uniformly. So some overetch is always done in manufacturing to ensure that the etch goes to completion everywhere. This is illustrated in Figure 10–4(b). In this figure we again assume that the selectivity with respect to the mask and substrate are infinite and that the etching is purely isotropic. The bias b now increases with time as the overetch continues. The amount of overetching needed, usually measured in terms of time or % time, can be determined by estimating the uncertainties in etch rates and in the nonuniformities of the thicknesses and then calculating the worst case etch time needed. Ten to twenty percent overetches are common. This ensures that while some of the structures may be overetched, none are underetched. We will see later that when the etching is anisotropic, overetching is also required to remove residual film from steps in the topography.

Example A 0.5-μm layer of silicon dioxide on a Si substrate needs to be etched down to the Si. Assume that the nominal oxide etch rate is r_{ox} (in μm min^{-1}). There is a ± 5% variation in the oxide thickness and a ± 5% variation in the oxide etch rate. (a) How much overetch is required (in % time) in order to ensure that all the oxide is etched? (b) What selectivity of the oxide etch rate to the Si etch rate is required so that a maximum of 5.0 nm of Si is etched? (Assume the overetch calculated in part (a) is done.)

Answer (a) The nominal etch time = $x/r_{ox} = 0.5/r_{ox}$ (x = thickness in μm.) The overetch is done to make sure all the oxide is etched for the worst case condition; that means for the thickest oxide and the slowest etch rate. The thickest oxide =

0.5 * 1.05 = 0.525 μm. The slowest etch rate = r_{ox} * 0.95 μm min^{-1}. The time to etch the worst case = x/r = 0.525/(r_{ox} * 0.95) = 0.553/r_{ox} min. The amount of overetch, in % time, is just the worst case etch time divided by the nominal etch time = (0.553/r_{ox})/(0.5/r_{ox}) = 1.106 or 10.6% overetch.

(b) The maximum amount of Si that will be etched will occur under the thinnest oxide being etched at the fastest etch rate. The thinnest oxide is 0.95 * 0.5 μm = 0.475 μm. The fastest etch rate is r_{ox} * 1.05 μm min^{-1}. The time to etch through that oxide is x/r = 0.475/(r_{ox} * 1.05) = 0.4523/r_{ox} min. The time that the Si is exposed to the etch is equal to the total etch time (0.553/r_{ox}, as calculated in part (a)) minus the time it took to etch the thinnest oxide, etching at the fastest rate, or

$$(0.553/r_{ox}) - (0.4523/r_{ox}) = 0.1003/r_{ox} \text{ min}$$

We want to etch 0.005 μm or less of the Si, so

$$r_{Si} = x/\text{time} = 0.005/(0.1003/r_{ox}) \text{ μm min}^{-1}$$
$$= 0.050 * r_{ox} \text{ μm min}^{-1}$$

The selectivity of oxide etch rate to Si etch rate is therefore:

$$r_{ox}/r_{Si} = r_{ox}/(0.050 * r_{ox})$$
$$r_{ox}/r_{Si} = 20{:}1$$

Since overetching is usually done, it is important that the selectivity with respect to the layer below be as high as possible or else some etching of that layer will occur during the overetch. On the other hand, if the selectivity is very large, one cannot indiscriminately overetch. This would cause unwanted undercutting or bias.

The above examples assume infinite etch selectivity of the film with respect to the mask; that is, the etch rate of the mask is zero. But there will be some finite etch rate of the mask, which is usually photoresist. What will happen? Obviously the mask will etch at some rate during the etching of the film. This is shown in Figure 10–5(a) for the case of isotropic etching and a rectangular-shaped mask. Here Δm is the amount of mask etched, the same in all directions in this case. This is called mask erosion. Obviously the extent of lateral etching will increase beyond what would be expected without any mask erosion. For completely anisotropic (vertical) etching, etch erosion should not be a problem as long as the mask is perfectly rectangular in shape. The etch will only occur from the top. Only if the mask is etched all the way through would any difficulties arise.

In fact, the mask shape is usually not perfectly rectangular but often has rounded corners or sloped sides. Mask erosion leading to extra undercutting can occur even for completely anisotropic etching, as illustrated in Figure 10–5(b). With a finite etch rate for the mask, there will be equal vertical etching at all points of the mask. At the sloped edge, a finite amount of the mask will be etched away. This will allow lateral erosion of

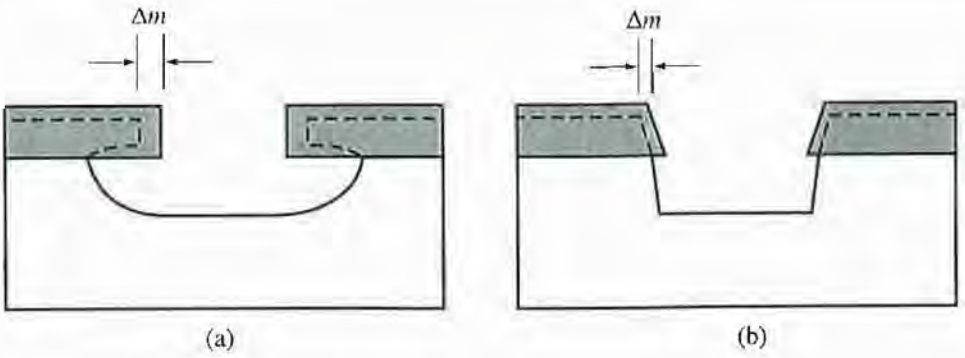

Figure 10–5 Illustration of mask erosion. In (a), there is an idealized rectangular mask and isotropic etching of the mask and the substrate. In (b), there is a sloped mask and anisotropic etching. In both cases more lateral etching of the substrate occurs as a result of the mask etching, or erosion.

the mask, shown by Δm in the figure, and etching of the film in this thin region will occur. This will result in undercutting with respect to the original mask edge, even for perfectly anisotropic, or vertical, etching.

Etching is often neither perfectly anisotropic nor perfectly isotropic, as was shown in Figure 10–3(b). We can define the degree of anisotropy of a film, A_f, as

$$A_f = 1 - \frac{r_{\text{lat}}}{r_{\text{ver}}} \tag{10.5}$$

where r_{lat} and r_{ver} are the etch rates in the lateral and vertical directions, respectively. In Figure 10–4, this would correspond to

$$A_f = 1 - \frac{b}{d} \tag{10.6}$$

Isotropic etching [Figure 10–3(a)] would have an anisotropy of 0, completely anisotropic etching [Figure 10–3(c)] would have an anisotropy of 1, and general anisotropic etching [Figure 10–3(b)] would have an anisotropy between 0 and 1. For a structure etched exactly to the bottom of the film, with no overetch, the anisotropy would be

$$A_f = 1 - \frac{b}{x_f} \tag{10.7}$$

where b is the bias and x_f is the film thickness.

Example Consider the structures shown in Figure 10–6. A 0.5-μm-thick oxide layer is etched to achieve equal structure widths and spacings (S_f). The etch process produces

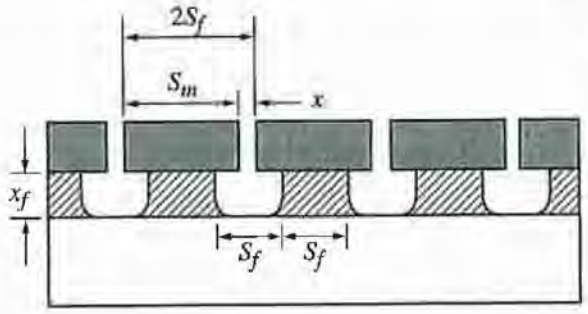

Figure 10–6 Structure and dimensions needed to fabricate features with equal widths and spaces (S_f) for a mask size of S_m, film thickness of x_f, and degree of anisotropy of A_f. x is the spacing between the masks and is usually the minimum feature size. (After [10.1].)

a degree of anisotropy of 0.8. If the distance between the mask edges, x, is 0.35 μm, what structure spacings and widths are obtained? Neglect any overetch.

Answer To obtain equal widths and spacings, S_f, the mask width, S_m, must be made larger to take into account the anisotropic etching. Since

$$S_m = S_f + 2b$$

where b is the bias on each side, and using Eq. (10.6) we can now rewrite S_m as

$$S_m = S_f + 2x_f(1 - A_f)$$

which is the dimension of the mask needed. [This makes sense. For isotropic etching ($A_f = 0$), S_m is a maximum, and for completely anisotropic etching ($A_f = 1$), S_m is a minimum and equal to S_f.] The distance between the mask edges, x, is usually the minimum feature size that can be lithographically resolved. Since

$$x = 2S_f - S_m$$

x can be rewritten, using the expression above for S_m, as

$$x = S_f 2x_f(1 - A_f)$$

Rewriting this in terms of S_f gives

$$S_f = x + 2x_f(1 - A_f) \tag{10.8}$$

Plugging the above numbers into Eq. (10.8) gives

$$\begin{aligned} S_f &= 0.35\mu m + 2(0.5\mu m)(1 - 0.8) \\ &= 0.55\ \mu m \end{aligned}$$

Eq. (10.8) shows that the structure size approaches the minimum lithographic dimension only when the film thickness gets very small (not always practical) or as the anisotropy gets close to 1. This means completely vertical etching. Unfortunately, wet etching does not result in vertical etching, but rather isotropic etching where A_f is a minimum. These considerations give some insight into the reasons wet chemical etching is not very commonly used in chip fabrication today.

Table 10–1 Common wet chemical etchants for various thin films used in IC fabrication

Material	Etchant	Comments
SiO_2	HF (49% in water) "straight HF"	Selective over Si (i.e., will etch Si very slowly in comparison). Etch rate depends on film density, doping.
	NH_4F:HF (6:1) "Buffered HF" or "BOE"	About $^1/_{20}$ th the etch rate of straight HF. Etch rate depends on film density, doping. Will not lift up photoresist like straight HF.
Si_3N_4	HF (49%)	Etch rate depends strongly on film density, O, H in film.
	H_3PO_4:H_2O (boiling @ 130–150°C)	Selective over SiO_2. Requires oxide mask.
Al	H_3PO_4:H_2O:HNO_3:CH_3COOH (16:2:1:1)	Selective over Si, SiO_2, and photoresist.
Polysilicon	HNO_3:H_2O:HF (+ CH_3COOH) (50:20:1)	Etch rate depends on etchant composition.
Single crystal Si	HNO_3:H_2O:HF (+ CH_3COOH) (50:20:1)	Etch rate depends on etchant composition.
	KOH:H_2O:IPA (23 wt. % KOH, 13 wt. % IPA)	Crystallographically selective; relative etch rates: (100): 100 (111): 1
Ti	NH_4OH:H_2O_2:H_2O (1:1:5)	Selective over $TiSi_2$.
TiN	NH_4OH:H_2O_2:H_2O (1:1:5)	Selective over $TiSi_2$.
$TiSi_2$	NH_4F:HF (6:1)	
Photoresist	H_2SO_4:H_2O_2 (125°C)	For wafers without metal.
	Organic strippers	For wafers with metal.

Before we describe techniques which do result in more anisotropic etching, we will look at some of the wet etch solutions that have been used, and are still used to some extent, in the Si integrated circuit industry. Table 10–1 gives a brief sampling of some of the common solutions and compositions.

The etch rate of a material in a particular etchant depends on the specific composition of the etch solution, and the etch rate can vary by orders of magnitude for different compositions of the same constituents. The rates of chemical reactions are strong functions of temperature, so the etch rates using wet, or chemical, etchants should be functions of temperature as well. For example, the etch rate of silicon dioxide in a 6:1 Buffered Oxide Etch (BOE) can change by a factor of two for a 10°C change in temperature [10.2].

Etch rates also depend on the composition and density of the films. For example, the etch rate of undoped, undensified CVD silicon dioxide in BOE is about 5 nm sec^{-1}, while the etch rate of phosphorus doped (8 atomic %) CVD silicon dioxide is about 9 nm sec^{-1}. On the other hand, the etch rate of densified (annealed at 800–1000°C), undoped CVD silicon dioxide is only about 1.5 nm sec^{-1}, the same as for thermal oxide. Clearly, the composition of the etchant and film, the density of the film, and the etch solution all affect the wet etch rates significantly.

The etch rate can also depend on the crystallographic orientation in some cases. In Table 10–1, the etch rates for two crystallographic directions in Si using KOH are given.

One sees that the etch rate for this etchant in the [100] direction is much faster than in the [111] direction, because the (111) planes are the most closely packed and etch more slowly. This can be used to form v-grooves on the surface of a (100) Si wafer because the {111} planes intersect the surface at a 54.7° angle.

10.2.2 Plasma Etching

While wet chemical etching was successfully used in the early days of integrated circuit manufacturing, it has been largely replaced by plasma, or dry, etching. There are two main reasons for this.

The first reason is that very reactive chemical species are produced in a plasma which can often etch more vigorously than species in a nonplasma environment. One of the first widespread uses of plasma etching was in the early 1970s to etch PECVD silicon nitride, used as the final passivation layer in integrated circuits. This film is difficult to etch with wet chemical etchants. Wet HF can etch it, but the etch rate is quite low, and the process is not selective with respect to SiO_2 which is often underneath. Boiling phosphoric acid was commonly used instead. But this etchant often lifts off photoresist masks so that a mask of SiO_2 has to be used, requiring its own masking and etching steps. A better etch process was sought. It was soon discovered that by using a CF_4/O_2 gas mixture in a plasma atmosphere, atomic fluorine is produced which can etch the nitride film much more easily. In addition, a photoresist mask is not lifted off in this process, making the process much simpler.

The second, and more important, reason why plasma etching has supplanted wet etching is the fact that directional or anisotropic etching is possible with plasma etch systems. Directional etching is needed to minimize underetching and etch bias, which then allows smaller and more tightly packed structures. This directional etching is due to the presence of ionic species in the plasma and the electric fields that direct them normal to the wafer surface.

Plasma etch systems can be designed so that either the reactive chemical components or the ionic components dominate. In many cases, a combination of ionic and reactive chemical species, usually acting in a synergistic manner, are utilized. The net etch rate can be much faster than the sum of the individual etch rates when they are acting alone. In addition, by utilizing both components this way, directionality can be achieved while maintaining an acceptable degree of selectivity.

A basic plasma etch system is shown schematically in Figure 10–7. This is similar to the PECVD and RF sputtering systems shown earlier in Figures 9–14 and 9–28. In these etch systems a low-pressure (1 mtorr – 1 torr) gas is used in the chamber. By applying a high electric field across two electrodes, some of the gas atoms are ionized, producing positive ions and free electrons and creating a plasma, as discussed in Chapter 9. In plasma etching systems, the energy is supplied by an RF generator, usually operating at 13.56 MHz. As we saw in Chapter 9, a voltage bias develops between the plasma and the electrodes. This is due to the difference in mobility of the electrons and the ions. Initially, the more mobile electrons are lost to the electrodes at a faster rate than the slower ions. This results in the plasma being biased positively with respect to the electrodes. For a symmetric RF plasma system, with the two electrodes of equal area, the

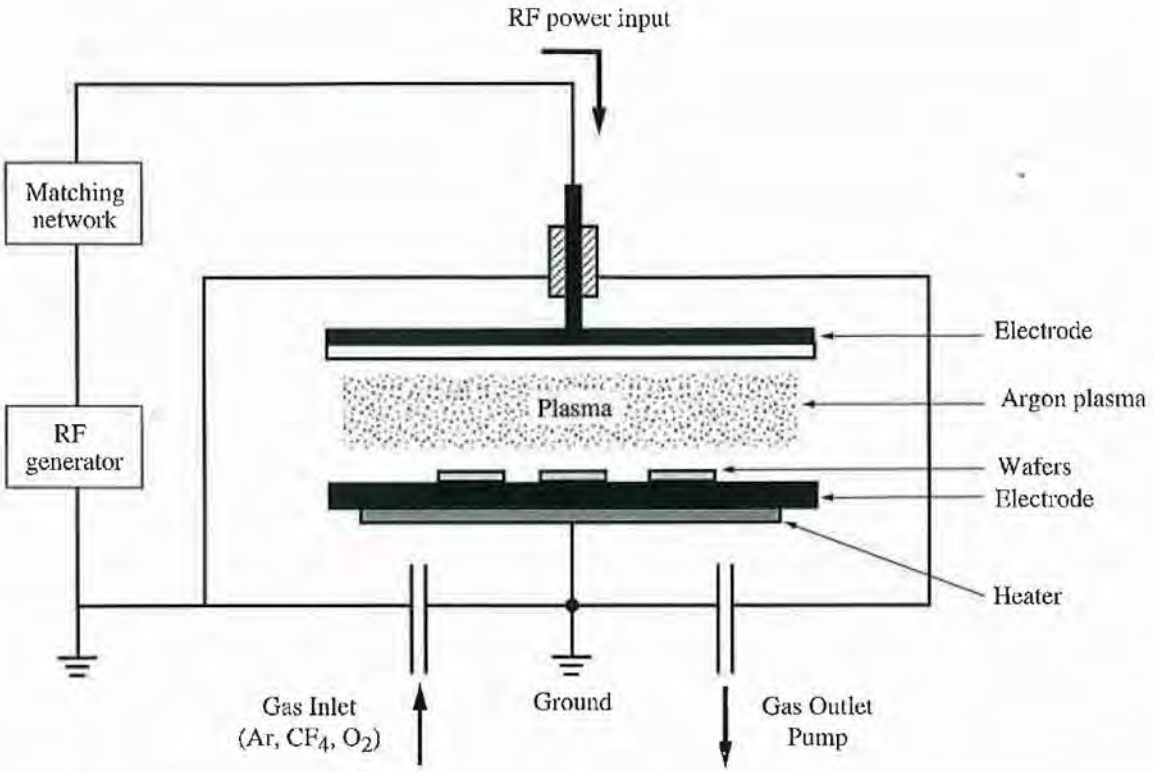

Figure 10–7 Schematic diagram of an RF-powered plasma etch system.

voltage distribution that develops is shown by the solid curve in Figure 10–8 (repeated from Figure 9–29). The sheaths are the regions next to each electrode where the voltage drops occur and correspond to the dark regions of the plasma.

The sheaths form to slow down the electron loss so that it is equal to the ion loss per RF cycle and so the average current to the electrodes is zero. Whereas the heavy ions respond to the average sheath voltage, the light electrons respond to the instantaneous voltage. However, due to the self-biasing, they only cross the sheath during a short period per cycle when the sheath voltage and thickness are near their minimum. During most of the RF cycle the electrons are turned back at the sheath edge, resulting in the sheaths, on the average, being deficient or depleted of electrons. This depletion of electrons in the sheath results in the sheath being dark because of the lack of electron/atom collisions and subsequent relaxation by light emission. In the bulk of the plasma, both ionization and excitation events occur, the latter producing the characteristic glow of the plasma. The lack of electrons in the sheath also increases the impedance of the sheath compared to the plasma bulk resulting in the main voltage drop being across the sheaths. If one of the electrodes is made much smaller in area, then the voltage distribution becomes asymmetric with a much larger voltage drop occurring from the plasma to the smaller electrode, as shown by the dashed curve in Figure 10–8. As described in Chapter 9, this happens because the RF current density must be much higher at the smaller electrode in order to maintain current continuity throughout the system. Therefore the fields must be higher at the smaller electrode.

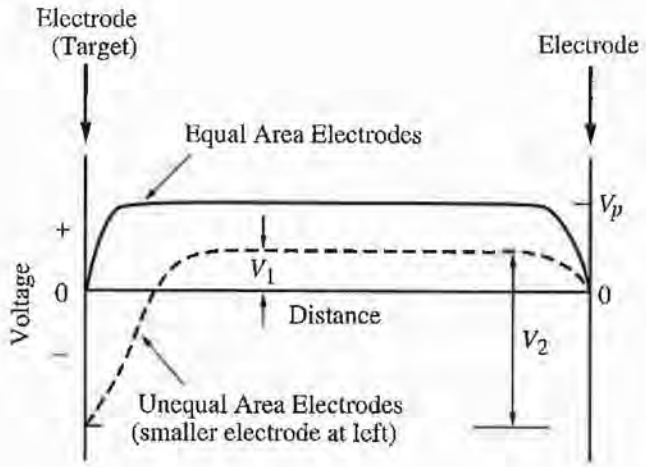

Figure 10–8 Steady-state voltage distribution in RF-powered plasma etch systems. (See Figure 9–27 and the associated text for a more complete discussion of this figure.)

For etching materials other than photoresist (for which O_2 is used) the reactant gases in the plasma are usually halide containing species, such as CF_4, Cl_2, and HBr. Sometimes small amounts of other gases, such as H_2, O_2, and Ar are also added. The high-energy electrons in the plasma can cause a variety of reactions to occur with the reactant gases, including electron-induced ionization, dissociation, recombination, and excitation reactions, as illustrated in Figure 10–9. The plasma will thus contain free electrons, ionized molecules, neutral molecules, ionized fragments of broken-up molecules, and free radicals. For example, in a CF_4 plasma, there will be electrons, CF_4, CF_3 (and other CF_X fragments), CF_3^+ and F. CF_3 and F are free radicals—very reactive species. Typically there will be about 10^{15} cm^{-3} neutral species (1 to 10% of which may be free radicals) and 10^8–10^{12} cm^{-3} ions and electrons. The actual numbers of the different species will depend on the balance between generation and loss or recombination reactions for each.

In standard plasma etch systems, the plasma density (the concentration of ions or electrons in the main plasma, which are virtually equal) and the ion energy as it travels across the sheath to strike the wafer are closely coupled. As the RF power increases, more ions are created, increasing the density, and the sheath voltage increases as well, increasing the ion energy. This coupling comes about because of the capacitive nature of the coupling between the RF current and the plasma [10.3].

10.2.2.1 Plasma Etching Mechanisms

The two main types of species involved in plasma etching are the reactive neutral chemical species and the ions. It is the reactive neutral species—free radicals in many cases, but sometimes other reactive species such as Cl_2—which are primarily responsible for the chemical component in plasma etch processes. It is the ions that are responsible for the physical component. They can work independently, or as we will see, they can work

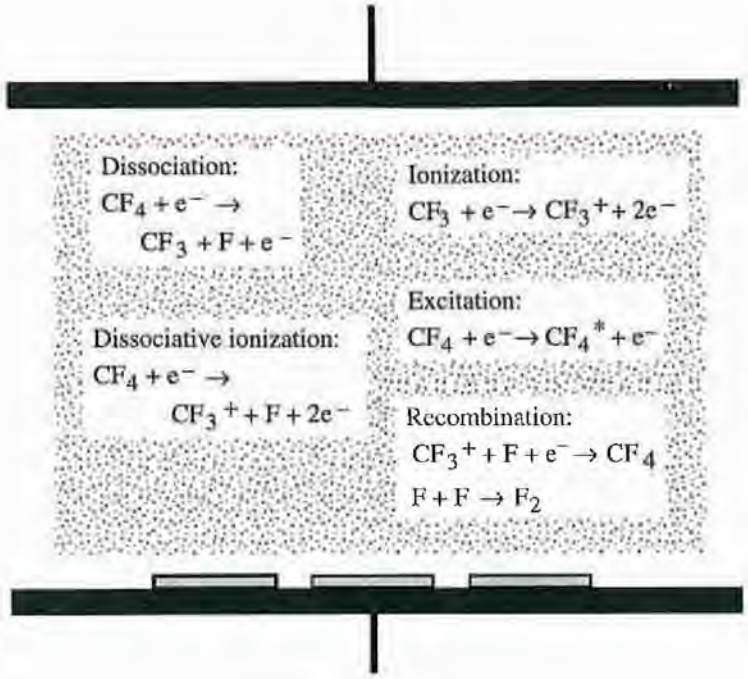

Figure 10–9 Typical reactions and species present in a plasma used for plasma etching.

together in a synergistic manner. When the reactive neutral species act by themselves, the process or mechanism is called chemical etching. Ions acting by themselves can result in physical etching, or sputtering. When the reactive neutral species and ions act in a synergistic fashion, this is called ion-enhanced etching. We will now describe each of these mechanisms.

Chemical Etching Chemical etching of materials in plasma systems is commonly done by free radicals. Free radicals are electrically neutral species that have incomplete bonding; that is, they have unpaired electrons. Examples are the fluorine free radical, F, and the neutral CF_3 free radical. Both of these can be produced by the reaction between free electrons in the plasma and CF_4, for example,

$$e^- + CF_4 \rightarrow CF_3 + F + e^- \tag{10.9}$$

Because of their incomplete bonding structure, free radicals are highly reactive chemical species. Free fluorine, with seven electrons in its outer shell instead of the energetically favorable set of eight, would much rather be bonded to other atoms so that all electrons are paired. It will therefore react quickly with other species to achieve that bonded state. The idea in plasma etching is for the reactive neutral species to react with the material to be etched, Si for example. The byproduct should be a volatile species

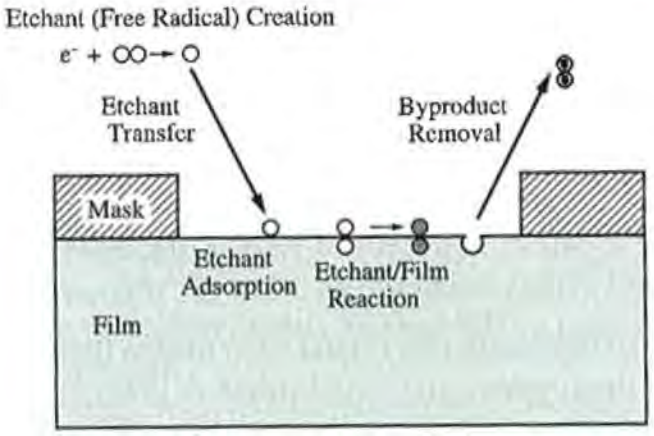

Figure 10–10 Processes involved in chemical etching during plasma etch process.

(that is, easily vaporized or evaporated into the gas phase) that leaves the surface, exposing more Si to be etched. This is illustrated in Figure 10–10. In fact, the fluorine free radical will react with Si via the reaction

$$4F + Si \rightarrow SiF_4 \tag{10.14}$$

to produce the gas SiF_4 (although there may be intermediate steps and species in that reaction). Thus a CF_4 plasma can be used to etch Si. It is important in etch processes that the byproducts of the etch, as well as any unreacted etchants, be removed from the surface for the etching to continue. Figure 10–10 shows a simple etch process, but more complex reactions may occur, involving multiple species and steps.

Gas additives can be used to increase the production of the reactive etch species and thereby increase the chemical etching rate. For example, oxygen is often added to a CF_4 plasma. The oxygen reacts with dissociated CF_4 species (CF_3 or CF_2, for example), which reduces the recombination of those species with F, thereby increasing the amount of free F. This enhances the chemical etching. However if too much O_2 is added, the etchant becomes too diluted or the surface is oxidized by the O_2, and the etching decreases.

We saw earlier that wet chemical etching occurs isotropically. The chemical component of plasma etching, when acting by itself, also occurs isotropically. Two factors lead to this behavior: an isotropic arrival angle distribution and a low sticking coefficient, with the latter usually more important. When we model and simulate plasma etching later in this chapter, we will closely follow the approach we used for deposition processes. That is, we will consider the incoming fluxes of species from the plasma to the wafer surface including the angular distribution of the various fluxes. This will be quantified by the $\cos^n\theta$ distribution, where a value of 1 for n describes an isotropic distribution (that is, coming in at all angles equally so that the normal component to the surface follows a $\cos\theta$ distribution) and a value of n larger than 1 describes a more directed distribution toward the surface. As in the deposition models we assume that all un-ionized

chemical species arrive at the wafer surface at all angles. Since the free radicals are neutral species, the flux of the chemical component of plasma etching has an isotropic arrival angular distribution, or n equals 1.

We also consider the sticking coefficient of the species, as defined by Eq. (9.25), and repeated here.

$$S_c = \frac{F_{\text{reacted}}}{F_{\text{incident}}} \tag{10.11}$$

A high sticking coefficient (equal to 1) means that the reaction, and etching, occur the first time the species strikes the surface. A lower sticking coefficient means that before the species reacts, it can leave the surface—usually assumed to occur at a random angle—and strike the surface somewhere else. One might guess that free radicals would have very high sticking coefficients, since they are so reactive. However, as in deposition processes, the chemical reactions that occur in chemical etching often require multiple steps, and multiple atoms or species, all in the right proportions. In addition, the availability of suitable sites on the surface, competing reactions, and surface recombination of species can be issues. These issues result in most free radicals having effective sticking coefficients that are low, on the order of 0.01 for CF_4/O_2 etching of Si, for example. Thus the chemical etch species bounce around a lot and tend to maintain random angles of flight even under masks and other surface features. The fluxes for chemical species are illustrated in Figure 10–11(a).

These two properties of the reactive neutral species in plasma systems—isotropic arrival angles and low sticking coefficients—result in the fact that purely chemical etching acts in an isotropic or near-isotropic manner. Only if the sticking coefficient is relatively large would shadowing by the topography and etch mask become important. (Sticking coefficients on the order of 0.1 may actually occur for some purely chemical

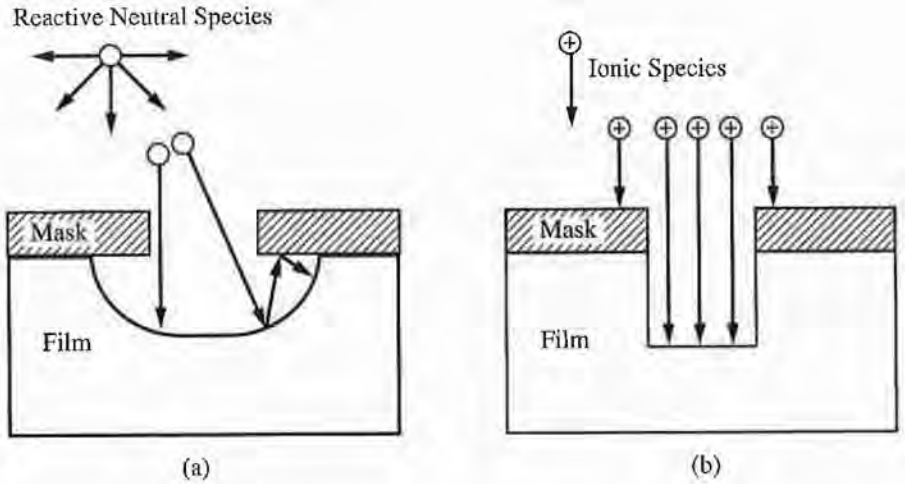

Figure 10–11 Fluxes of species in plasma etching: (a) fluxes of reactive neutral chemical species (such as free radicals), with a wide arrival angle distribution and low sticking coefficient; (b) fluxes of ionic species, with a narrow, vertical arrival angle distribution and high sticking coefficient (assumed equal to 1).

plasma etching cases [10.4].) In that event somewhat more anisotropic etch profiles can result. In general, though, pure chemical plasma etching results in isotropic or near-isotropic etching. We must reiterate that this is only for the chemical component acting independently, or for a purely chemical etch process. As we shall see shortly, the reactive chemical species and the ions can work in a synergistic manner so that the etching is not isotropic at all.

Since free radicals etch by chemically reacting with the material to be etched, the process can be very selective, as discussed earlier for wet chemical etching. Thus any chemical component to plasma etching will be relatively selective, as compared to the physical component.

In some cases, reactive neutral species other than free radicals may be involved in chemical etching. An example is in the plasma etching of Al by the Cl_2 molecule. For the same reasons given above, these reactive neutral species would have the same etching properties as the free radicals, namely isotropic profiles and selective chemical etching.

Physical Etching The other main species that participate in plasma etching are the ions that are present. Because of the voltage drop between the plasma and each electrode and the resulting electric field across each sheath, positive ions will be accelerated toward each electrode. Since the wafers are placed on one of the electrodes, then the ionic species such as Cl^+ (or Ar^+ in a purely physical sputter etch system) will be accelerated to the wafer surface. This striking of the wafer surface results in the more physical component of etching. The flux of ions toward the wafer surface is much more directional than the flux of neutral free radicals because of the directionality of the electric field from the plasma to the surface of the wafer. Thus the etching will be much more directional and more anisotropic. In fact it is usually assumed that all the ions arrive normal to the wafer surface, as shown in Figure 10–11(b), meaning that n is very large in the $\cos^n\theta$ distribution. Because the ions etch by more physical processes, rather than chemical, the etching will be much less selective. It is also assumed that once each ion strikes the surface, it does not strike the surface again somewhere else (or if it does, it does not have enough energy to do anything). This means that the sticking coefficient is effectively one. This also leads to more anisotropic behavior.

Ions can etch material in plasma systems by physical sputtering. Sputtering was discussed in Chapter 9, where we saw how sputtering of elemental, compound, or alloy targets was utilized to deposit material on the surfaces of wafers. We also saw how sputtering of the wafer itself could be done during a simultaneous deposition of material in order to planarize or fill features. Sputtering occurs by ions, accelerated by an electric field, striking the atoms on the surface of a material and physically dislodging them. The sputtering process is often described by a sputtering yield defined in Chapter 9 as the number of atoms or molecules ejected from the target per incident ion and is a measure of the sputter efficiency. As discussed in Chapter 9, the sputter yield is a function of incident angle, and it is this property that allows for planarization and filling in bias-sputtering deposition techniques. Because of the physical nature of sputtering by ion species arriving normal to the wafer surface, very anisotropic etching occurs, with profiles similar to that shown earlier in Figure 10–3(c).

When ionized reactive species such as Cl^+ and CF_3^+ are involved in sputtering, chemistry can also be important in determining the effective sputtering yield since these ions do have the potential of reacting chemically with surface atoms. Such effects can enhance the sputtering yield compared to Ar^+ sputtering, although the magnitude of these effects is usually small compared to ion-enhanced etching, which involves both neutral an ionized species and which we describe next.

Ion-Enhanced Etching Another way that ions participate in plasma etch processes is through ion-enhanced etching. It was recognized many years ago that the ions and reactive neutral species do not always act independently in an etch process. This was easily deduced when it was observed that the etch rate measured is not always just the sum of the two components acting independently; in many cases it is much higher. As a classic example of ion-enhanced etching, Figure 10–12 shows the etch rate of silicon as XeF_2 gas and Ar^+ ions are introduced to the surface. First XeF_2 gas (not a plasma, so there is no physical component at all) is introduced into the chamber and very little etching occurs. Then an Ar^+ ion beam is directed toward the surface and the etch rate increases ten times. Finally the XeF_2 gas is turned off so that only ion sputtering occurs, and the etch rate becomes almost zero. The chemical and physical species are obviously working in a synergistic manner.

This synergistic process is also demonstrated in the shape of the etch profiles that result. In ion-enhanced etching, both chemical and physical components are acting, but the profiles are not just a linear combination of isotropic chemical etching and anisotropic physical etching as illustrated in Figure 10–3(b). Instead, the profile for ion-enhanced etching is much more like the case for physical etching acting alone, as in Figure 10–3(c). If the chemical component in the etch system is increased, the vertical etching is increased but not the lateral etching, which is not what would be expected from chemical etching. In addition, the etch rate in many cases is much higher (>10 times) than that calculated from the measured ion flux assuming normal sputtering yields in the absence of neutral species; hence it is not likely that the observed plasma

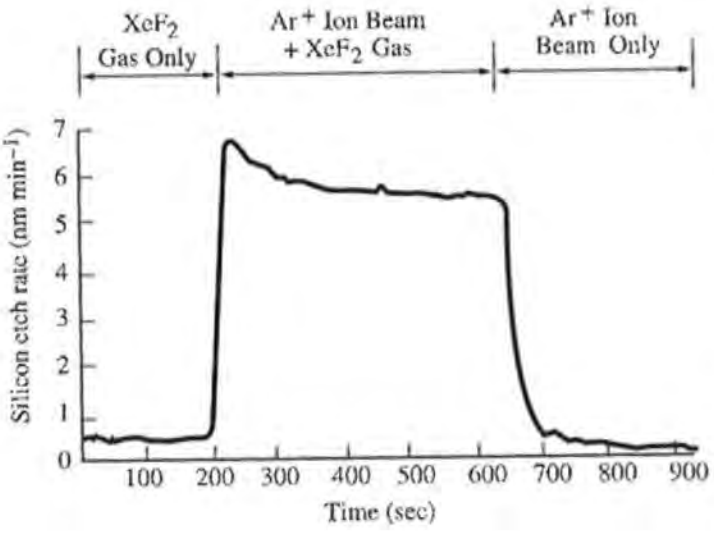

Figure 10–12 Etch rate of silicon as XeF_2 gas (not plasma) and Ar^+ ions are introduced to the silicon surface. Only when both are present does appreciable etching occur, illustrating synergistic etching behavior between chemical and ionic species. (After [10.5, 10.6].)

etch behavior in these cases is due simply to an ionized chemical species, such as CF_3^+. Observed saturation effects, which we will describe later, also support this dual component model involving both neutral and ion species.

In this type of etch process, the chemical components (the free radicals or other reactive neutral species) and the physical components (the ions) in the plasma somehow work together to etch the material, with both components required at each point for the process to work. In addition, because of the required chemical component and the chemical processes that occur, good selectivity can be obtained. So not only can one achieve very anisotropic etch profiles, but also a high degree of selectivity as well. This is something that is very difficult to achieve when the two components work independently. In addition, high etch rates can be achieved. Thus gas chemistries and etch conditions are usually chosen to promote ion-enhanced etching and to suppress independent chemical and physical etching.

There have been many explanations and mechanisms proposed for this type of cooperative etch process. In most of the models the ion bombardment enhances one of the steps of the chemical etch process, such as surface adsorption, etching reaction and formation of byproduct, or removal of byproduct or unreacted etchant. For example, the ion bombardment causes damage (breaking bonds, etc.) which makes the surface more apt to chemically react with the radical. Or the ion bombardment accelerates the formation of volatile byproducts. Or the ion bombardment may dislodge or sputter away etch byproducts which would otherwise tend to stay on the surface and impede the etch process. The removal of byproducts can also include secondary, or indirect, byproducts. It has been found that a chemically inert residue is often formed from etching or sputtering photoresist. In addition, species from the plasma may deposit on the surface. These layers that are formed or deposited, whatever their source, inhibit chemical etching either by physically blocking the chemical etch species or by reacting with them and depleting them. But the inhibitors can be removed, or prevented from forming, by ion bombardment, leading to anisotropic etching. In fact, these inhibitors are often intentionally produced on the surface by providing sources in the gas which enhance their formation so that more directional etching is achieved. In a particular ion-enhanced etch process, one or more of the above mentioned mechanisms (damage induced, inhibitor removal, etc.) may be occurring.

Whatever the exact mechanism for ion-enhanced etching, and it is often difficult to distinguish between them, directional etching occurs because of the directionality of the ions. The enhancement of the process, whether of enhancement of etch reaction or of inhibitor removal, occurs only where the ions strike the surface. Since the ions are striking normal to the wafer surface, the enhancement will occur normal to the wafer surface. This directional enhancement will result in directional, anisotropic etching. This is illustrated in Figure 10–13(a) for the case when the chemical etch reaction is enhanced by ion bombardment. It is assumed in this figure that etching by the neutral (radical) species alone without ion bombardment, the so-called "spontaneous chemical etching," is negligible, leading to no lateral etching at all. In Figure 10–13(b) the process of inhibitor formation and removal is illustrated, which also results in anisotropic etching. Ion bombardment on the horizontal surfaces removes the inhibitor there but leaves the sidewall inhibitor to protect against lateral chemical etching.

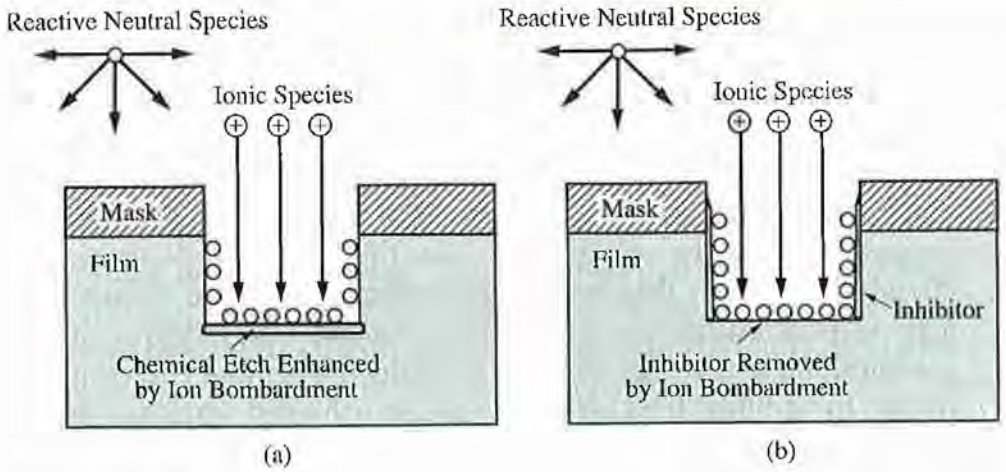

Figure 10–13 Illustration of ion-enhanced etching. In (a) the chemical etch reaction is enhanced by ion bombardment. In (b) an inhibitor is formed which is removed by ion bombardment, allowing chemical etching to proceed. In both cases, anisotropic etching results.

Even with ion-enhanced etching and no spontaneous chemical etching, the slope of the resulting sidewall may not always be exactly vertical. For example, the ion flux may not always be perfectly normal to the wafer surface and can result in slightly bowed sidewalls. Also, when inhibitors form or are deposited during etching, sloped sidewalls can result if the inhibitor formation or deposition rate is very high relative to the etch rate of the inhibitor and substrate, as demonstrated in Figure 10–14. Changing the relative deposition/etch rates by changing the etch chemistry or plasma conditions can be used, in fact, to tailor the desired slope if less than vertical sidewalls, without any undercutting, are desired in specific applications. But the important point still is that very anisotropic etching can be achieved when ion-enhanced etching is occurring.

Selectivity occurs in ion-enhanced etching because chemical reactions are involved in this type of etch process, and chemical reactions are by nature selective. To achieve the needed selectivity, one must choose the proper etchant gas (either F, Cl, or Br based, with or without oxygen or other species added, for example) and other etching conditions which favor the etching of the film over the mask material and the underlying films. With ion-enhanced etching, one can achieve good etch selectivity while still obtaining very directional etch profiles.

10.2.2.2 Types of Plasma Etch Systems

Different types of plasma etch systems and modes of operation have been developed to etch films in semiconductor structures which take advantage of the different species present in a plasma and the different etch mechanisms that can operate. We will now describe these different etch systems or modes, focusing on the advantages and disadvantages of each.

Plasma Etching in Barrel Etchers One of the earliest plasma systems developed was the barrel etcher, named because of its shape, and is shown in Figure 10–15. In this

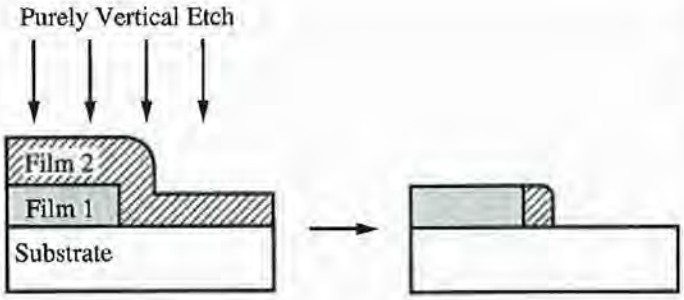

Figure 10–23 Overetching of a film over a step would be required to remove material at corner. In this example, completely anisotropic etching is assumed to emphasize this concept. This phenomenon can be used in producing self-aligned oxide spacers, in which no extra overetching is done so as to intentionally produce the structure at the right.

case of via etching, it may make it worse since sputtering of the underlying metal onto the sidewalls makes the sidewall residue harder to remove. One safe way to remove etch residues is by using a selective chemical etch after the plasma etch.

10.3.2 Plasma Etch Methods for Various Films

Most reactant gases for plasma etching contain halogens, generally Cl and F, and sometimes Br. Free radicals of these species can easily be produced in a plasma which can efficiently etch many films, and volatile etch products commonly result with these species. The exact choice of reactant gases to etch a specific film depends on a variety of factors, the most important being (1) etch selectivity to underlying films; (2) anisotropic etching; and (3) volatility of main etch byproducts.

Since in an etch process it is necessary to remove the byproducts from the surface, keep them off the wafer and pump them out of the system, the volatility of the main byproducts is an important consideration in choosing the gas chemistry. The volatility, or tendency to evaporate, depends on how tightly bound the species is to the surface. Table 10–2 lists the boiling points of typical etch products. The boiling point is an indicator of the volatility of a species—the lower the boiling point, the lower the surface binding energy and the higher the volatility. One can see that if Al is etched with a fluorine chemistry, AlF_3 would be the main byproduct. But the boiling point of AlF_3 (actually the sublimation temperature since it transforms directly from solid to gas without going through a liquid phase) is quite high at 1291°C. This means that the etch byproduct will be tightly bound to the surface with a very low vapor pressure, and even with ion bombardment it would be difficult to remove and keep off the surface. On the other hand, the byproduct of a chlorine-based etchant would be $AlCl_3$ with a boiling point (sublimation temperature again) of 177.8°C. This indicates looser bonding to the surface and a much more volatile species, so that a chlorine-based etchant would be a much better choice to etch Al. Having a volatile, weakly bound main etch byproduct is important when choosing an etch chemistry for a particular film. It is important to note that other, sometimes intermediate, byproducts may form during the etch process, and their removal or conversion to a more volatile species can be a rate limiting step in the etch process.

Table 10–3 lists some common chemistries used in plasma etching films in integrated circuit fabrication. These etchants are chosen mostly based on issues associated with byproduct volatility, etch selectivity, and etch profiles. The exact etch profiles and

Table 10–2. Boiling points of typical etch products (After [10.12]).

Element	Chlorides	Boiling Point (°C)	Fluorides	Boiling Point (°C)
Al	$AlCl_3$	177.8 (subl.)	AlF_3	1291 (subl.)
Cu	CuCl	1490	CuF	1100 (subl.)
Si	$SiCl_4$	57.6	SiF_4	−86
Ti	$TiCl_3$	136.4	TiF_4	284 (subl.)
W	WCl_6	347	WF_6	17.5
	WCl_5	276	WOF_4	187.5
	$WOCl_4$	227.5		

selectivities will depend on the specific system and plasma conditions. We will next discuss the plasma etching of three films, silicon dioxide, polysilicon, and aluminum, in a little more detail. In doing so we hope to illustrate some of the major issues involved in plasma etching and in choosing the etch chemistries.

10.3.2.1 Plasma Etching Silicon Dioxide

Fluorine-based chemistries are usually used to etch silicon dioxide. The volatility of the main etch product, SiF_4, is very high as indicated by its low boiling point in Table 10–2 and thus the SiF_4 can easily leave the surface via thermal desorption. Fluorine can etch the SiO_2 by reaction of the fluorine free radical, with the overall reaction given by

$$4F + SiO_2 \rightarrow SiF_4 + O_2 \tag{10.12}$$

CF_2 and CF_3 free radicals may also be directly involved in the case of fluorocarbon etchants, in which case CO-containing byproducts would also be produced.

Etching with a significant isotropic component can result using source gases such as NF_3, SF_6, CF_4, or CF_4/O_2 mixtures. CF_4 dissociates to form free F in a plasma via the reaction given in Eq. (10.9). As mentioned earlier, adding O_2 to the CF_4 gas significantly increases the F free radical production and hence increases the etch rate as well as makes the etch profile more isotropic. Too much oxygen, though, can decrease the etch rate due to dilution [or by reversing the reaction in Eq. (10.12) and redepositing oxide on the surface]. These gas sources with relatively high free F concentrations will etch the oxide; however, depending on the conditions, the etching can result in undercutting. In addition, the etch selectivity over silicon or polysilicon, which is often underneath the oxide, is poor. The anisotropy can be enhanced by increasing the ion bombardment and lowering the pressure [10.14], but the selectivity over silicon is still generally poor with these chemistries.

The key to making the etching of silicon dioxide very anisotropic as well as much more selective over Si is to reduce the amount of F free radical production and increase the carbon content. Doing so will decrease any spontaneous isotropic etching by free

Table 10–3. Typical or representative etch plasma etch gases for films used in IC fabrication (After [10.1, 10.4, 10.13, 10.14].)

Material	Etchant	Comments
Polysilicon	SF_6, CF_4	Isotropic or near isotropic (significant undercutting); poor or no selectivity over SiO_2.
	CF_4/H_2, CHF_3	Very anisotropic; nonselective over SiO_2.
	CF_4/O_2	Isotropic; more selective over SiO_2.
	HBr, Cl_2, Cl_2/HBr/O_2	Very anisotropic; most selective over SiO_2.
Single-crystal Si	same etchants as polysilicon	
SiO_2	SF_6, NF_3, CF_4/O_2, CF_4	Can be near isotropic (significant undercutting); anisotropy can be improved with higher ion energy and lower pressure; poor or no selectivity over Si.
	CF_4/H_2, CHF_3/O_2, C_2F_6, C_3F_8	Very anisotropic; selective over Si.
	CHF_3/C_4F_8/CO	Anisotropic; selective over Si_3N_4.
Si_3N_4	CF_4/O_2	Isotropic; selective over SiO_2 but not over Si.
	CF_4/H_2	Very anisotropic; selective over Si but not over SiO_2.
	CHF_3/O_2, CH_2F_2	Very anisotropic; selective over Si and SiO_2.
Al	Cl_2	Near isotropic (significant undercutting).
	$Cl_2/CHCl_3$, Cl_2/N_2	Very anisotropic; BCl_3 often added to scavenge oxygen.
W	CF_4, SF_6	High etch rate; nonselective over SiO_2.
	Cl_2	Selective over SiO_2.
Ti	Cl_2, $Cl_2/CHCl_3$, CF_4	
TiN	Cl_2, $Cl_2/CHCl_3$, CF_4	
$TiSi_2$	Cl_2, $Cl_2/CHCl_3$, CF_4/O_2	
Photoresist	O_2	Very selective over other films

fluorine. The etching that then occurs involves other fluorine species and is strongly ion enhanced. In addition, polymer inhibitor formation will be increased, also suppressing lateral etching. Furthermore, selectivity over Si will be greatly enhanced.

One way to reduce the relative amount of F free radicals is by adding H_2 to the CF_4 gas. The hydrogen reacts with the fluorine to produce HF which reduces the amount of atomic F. Another possibility is to use more fluorine-poor and more carbon-rich gas species, such as CHF_3, C_3F_8, or C_2F_6. With fewer F free radicals, any spontaneous (non-ion-enhanced) chemical etching by these species is reduced. Appreciable etching happens only where ion bombardment occurs, perhaps involving an ion-induced damage or bond-breaking process, leading to anisotropic etching. CF_2 or CF_3 free radicals, rather than free F, are believed to be involved in the etch mechanism in these fluorine-deficient gas mixtures [10.14]. Increasing the carbon-to-fluorine ratio with these etchants also increases the amount of polymer inhibitor formed, or deposited, during the etching [10.6]. With a CF_4 or CF_4/O_2 gas, little or no polymer forms on the surface. However with a higher carbon concentration relative to fluorine in the gas mixture,

more polymer formation will occur relative to etching. With a thick polymer inhibitor layer present, etching can not occur on vertical surfaces where there is little ion bombardment to remove the polymer inhibitor layer. Any lateral etching is suppressed, and near vertical sidewalls can result.

Of course, as even more carbon is added to the system and the deposition of the polymer becomes faster relative to the etching of the polymer and film, then sloped sidewalls (without mask undercutting) can result, as discussed earlier and illustrated in Figures 10–14 and 10–21. CHF_3 has a higher ratio of C to F than CF_4 does so that more polymer forms, in fact enough so that sloped sidewalls often result. If pure CHF_3 produces excess inhibitor deposition for particular plasma conditions, then O_2 could be added to react with the carbon and reduce the polymer deposition and change the etch profile. Other species such as CO_2 or CF_4 may also be added to CHF_3 for this purpose. Figure 10–24 conceptually illustrates the different etch profiles in SiO_2 that could result when different amounts of inhibitor formation occur.

Reducing the fluorine-to-carbon ratio also greatly improves the etch selectivity of SiO_2 over Si. This is because polymer formation preferentially occurs on Si surfaces compared to SiO_2 surfaces during the etch process. This may be due to the fact that carbon in the polymer reacts with the oxygen from the oxide, forming carbon monoxide, which results in much less accumulation of polymer on the oxide surface. With more polymer on the Si, its etching is slower, resulting in a higher SiO_2 to Si etch selectivity. Figure 10–25 shows how the etch rates of Si and SiO_2, and hence the selectivity, vary as H_2 is added to a CF_4 gas mixture [10.15]. The etch selectivities of oxide over $TiSi_2$ and TiN, which are important when etching contact or via holes in multilayer structures, are enhanced in this same manner by adding H_2 to the fluorocarbon etchant.

Changing the etchant chemistry may also change the etch rate of the mask. As seen in Figure 10–25, the etch rate of photoresist changes significantly as the H_2 content in the H_2/CH_4 gas mixture changes, with a lower H_2 content (and less F production) resulting in a higher etch rate. Adding O_2 to a CHF_3 mixture similarly increases the etch rate of the photoresist. The resulting mask erosion can cause the etch profile of the SiO_2 to change, as we will illustrate in the models and simulation section (Figure 10–33), and may alter the profiles shown above in Figure 10–24.

A relatively large bias voltage is often used for SiO_2 etching (in the 400- to 500-V range). This is needed to help remove the polymers. After the SiO_2 etching is complete, including a suitable overetch, any remaining polymer residue must be removed, espe-

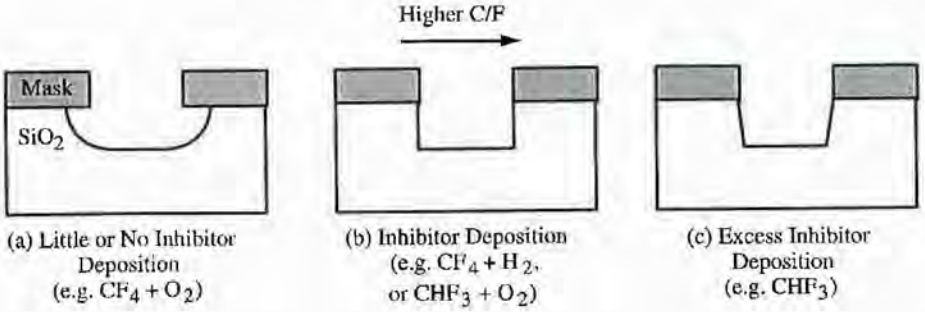

Figure 10–24 Effect of inhibitor deposition on resulting etch profile. The profiles shown assume no mask erosion, which may occur for different etch chemistries. The actual profiles obtained will also depend on the specific etch system and plasma conditions.

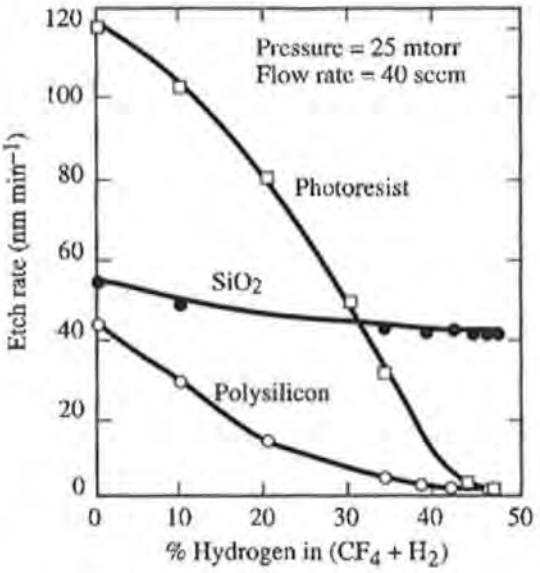

Figure 10–25 The etch rates of Si and SiO_2 (and photoresist) as H_2 is added to a CH_4 gas mixture. (After [10.15].)

cially in contact holes. The polymer can be removed by an O_2 or CF_4 plasma treatment. CF_4 is better at removing all the polymer and even some damaged substrate, but it cannot be over done since it has poor selectivity over Si and there can be lateral etching. A lower energy is used for the cleanup step so as not to damage the exposed material below. Also helpful is an anneal treatment; the polymer appears to break down at temperatures higher than $\approx 280°C$.

It should be clear from the discussion in this section that there are many details and subtleties associated with developing a particular manufacturing recipe for plasma etching. We will return to some of these issues in the models and simulation section.

10.3.2.2 Plasma Etching Polysilicon

The choice of etch chemistry for polysilicon depends on the same factors as etching silicon dioxide: byproduct volatility, selectivity, and anisotropy. Another issue is the presence of the ubiquitous native oxide layer on the film's surface, requiring an initial "breakthrough" etch step to remove it.

F-based etchants could be used for polysilicon etching because of the volatility of SiF_4. However, etching polysilicon with SF_6, CF_4, or CF_4/O_2 has a relatively large isotropic component due to the high concentration of F free radical species produced in such plasmas combined with little polymer formation. We saw that with etching SiO_2, anisotropy can be enhanced by reducing the relative F component and increasing the carbon component, thus decreasing the spontaneous etching, increasing polymer formation, and enhancing ion-induced etching. Adding H_2 or using CHF_3 can accomplish this. But doing so also increases the etch rate of SiO_2 relative to Si by preferentially forming a polymer on the Si. This is good when etching SiO_2 but bad when etching polysilicon when a high selectivity over the underlying SiO_2 is required. (Recall the structure in Figure 2–26 in the CMOS process flow in Chapter 2 for an example of why good selectivity over SiO_2 might be important.) Thus with a fluorine chemistry it is difficult