Chapter 2
Optical Pattern Formation

2.1 The Problem of Imaging

The basic problem of imaging is shown in Fig. 2.1. Light from an illumination source passes through a photomask, which defines the patterns. The simple photomask illustrated here consists of two types of complementary areas—one type that is opaque, while the other is transparent. In this example, the transparent (or "clear") area is a long space of uniform width, and the optical and resist profiles shown in Fig. 2.1 are cross sections for this geometry. Some of the light that passes through the mask continues through a lens, which projects an image of the mask pattern onto a wafer. The wafer is coated with a photosensitive film, a photoresist that undergoes a chemical reaction upon exposure to light. After exposure, the wafer is baked and developed, leaving regions covered by photoresist and complementary regions that are not covered. The patterning objective of microlithography is to produce well-defined resist features, sized within specifications. This is a challenge because of the shape of the light-intensity distribution produced at the wafer plane by a lens of finite resolution. This distribution lacks a clearly defined edge (Fig. 2.1). From the light-intensity distribution alone, it is not possible to know where the edges of the feature are. If the light-intensity distribution had the shape shown in Fig. 2.2, there would be no such a problem, because a clear delineation would exist between areas of the resist exposed to light and areas unexposed to light.

The light distribution shown in Fig. 2.1 was not drawn arbitrarily or artistically. Because light is a form of electromagnetic radiation, it is possible to use equations that describe the propagation of electromagnetic waves to calculate optical image formation,[1] and the light-intensity distribution shown in Fig. 2.1 was generated accordingly. The physics of image formation will be discussed in more detail in this and later chapters.

Lithographic processes are easier to control the closer the actual optical images resemble the ideal ones. If the light profile at the wafer plane is represented by the distribution shown in Fig. 2.2, the edges of the feature could be clearly identified by looking at the light-intensity distribution. The photoresist on the wafer would be cleanly separated into exposed and unexposed areas. In the situations that actually occur, the photoresist receives continuously varying doses of light in the regions corresponding to features on the mask. Proper pattern definition on the wafer

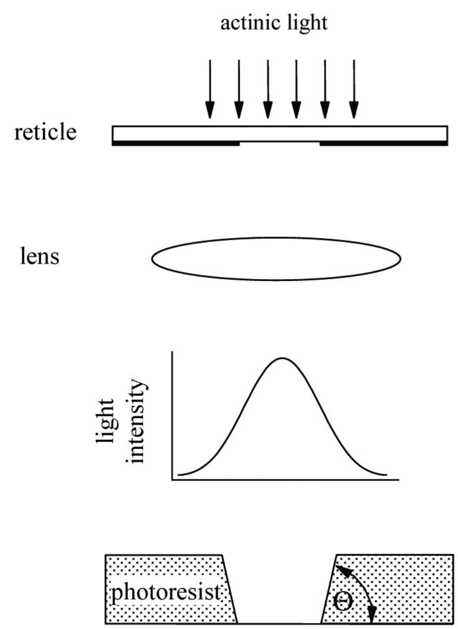

Figure 2.1 An illustration of the imaging process. Light passes through a reticle (photomask). The resulting pattern is imaged onto a photoresist-covered wafer by a lens. The finite resolution of the lens results in a light-intensity distribution that does not have clearly defined edges. The particular light-intensity distribution shown in this figure was calculated using PROLITH 1.5, for a $0.25\text{-}\mu m$ space on a $0.65\text{-}\mu m$ pitch (i.e., for a grating pattern of $0.25\text{-}\mu m$ spaces and $0.4\text{-}\mu m$ lines), where the numerical aperture (NA) of the aberration-free lens is 0.5, imaging at a wavelength of 248 nm, and with 0.5 partial coherence. The parameter "numerical aperture" will be explained later in this chapter, and partial coherence is discussed in Appendix A. *Actinic* refers to light that drives chemical reactions in the photoresist.

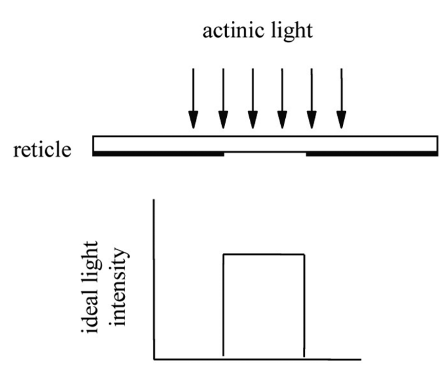

Figure 2.2 An ideal light-intensity distribution.

requires that small differences in exposure doses at the edges of pattern features be distinguished through the resist processing. Edge control will be better for light-intensity distributions that closely resemble the ideal case illustrated in Fig. 2.2.

Light-intensity profiles produced by two sets of optics, one of "higher resolution" than the other, are shown in Fig. 2.3. (The parameters that affect image sharpness and resolution will be discussed later in this chapter.) The image produced by the higher-resolution lens is closer to the ideal image than the one produced by the lens with lower resolution. For a given feature size, it is possible to have better pattern definition with higher-resolution optics. However, because of the highly competitive nature of the microelectronics industry, lithographers need to operate their processes at the limits of the best available optics. With a given set of optics, the light-intensity distributions are degraded further, relative to the ideal case, with smaller features (Fig. 2.4). As features become smaller, the edge slope of the light-intensity profiles become more sloped and the peak intensity decreases. The challenge for lithographers is to produce the smallest possible features on the wafers, with shape and size well controlled, for a given generation of optics.

2.2 Aerial Images

The origin of the image-intensity profile of Fig. 2.1 is the physical phenomenon of diffraction, a subject that was studied extensively in the nineteenth century and is well understood.[1] Diffraction effects are inherent to the wave nature of light. The resolution of optical tools is fundamentally limited by the physical phenomenon of diffraction, and as device geometries shrink, the lithography engineer must ultimately contend with this barrier imposed by the laws of physics. This is illustrated in Fig. 2.4, where aerial images are shown for features of varying sizes.

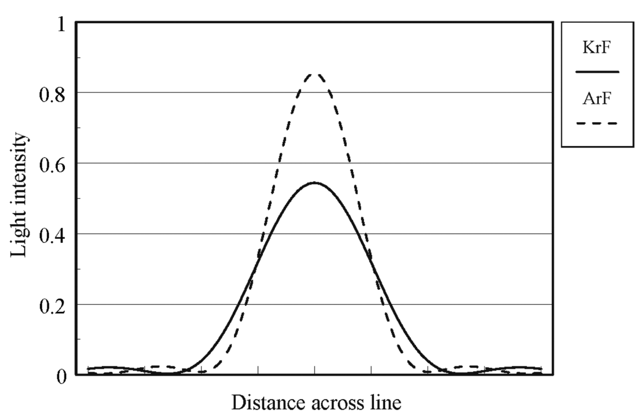

Figure 2.3 Light-intensity profiles for 150-nm isolated spaces (nominal). One image was generated for an aberration-free 0.7-NA lens, with a partial coherence of 0.6, at a wavelength of 248 nm (KrF), while the other image was generated for a 193-nm (ArF) unaberrated lens, with a numerical aperture of 0.75 and a partial coherence of 0.6. These images were calculated using the simulation program Solid C.[2] The ArF lens has higher resolution than the KrF lens. All images are in the plane of best focus.

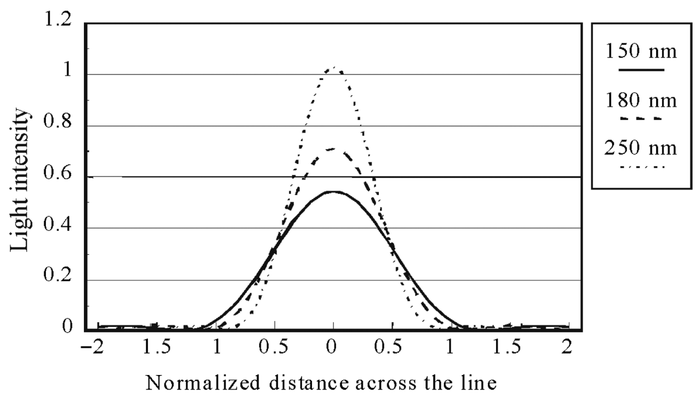

Figure 2.4 Calculated light-intensity distributions for isolated spaces of varying sizes, using parameters for an aberration-free 0.7-NA lens, with a partial coherence of 0.6, at a wavelength of 248 nm. All of the features are normalized to a width of 1.0. All images are in the plane of best focus.

As the feature size shrinks, the edge acuity of the light-intensity distribution is degraded. If even smaller features are considered, at some point one must say that features are no longer resolved, but it is clear from Fig. 2.4 that there is a gradual transition from "resolved" to "unresolved." A definition of resolution is not obvious because of this lack of a clear delineation between "resolved" and "unresolved." Simple diffraction analyses lead to the most frequently cited quantitative definition of resolution, the Rayleigh criterion, which will be introduced shortly. While sufficiently instructive to justify discussion, the Rayleigh criterion does not provide a measure directly applicable to the situation encountered in photolithography, and one should use it with care. Discussion of the Rayleigh criterion in this book emphasizes its assumptions, and therefore its applicability.

The subjects of diffraction and image formation are quite involved and, at best, can only be introduced here. The mathematical formalism is particularly complex, and the interested reader is referred to the references cited at the end of this chapter. The phenomenon of diffraction will be introduced through two examples—the diffraction grating and the circular aperture.

The origin of the resolution limits of optical systems can be appreciated by considering the system shown in Fig. 2.5. A reticle with a diffraction grating with a periodicity $2d$ (equal lines and spaces) is illuminated by a source of coherent light of wavelength λ. (For readers unfamiliar with coherence, a short introduction is provided in Appendix A.) For an angle of incidence of θ_i, the grating diffracts the light into various beams whose directions are given by[1]

$$\sin(\theta) - \sin(\theta_i) = \frac{m\lambda}{2d}, \quad m = 0, \pm 1, \pm 2, \ldots. \tag{2.1}$$

Here we are assuming that the index of refraction of air, which surrounds the mask and lens, is ≈ 1.

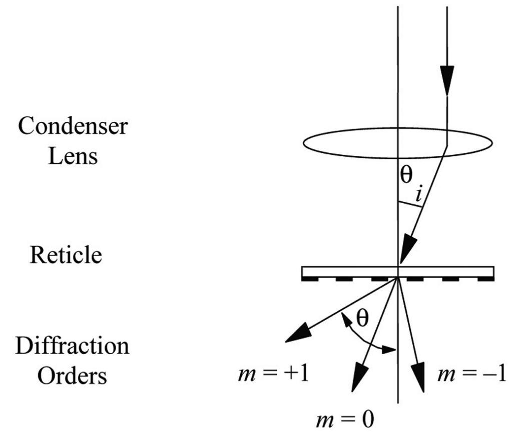

Figure 2.5 A situation with coherent light and a diffraction grating on the reticle. A ray of light is diffracted into beams that propagate in distinct directions given by Eq. (2.1).

Consider a lens with a collection angle of θ_0 (Fig. 2.6) used to image the grating. Due to Eq. (2.1), the lens collects only a finite number of diffraction beams. Coherent light illuminating a grating diffracts into beams that correspond to the object's Fourier components, and a diffraction-limited lens will recombine those beams that pass through the entrance pupil of the lens. If the lens collects all of the diffracted beams, then the grating will be fully reconstructed in the image. However, since the lens collects only a finite number of beams the image is only a partial reconstruction of the original grating pattern. As light from increasingly larger angles is collected, the image consists of more terms in the Fourier-series expansion of the light-intensity distribution of the grating (see Fig. 2.7). For on-

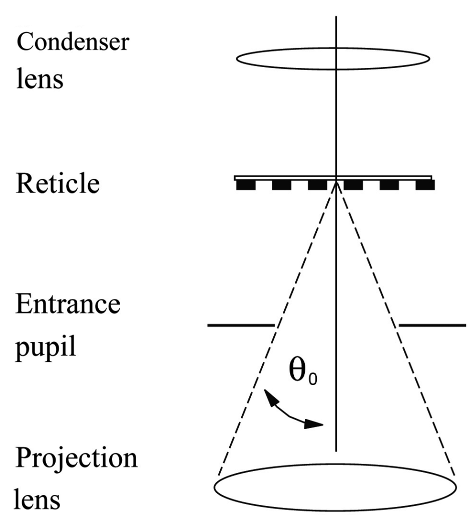

Figure 2.6 The finite size of lenses limits the angles at which diffracted beams are collected.

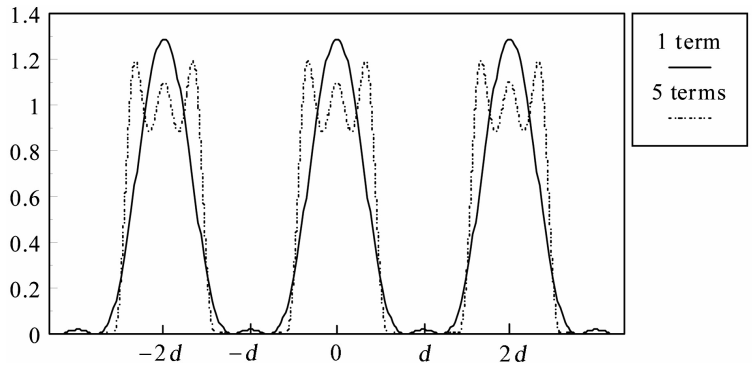

Figure 2.7 Partial sums of Eq. (2.2), with $I_0 = 1$. The image more closely approximates a rectangular grating when more terms are added.

axis illumination ($\theta_i = 0$),

$$I(x) = I_0 \left| \frac{1}{2} - \frac{2}{\pi} \sum_{m=1}^{\infty} \frac{(-1)^m}{m} \sin \frac{m\pi}{2} \cos \frac{m\pi x}{d} \right|^2 . \tag{2.2}$$

If the lens captures only a single beam, then there is no pattern in the image, since a single plane wave has no spatial variation in its intensity:

$$\text{Intensity of a single plane wave} = \left| A_0 e^{i\vec{k}\cdot\vec{x}} \right|^2 = |A_0|^2 . \tag{2.3}$$

At least two interfering plane waves are needed to generate spatial variation. To retain more than the first (constant) term in the expansion of $I(x)$, θ_0 must be large enough to capture diffracted beams for $m \geq 1$. Equation (2.1) then leads to the following expression for a minimally resolved feature:

$$d = 0.5 \frac{\lambda}{\sin \theta_0} . \tag{2.4}$$

The resolution is found to be directly proportional to the wavelength of the light and inversely proportional to $\sin \theta_0$. Resolution is limited because only a finite number of diffracted beams pass through the entrance pupil of the lens.

The phenomenon of diffraction can be appreciated further by considering the situation in which uniform illumination is normally incident on a masking screen where there is a circular hole (Fig. 2.8). Such a situation arises, for example, if the light is produced by a point source located far from the masking screen. The light that passes through the hole illuminates another screen. In the absence of diffraction, the incident light would simply pass through the aperture and produce a circular illuminated spot on the second screen. Because of diffraction, in addition to the normally incident rays of light, some light propagates at divergent angles.