

Impact of lens aberrations on optical lithography

by T. A. Brunner

All optical projection systems for microlithography depart from perfection because of various lens aberrations, especially when large image field size is combined with high numerical aperture (NA). Such aberrations have a variety of effects on lithographic imaging: shifts in the image position, image asymmetry, reduction of the process window, and the appearance of undesirable imaging artifacts. These undesirable effects are sometimes exacerbated through use of resolution enhancement techniques such as phase-shift masks or nonstandard illumination. This paper examines the impact of different types of aberrations on lithographic imagery through simulation. New techniques for measuring aberrations by examining lithographically printed resist patterns are considered.

Introduction

Modern lithography tools used in integrated circuit (IC) manufacturing lines are capable of imaging a complex chip pattern with billions of pixels, in an exposure lasting a fraction of a second. Progress in lithographic projection optics has been steady over the past two decades, made possible by advances in optical design, optical materials, mounting techniques, interferometric metrology, antireflective coatings, and precision engineering. Fueled

by the strong economics of shrinking IC features, optical projection systems of prodigious capability are now commercially available, e.g., a 0.63-NA I-line projector [1] with a resolution of 350 nm (less than the imaging wavelength of 365 nm!) and image distortion less than 30 nm over a 22-mm-square field. The goal in building such a projector is "diffraction-limited imagery," that is, optical performance which is not limited by lens imperfections. In real optical systems, this goal is never fully attained because of lens aberrations, both in the optics design and, more significantly, in the manufacturing of the optics.

Guidelines for resolution in a diffraction-limited projector are found in the Rayleigh scaling equation

$$W_{\min} = k_1 \times \lambda / NA, \quad (1)$$

where $W_{\min}$ is the minimum resolved linewidth, λ is the imaging wavelength, NA is the numerical aperture of the projection optics, and k_1 is a dimensionless number of order unity. In the past, an IC manufacturing line would be expected to print minimum features with $k_1 = 0.8$. More recently, k_1 has been driven to values near 0.6, and innovative techniques [2] such as phase-shift mask, off-axis illumination, and optical proximity correction potentially may allow $k_1 < 0.5$. These trends require that aberrations must be reduced below levels tolerable for $k_1 = 0.8$ lithography.

This paper examines optical aberrations in the context of IC lithography. The first section describes how to simulate aberrated images. The next section examines the

©Copyright 1997 by International Business Machines Corporation. Copying in printed form for private use is permitted without payment of royalty provided that (1) each reproduction is done without alteration and (2) the *Journal* reference and IBM copyright notice are included on the first page. The title and abstract, but no other portions, of this paper may be copied or distributed royalty free without further permission by computer-based and other information-service systems. Permission to republish any other portion of this paper must be obtained from the Editor.

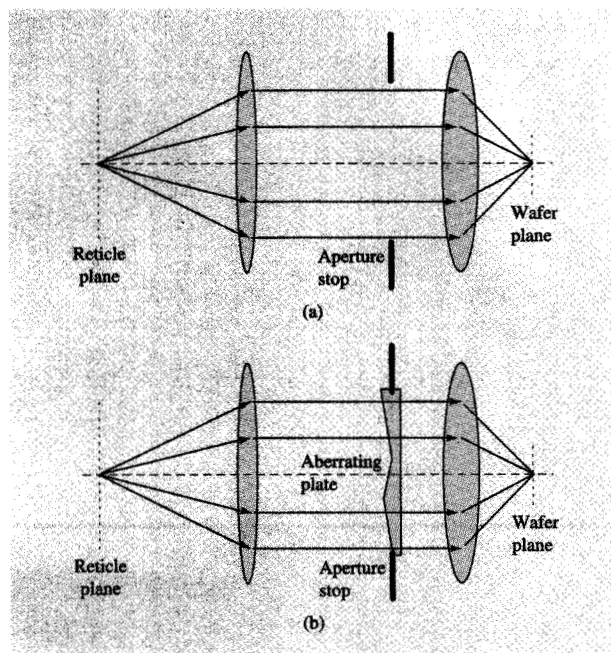


Figure 1

Schematic diagrams of a simple 2× reduction projector: (a) Aberration-free projector, such that $OPD = 0$ for every path from object to image. (b) Aberrated projector, where a transmitting plate causes each ray from object to image to have unique OPD.

ways in which different types of aberrations affect lithographic pattern quality. Finally, some techniques for measuring aberrations by examining lithographically printed resist patterns are considered.

Simulation of aberrated images

Let us consider a simple aberration-free 2× projector, as schematically shown in **Figure 1(a)**. Several light rays from the object point on the reticle to the image point at the wafer plane are shown. For each such ray, the optical path can be defined as the distance along the ray times the local index of refraction. For an aberration-free projector, all possible rays from object point to image point have exactly the same optical path. By definition, lens aberrations occur when different rays have different optical paths. Any desired aberration can be introduced by adding a suitably shaped transmitting plate in the aperture stop, as shown in **Figure 1(b)**.

The optical path difference (OPD) of a particular ray is defined to be the difference between the optical path of that ray and the optical path of the reference ray which passes through the center of the aperture stop. Since every different optical ray passes through a different part of the aperture, the OPD may be defined as a surface across the

aperture whose shape is much like the aberrating plate. The location of the ray within the aperture is specified by cylindrical coordinates ρ, ϕ , where ρ is 0 at the center and 1 at the extreme edge of the aperture. For more complicated systems than the simple example of **Figure 1**, the OPD surface can usually be defined across the exit aperture of the optics. The shape of the $OPD(\rho, \phi)$ surface fully specifies the aberrations at a particular point in the image field, and is normally represented as a sum of Zernike polynomials¹,

$$OPD(\rho, \phi) = \sum a_j \lambda Z_j(\rho, \phi). \quad (2)$$

This paper uses orthonormal Zernike polynomials Z_j , as defined using Mahajan's convention [3]. The a_j coefficient determines the contribution of the j th Zernike term measured in waves, i.e., in units of wavelength λ . In this representation, the a_j coefficient represents the root-mean-square deviation of the OPD surface contributed by the j th Zernike term. Note that this approach represents the aberrations within some small portion of the projected image field. To fully characterize a lens system, one must independently measure the Zernike coefficients at many points across the image field.

Simulations of aberrated images were calculated by the VCIMAGE program, an internal IBM program. VCIMAGE includes full vector diffraction on the wafer side of the optics, following the thesis work of Flagello [4]. Partial coherence is treated by breaking up the illumination into a variable number of discrete sources. The program can model images projected by ordinary binary masks as well as phase-shifted masks. Lens aberrations may be specified by a set of up to 37 Zernike polynomial coefficients, $\{a_1, a_2, \dots, a_{37}\}$.

A simple test based on the Strehl ratio can be used to verify the accuracy of simulation of aberrated images. A subresolution, isolated contact hole is used as the mask structure, and the image intensity is calculated at the center of the contact hole, both with and without aberration. The Strehl ratio S is defined as the ratio of the aberrated center intensity to the unaberrated center intensity. Previous theoretical work [5] has shown that for small aberrations, the Strehl ratio depends only on the total sum of the squares of the Zernike coefficients,

$$S \approx \exp \left[-4\pi^2 \sum_{j=2} a_j^2 \right], \quad (3)$$

excluding the $j = 1$ term, which simply represents an unimportant constant phase shift. In particular, an aberrated optical system with 0.1 waves of any single Zernike polynomial, e.g., $\{0, a_2 = 0.1, 0, 0, 0, 0, \dots\}$ or

¹ The interested reader is referred to the extensive literature on lens aberration theory, e.g., V. N. Mahajan, *Aberration Theory Made Simple*, SPIE Press, Bellingham, WA, 1991.

$\{0, 0, \dots, 0, a_{18} = 0.1, 0, \dots\}$, will have almost the same Strehl ratio.

Strehl ratios were calculated using 0.1 waves of each individual Zernike term from $j = 2$ to $j = 37$, using three image simulation programs: VCIMAGE, SPLAT v5.0 [6], and FAIM v2.3 [7]. Figure 2 shows that VCIMAGE and SPLAT are in excellent agreement with each other, as well as the theoretical expectation of $S \approx 0.67$, from Equation (3). FAIM showed incorrect results for the $j = 4, 11, 22$, and 37 aberrations. This problem was traced to a program error in the normalization of the spherically symmetric Zernike functions, and was corrected in a subsequent release of FAIM. The Strehl test can easily be applied to any simulation program which calculates aberrated images.

Lithographic impact of aberrations

Lens aberrations have a variety of effects on lithographic imaging [8, 9], such as shifts in the image position, image asymmetry, reduction of the process window, and the appearance of undesirable imaging artifacts. In this paper, X and Y axes are in the image plane, i.e., the wafer plane, while ΔZ refers to the defocus direction perpendicular to the image plane.

The present work concentrates on the first 11 Zernike polynomials, listed in Table 1. Z_1 , a constant phase across the aperture, does not affect imagery and is not considered further.

• Image shifts— Z_2, Z_3

These two aberrations represent a simple tilt of the OPD surface, and the imaging consequence is a positional shift of the image in the plane of the wafer. The shift can be represented as a vector $(\Delta X, \Delta Y)$ which is proportional to the Zernike coefficients as

$$(\Delta X, \Delta Y) = (a_2, a_3) \times 2\lambda/NA. \quad (4)$$

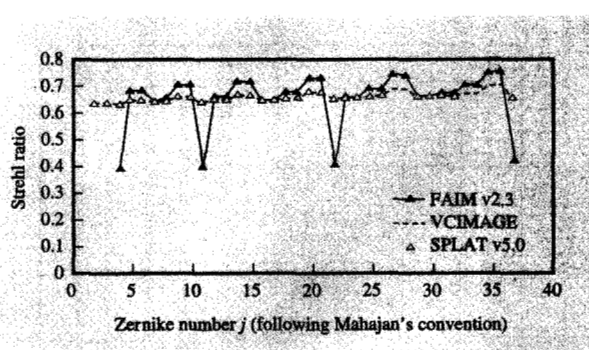


Figure 2

Strehl ratios calculated by three different simulation programs with 0.1 waves of each individual term of the first 37 Zernike aberrations. An approximate value of 0.67 is expected for each individual aberration. FAIM v2.3 miscalculated the spherical aberrations $Z_4, Z_{11}, Z_{22}, Z_{37}, \dots$, resulting in an incorrect Strehl ratio. The Strehl ratio was calculated using an isolated sub-resolution ($k_1 = 0.1$) contact hole as the mask structure.

Thus, an aberration of $a_2 = 0.05$ waves causes the image to shift in the X direction by $0.1\lambda/NA$ relative to that of an unaberrated lens. The amount of shift is completely independent of the complexity or feature size of the pattern. If a_2 and a_3 were constant across the entire lens field, a simple realignment of the wafer would correct the positional error. However, in most lithographic optics there is significant variation of a_2 and a_3 across the lens field, resulting in lens distortion which ultimately causes overlay errors. In an IC manufacturing line with many lithographic tools, lens distortion variations among the tools are one of the most important limits to overlay performance.

Table 1 The first eleven Zernike polynomials.

Z_j	Name	Equation for Z_j	Imaging consequence
Z_1	Piston	1	None
Z_2	Image translation X	$2\rho \cos(\phi)$	Shift of image, independent of pattern
Z_3	Image translation Y	$2\rho \sin(\phi)$	
Z_4	Defocus	$\sqrt{3}(2\rho^2 - 1)$	Image degradation
Z_5	Astigmatism $\pm 45^\circ$	$\sqrt{6}\rho^2 \sin(2\phi)$	Orientation-dependent shift of focus
Z_6	Astig. Hor./Vert.	$\sqrt{6}\rho^2 \cos(2\phi)$	
Z_7	Coma Y	$\sqrt{8}(3\rho^3 - 2\rho) \sin(\phi)$	Image asymmetry and pattern-dependent shift of image
Z_8	Coma X	$\sqrt{8}(3\rho^3 - 2\rho) \cos(\phi)$	
Z_9	Three-leaf clover	$\sqrt{8}\rho^3 \sin(3\phi)$	Imaging anomalies with threefold symmetry
Z_{10}	Three-leaf rotated 30°	$\sqrt{8}\rho^3 \cos(3\phi)$	
Z_{11}	Third-order spherical	$\sqrt{5}(6\rho^4 - 6\rho^2 + 1)$	Pattern-dependent focus shifts

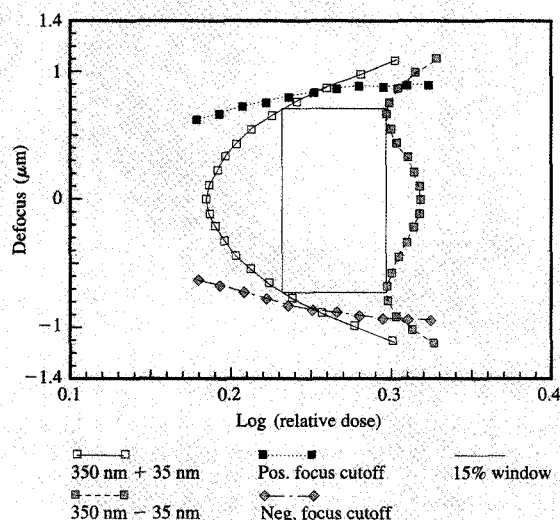


Figure 3

The process window for printing 350-nm line/space grating lines using a 0.5-NA projector with $\lambda = 248$ nm and $\sigma = 0.6$. The irregular area within the four curves represents the exposure-defocus space where the resist critical dimension (CD) prints within $\pm 10\%$ of the target 350 nm. Various rectangular process windows can be defined within this space, such as the one shown which has an exposure latitude of 15% and a 1450-nm depth of focus (DOF).

• Defocus— Z_4

The Z_4 aberration leads to a quadratic OPD surface which tends to degrade the image contrast, edge slope, pattern fidelity, and resolution relative to an aberration-free image. Under the paraxial assumption which is valid when $NA \ll 1$, a_4 is directly proportional to the defocus ΔZ through the relationship

$$\frac{\Delta Z}{\lambda/(2NA^2)} = 8\sqrt{3} \cdot a_4 \approx \frac{a_4}{0.072 \text{ waves}}; \quad (5)$$

i.e., $a_4 = 0.072$ waves is equivalent to one Rayleigh unit $\equiv \lambda/(2NA^2)$ of defocus. [At higher NA, when the paraxial assumption breaks down, small amounts of higher-order spherical aberrations are introduced by defocus, but Equation (5) still describes the dominant effect.]

Focus variation is ubiquitous in IC manufacturing because of the combined effects of many problems: wafer nonflatness, autofocus errors, leveling errors, lens heating, etc. A useful lithographic process must be able to print acceptable patterns in the presence of these unavoidable focus variations. Similarly, a useful lithographic process must be able to print acceptable patterns in the presence of variations in the exposure dose. To account for simultaneous variations of exposure dose and focus, it is

useful to map out the "process space," i.e., the exposure-defocus space [10], within which acceptable lithographic quality occurs. Figure 3 shows an example of a process space calculated for a 350-nm line/space grating pattern printed by an aberration-free 0.5-NA projector with $\lambda = 248$ nm and partial coherence $\sigma = 0.6$, and assuming an approximate model for APEX-E resist [11]. The irregular area within the four curves represents the exposure-defocus space where the resist linewidth, or critical dimension (CD), prints within $\pm 10\%$ of the target 350-nm CD. Various rectangular process windows can be defined within this space, such as the one shown in Figure 3, which has an exposure latitude of 15% and a 1450-nm depth of focus (DOF). The process window is the standard measure of the robustness of the process to variations [10, 12].

The effect of Z_4 aberrations is to shift the process window along the focus axis according to Equation (5). Process windows of all pattern types, orientations, and feature sizes are shifted by the same amount. If a_4 varies across the lens field, the surface of best imagery is not planar, and the usable depth of focus (UDOF) is correspondingly reduced [10]. For the example in Figure 3, if a_4 varied by 0.036 waves over the lens field, the best focus position would vary by 250 nm across the lens. Since IC production requires that all parts of the lens print simultaneously, the UDOF at 15% exposure latitude is reduced from 1450 nm to 1200 nm, corresponding to the overlapped area of the two extreme process windows.

• Astigmatism— Z_5, Z_6

These two Zernike polynomials are saddle-shaped OPD surfaces that are positively curved in one direction and negatively curved in an orthogonal direction. The effect of such aberrations on imagery is to cause lines with one orientation to be positively shifted in focus while lines of the orthogonal orientation are negatively shifted in focus. Pure Z_6 introduces a focus difference $\Delta Z_{H/V}$ between lines with horizontal and vertical orientations, given by

$$\frac{\Delta Z_{H/V}}{\lambda/(2NA^2)} = 8\sqrt{6} \cdot a_6 \approx \frac{a_6}{0.051 \text{ waves}}. \quad (6)$$

In a similar way, pure Z_5 causes focus differences between lines of $+45^\circ$ and -45° orientations. The proper combination of a_5 and a_6 can represent a more general astigmatism between two orthogonal lines of arbitrary orientation [13].

Because it is generally necessary to print both line orientations simultaneously, astigmatism has a degrading effect on the process window. Figure 4 illustrates a simulated process window similar to that of Figure 3, but with the addition of $a_6 = 0.05$ waves of astigmatism. The

process window for horizontal lines is identical in size and shape to the aberration-free calculation, but shifted by +243 nm in focus. Vertical lines are similarly shifted by -243 nm, resulting in a relative shift $\Delta Z_{H/V} = 486$ nm. The common process window which can print both orientations is substantially reduced, so that the UDOF at 15% exposure latitude is reduced from 1450 nm to 964 nm. The amount of focus shift for long line structures is dependent only on the line orientation, and not on the linewidth or proximity to other lines.

• Coma— Z_7 , Z_8

Coma occurs when image contributions from different pupil radii shift relative to one another, in contrast to the Z_2 , Z_3 image shift, where all image contributions shift by the same amount. The definition of Y coma in Table 1 can be rewritten as

$$Z_7 = \sqrt{2}(3\rho^2 - 2) \times Z_3, \quad (7)$$

illustrating that the shift for on-axis rays with ρ near 0 is different, and of the opposite sign, from the shift for off-axis rays with ρ near 1. As with image shift, two numbers, a_7 and a_8 , are required to characterize both X and Y coma. It should also be noted that Z_7 and Z_8 represent only the lowest-order coma terms, normally called third-order coma.

A consequence of coma is that symmetric patterns may print asymmetrically. Let us consider the example of a three-bar pattern, where three 250-nm-wide dark lines are separated by 250-nm spaces, imaged by a 0.5-NA projector with $\lambda = 248$ nm. **Figure 5** shows resist profiles simulated by the PROLITH/2™ simulation program² under various imaging assumptions. Figure 5(a) illustrates the aberration-free case, using an ordinary chrome (binary) mask and a partial coherence $\sigma = 0.6$. The image is symmetric about the center of the pattern, and the outside lines are slightly wider than the interior line. Figure 5(b) shows the addition of coma, $a_7 = 0.035$ waves, which causes the left line to be roughly 50 nm narrower than the right line. (Such a linewidth variation would be considered a major problem in an advanced CMOS gate-level process.) Figure 5(c) shows that for the case of $\sigma = 0.3$, the linewidth asymmetry is significantly increased, and the height of the left and right resist patterns is also different. Figure 5(d) shows annular off-axis illumination with $\sigma_{\text{outer}} = 0.7$ and $\sigma_{\text{inner}} = 0.6$. The direction of asymmetry has changed such that the left line is now wider than the right line. This can be understood through the different sign of the image shift for on-axis and off-axis light in Equation (7).

Not only can the linewidth of patterns change because of coma; the center position of the patterns can also

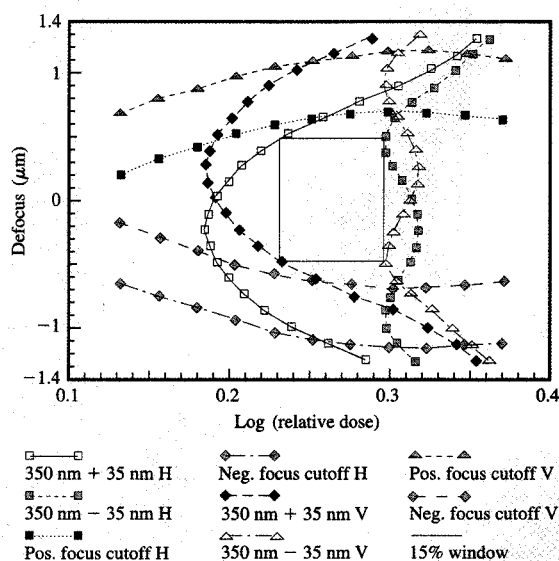


Figure 4

Process windows similar to Figure 3 were calculated assuming $a_6 = 0.05$ waves of astigmatism. The process window for horizontal lines is identical in size and shape to the aberration-free calculation, but shifted by +243 nm in focus. Similarly, the process window for vertical lines is shifted by -243 nm in focus. The common process window which can print both orientations with 15% exposure latitude has 964 nm usable depth of focus (UDOF), a substantial reduction from the aberration-free case in Figure 3.

change. This shift is highly dependent on the details of the mask pattern as well as illumination. For example, a small isolated contact hole with a feature size of $0.5\lambda/NA$ shifts less with coma than a large contact hole with a feature size of $1.5\lambda/NA$, using a $\sigma = 0.3$ projector. Therefore, the presence of coma destroys the concept of a single "lens distortion" map that can be unambiguously measured and applied to any mask pattern. Since relatively large patterns, e.g., "box-in-box," are almost universally used to measure overlay errors in IC processing, coma can cause a relative overlay shift between the measured overlay patterns and the actual device patterns. The current overlay error control scheme in IC processing is based on the assumption that only simple image shifts, i.e., Z_2 and Z_3 , are present.

• Three-leaf clover— Z_9 , Z_{10}

The next two Zernike terms represent OPD surfaces with threefold symmetry. Z_{10} is identical to Z_9 , except that it is rotated by 30° so that the proper combination of a_9 and a_{10} can represent a surface of any desired orientation. The main effect of three-leaf-clover aberration on lithography

² PROLITH/2™ Version 5.0 is a product of FINLE Technologies, Austin, TX.

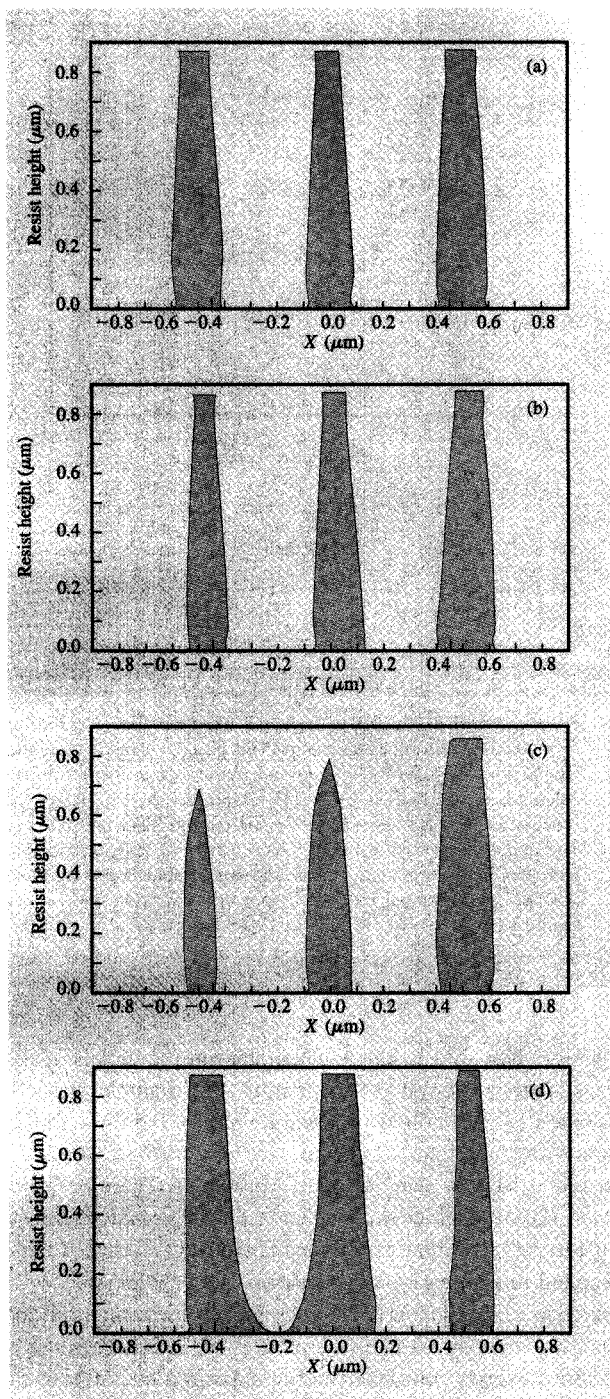


Figure 5

Simulated resist profiles of a "three-bar" pattern with 250-nm linewidths printed using a 0.5-NA projector with $\lambda = 248$ nm: (a) Aberration-free optics with $\sigma = 0.6$. (b) $a_g = 0.035$ waves of coma and $\sigma = 0.6$. The addition of coma has caused a linewidth asymmetry between left and right lines. (c) Same coma with $\sigma = 0.3$ showing enhanced asymmetry. (d) Same coma with annular off-axis illumination where $\sigma_{\text{outer}} = 0.7$ and $\sigma_{\text{inner}} = 0.6$, showing a reversed linewidth asymmetry.

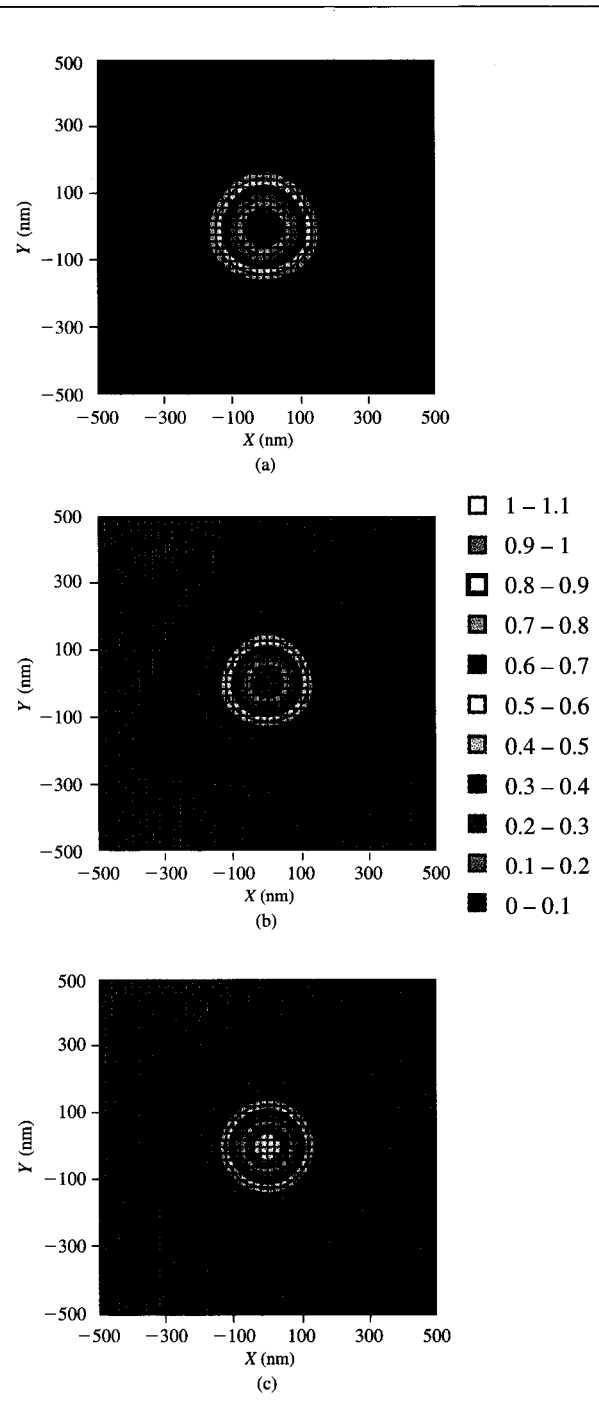


Figure 6

Aerial image contours from simulation of 350-nm contact hole, using a 0.5-NA projector with $\lambda = 248$ nm and $\sigma = 0.3$. Contours are drawn at 10% intensity intervals: (a) Binary mask assuming optics with $a_g = 0.045$ waves of three-leaf clover. (b) 13% attenuated phase-shift mask assuming optics with $a_g = 0.045$ waves of three-leaf clover. (c) 13% attenuated PSM with aberration-free optics.

is to cause undesirable imaging artifacts. One pattern which has high sensitivity to this aberration is the attenuated phase-shift mask (PSM) contact hole. **Figure 6** shows aerial image calculations of an isolated 350-nm contact hole using a 0.5-NA projector with $\lambda = 248$ nm and $\sigma = 0.3$. Figure 6(a) shows the image from a chrome (i.e., binary) mask with $a_9 = 0.045$ waves. While the peak intensity is down a few percent compared with that of an unaberrated image, the image contours show little evidence of an aberration. Figure 6(b) shows the image from a 13% attenuated PSM with the same three-leaf-clover aberration. Surrounding the main contact hole image are three secondary peaks with intensities of more than 0.3, and with the characteristic symmetry of the three-leaf-clover aberration. Figure 6(c) shows the image from the same PSM with no aberration, showing a secondary ring with intensity of roughly 0.22 surrounding the main contact image. Comparison of Figures 6(b) and 6(c) illustrates that the three-leaf-clover aberration breaks up the circular ring of an aberration-free image and concentrates the energy into three spots. Resist processes with insufficient contrast may partially print the secondary peaks [14], causing problems in the final etched contact hole structure.

- *Third-order spherical— Z_{11}*

Just as coma can be viewed as an image shift which depends on pupil radius ρ , so spherical aberration can be thought of as a focus shift which depends on ρ . By rewriting Z_{11} in terms of Z_4 ,

$$Z_{11} = (15/16)^{0.5} \times [(4\rho^2 - 2)Z_4 - 2/\sqrt{3}], \quad (8)$$

it is evident that the focus shift depends on ρ such that, for on-axis rays with ρ near 0, the shift is of opposite sign from that for off-axis rays with ρ near 1. As with coma, the effect of this aberration is highly dependent on the mask pattern and the method of illumination, since these determine which part of the aperture is used in image formation. By changing the pitch of an alternating PSM, different parts of the imaging aperture can be chosen.

Figure 7 displays a plot of focus shift versus feature size of a 1:1 alternating PSM grating, using a 0.5-NA projector with $\lambda = 248$ nm, $\sigma = 0.3$, and $a_{11} = 0.045$ waves of spherical aberration. The small feature sizes diffract light into the outer parts of the aperture, resulting in a focus shift with sign opposite to that of the larger feature sizes, in accord with Equation (8). Ordinary binary mask patterns also exhibit shifts of best focus with different feature sizes and feature types, though the focus shifts are smaller than in Figure 7.

Another aspect of spherical aberration is that the image displays asymmetric behavior through focus. **Figure 8** shows results for a 350-nm line/space grating mask imaged with the 0.5-NA, $\lambda = 248$ nm, $\sigma = 0.3$ projector. Aerial

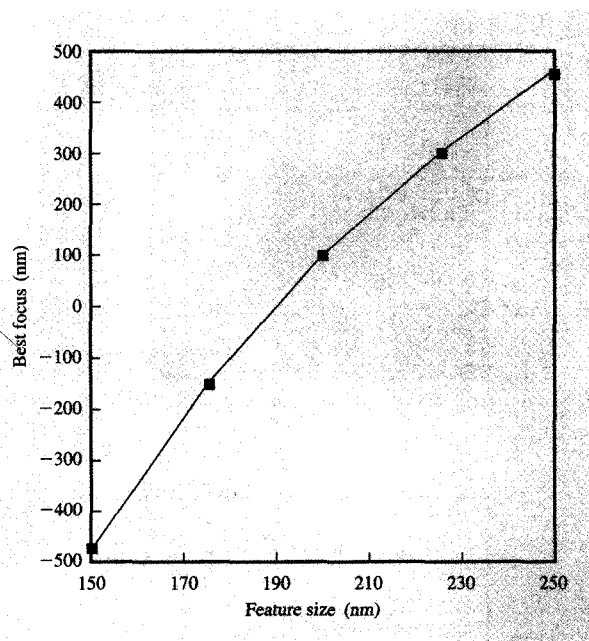


Figure 7

Dependence of best focus position on feature size of alternating PSM when spherical aberration is $a_{11} = 0.045$ waves. A 0.5-NA projector with $\lambda = 248$ nm and $\sigma = 0.3$ was assumed. If spherical aberrations were absent, best focus would be zero, independent of feature size.

image contours in the X - ΔZ plane are shown for an aberration-free projector in Figure 8(a); imagery is symmetric about best focus $\Delta Z = 0$. At roughly $\Delta Z = \pm 1600$ nm, the image is observed to reverse, with the opaque portion of the mask having higher intensity than the clear portion. Figure 8(b) shows similar image contours for $a_{11} = 0.045$ waves of spherical aberration; best focus is shifted upward by several hundred nm, and the image reversal is weakened at defocus $\Delta Z = 1900$ nm but strengthened at defocus $\Delta Z = -1300$ nm. Figures 8(c) and 8(d) respectively display the process window for the aberration-free and aberrated cases. The imaging asymmetry induced by spherical aberration cuts off the process window for negative values of defocus. For the process windows with 15% exposure range, the unaberrated DOF of 1370 nm is reduced by spherical aberration to approximately 1000 nm.

Measuring aberrations with resist patterns

An ideal aberration measurement would measure the OPD surface at many points across the pupil, and then fit with enough Zernike polynomials to represent the surface as a set of coefficients $\{a_1, a_2, a_3, \dots\}$. For a complete lens measurement, this process would be repeated at many

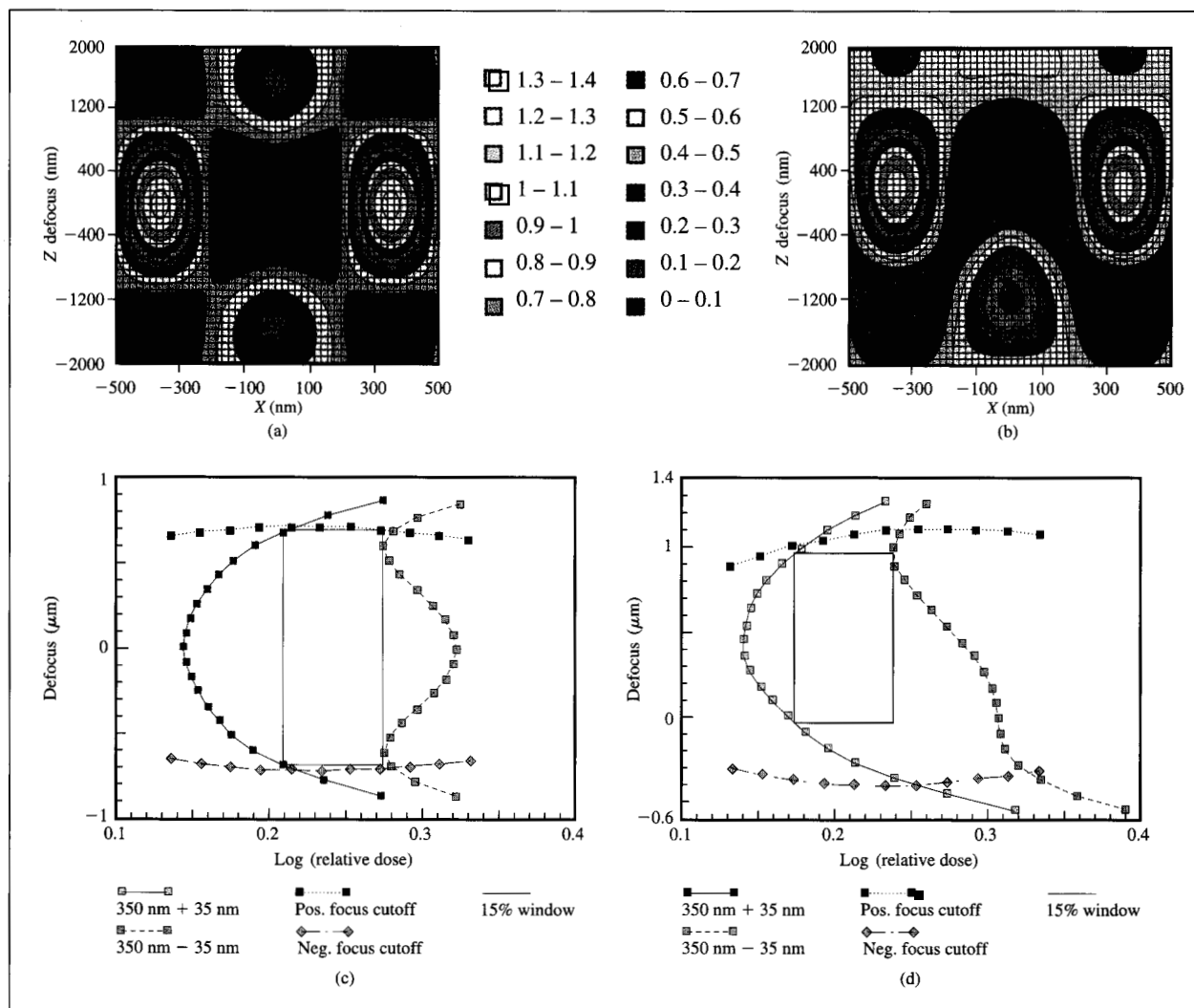


Figure 8

Image calculations for 350-nm line/space grating mask imaged with a 0.5-NA projector with $\lambda = 248$ and $\sigma = 0.3$: (a) Aerial image contours in the X- Δ Z plane from aberration-free simulation. Contours are drawn at 10% intensity intervals. (b) Aerial image contours with spherical aberration $a_{11} = 0.045$ waves. (c) Process window calculated for aberration-free image. (d) Process window calculated for spherical aberration $a_{11} = 0.045$ waves.

points across the lens field, since aberrations are expected to vary slowly across the field. Modern interferometry [15] can achieve such a complete lens characterization, with accuracy better than 0.01 waves. Unfortunately, this technique cannot be applied to a fully assembled lithographic tool. Another method with considerable potential for measuring aberrations *in situ* is the image monitor technique [16], in which the aerial image is measured directly. Although image monitors have been routinely implemented in lithographic tools for the automated measurement of baseline errors and focus setup, they are not generally available for detailed image

characterization and aberration measurement. In most practical situations, a lithographer who wishes to test the lens of a particular lithographic tool has no choice but to print photoresist patterns. In this section, several methods of examining resist patterns to determine aberrations are considered. For all of these methods, it is useful to simulate the particular imaging situation of the experiment, and adjust aberrations to replicate the experimental results.

• Pattern placement

Wide-line patterns (e.g., with $k_1 > 2$) are normally used to measure lens distortion, and it is assumed that all

patterns are shifted the same, as in Z_2 , Z_3 image shifts. However, coma can cause pattern-dependent shifts.

Figure 9 shows the image shift of an isolated clear line, with an opaque background, as a function of linewidth, for a 0.5-NA projector with $\lambda = 248$ nm, $\sigma = 0.3$, and $a_7 = 0.035$ waves coma. It is apparent that, under these conditions, the narrowest line is shifted less by coma than wider lines. This observation can lead to a coma measurement pattern. A "box-in-box" pattern can be designed with the inner box made from a narrow (250-nm) linewidth and the outer box made from a wide (500-nm) linewidth. Coma aberrations would then induce a shift of the center of the inner box with respect to the center of the outer box, a measurement that can be made with a few nm precision by modern optical overlay metrology tools. Since both X shifts and Y shifts are measured, information about both a_7 and a_8 can be obtained. Illumination that is not properly centered in the aperture could also be detected with this pattern by observing the slope of the overlay shift versus focus [16]. Increasing σ to 0.6 results in smaller shifts that depend less on feature size, as shown by the second curve in Figure 9.

• Pattern symmetry

The three-leaf-clover aberration is most clearly observed through the breaking of symmetry. A 13% attenuated PSM imaging small contacts with low σ , as in Figure 6, is a sensitive indicator of lens asymmetry. Sensitivity can be increased by deliberately increasing the exposure dose, i.e., overexposing, so as to bring out relatively small imaging artifacts. The presence of a symmetrical ring around the main contact image is a good indication that asymmetric aberrations are small. Three-leaf-clover aberrations, a_9 and a_{10} , break the ring into three spots. Coma aberrations, a_7 and a_8 , cause one side of the ring to be more prominent than the other side.

Another useful symmetry test uses three-line patterns to search for coma, as in Figure 5. Linewidth differences between the two outer lines are an indication of coma. By orienting such patterns in both horizontal and vertical orientations, one can determine both a_7 and a_8 . It is useful to adjust σ to as low a value as possible, resulting in the most sensitive detection of aberrations.

• Imagery through focus

Many techniques to measure astigmatism and focal plane nonflatness by tracking image performance through focus are well established. In the pin bar technique [13], the best focus is picked out by visual observation of a "microstepped" focus matrix. By measuring lines of different orientation at many locations across the lens field, astigmatism (a_5 , a_6) and focus plane nonflatness (a_4) can be accurately measured.

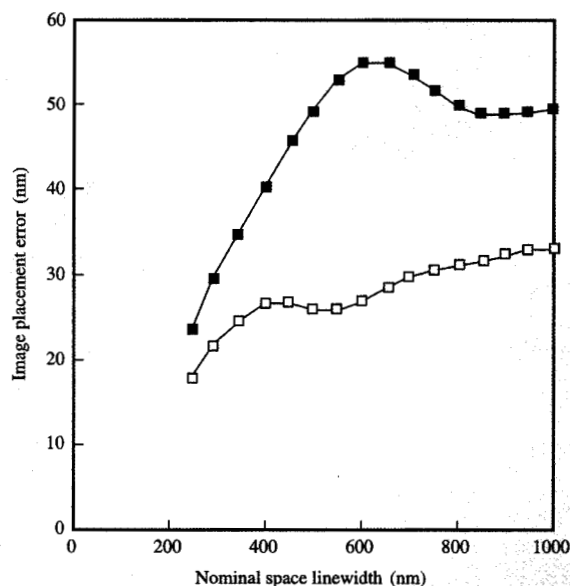


Figure 9

Image shift plotted as a function of the linewidth of an isolated space for a 0.5-NA projector with $\lambda = 248$ nm and $a_7 = 0.035$ waves of coma. The solid squares represent a simulation with $\sigma = 0.3$, while the open squares represent $\sigma = 0.6$. The smaller linewidths are observed to shift less than the larger linewidths, especially when σ is small.

The determination of spherical aberration is a more challenging problem. One approach is to look for a dependence of best focus on feature size. Figure 7 showed such a case using alternating PSM gratings of various sizes, imaged with small σ . Such a PSM is not commonly available, and the linewidths are extremely small. Similar results, albeit with reduced sensitivity, can be obtained with ordinary binary mask gratings. **Figure 10** shows the best focus as a function of feature size for imaging with spherical aberration $a_{11} = 0.03$ waves, at two values of partial coherence σ . The $\sigma = 0.6$ imagery is considerably less sensitive to spherical aberration than the $\sigma = 0.3$ imagery as a result of the greater averaging across the aperture. Unfortunately, real experimental data may also contain effects due to imaging into the relatively thick (e.g., 1000 nm) resist layer. Perhaps ultrathin imaging layers (e.g., 50 nm) might be used to circumvent these difficulties.

• 90° Phase-shift mask patterns—"Focus monitor"

An alternating PSM with phase near 90° possesses unusual optical properties that can be exploited to measure focus errors [17, 18]. It is possible to design a "box-in-box"

pattern, termed the focus monitor, in which the measured overlay error is proportional to the focus error. Focal plane nonflatness can be assessed by measuring focus monitor patterns across the lens field. Astigmatism information appears as differences between the ΔX overlay error and the ΔY overlay error measurement. This technology has proven to be particularly useful for assessing variations in focus across the wafer due to lens heating, misfocusing near the edge of the wafer, and wafer chuck flatness.

The focus monitor pattern is also sensitive to spherical aberration. Full resist simulations were performed to determine the calibration curve (i.e., the overlay shift versus focus offset) of a focus monitor pattern consisting of a 200-nm-wide chrome line with 90° phase shifter to the left and no phase shifter to the right. **Figure 11** plots such curves both with and without spherical aberration, and for two different values of partial coherence. The solid square points, representing an aberration-free projector with $\sigma = 0.5$, are less strongly dependent on focus than the open

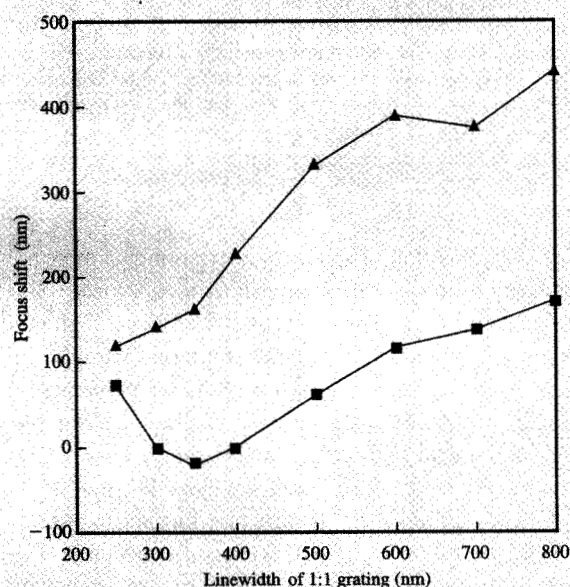


Figure 10

Simulation of the focus shift as a function of linewidth due to $a_{11} = 0.03$ waves of spherical aberration in a 0.5-NA projector with $\lambda = 248$ nm. The process window for equal line/space gratings of various linewidths was calculated, and a focus shift was calculated as the center focus in the process window with 15% exposure latitude. The simulations with $\sigma = 0.3$ (triangular data points) show significantly more variation in focus than the simulations with $\sigma = 0.6$ (square data points). If no spherical aberration were present, the best focus would be zero, independent of linewidth.

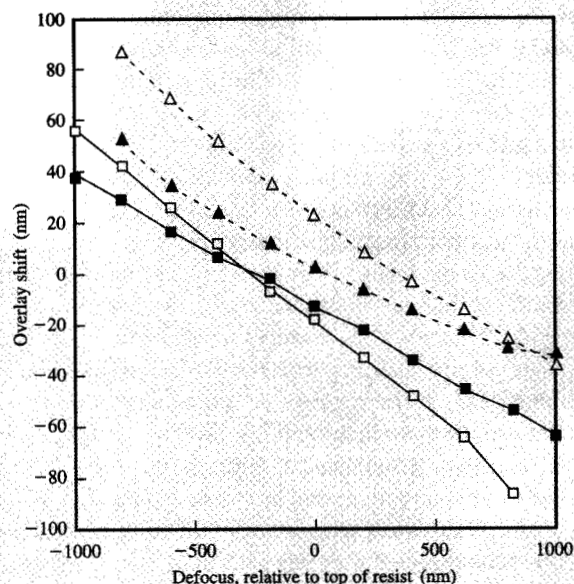


Figure 11

Full resist simulation of calibration curves, i.e., plots of pattern shift versus focus, for a focus monitor with 200-nm chrome linewidth imaged by a 0.5-NA projector with $\lambda = 248$ nm. The two solid lines represent an aberration-free projector where the solid squares assume $\sigma = 0.5$ and the hollow squares assume $\sigma = 0.3$. Note that these two curves cross at roughly zero pattern shift. The two dashed lines represent a projector with $a_{11} = 0.045$ waves of spherical aberration where the solid triangles assume $\sigma = 0.5$ and the hollow triangles assume $\sigma = 0.3$. The crossing point of the two dashed lines occurs at approximately -30 nm of pattern shift.

square points with $\sigma = 0.3$. The two curves cross at approximately zero overlay and focus offset of -250 nm, a focus very close to that for optimum resist imagery. Similar simulations, with an aberration of $a_{11} = 0.045$ waves, are shown in Figure 11 as the triangle data points. The aberrated curves are shifted relative to the aberration-free curves, with a significantly larger shift for $\sigma = 0.3$ than for $\sigma = 0.5$. The aberration has a huge impact on the crossing point of the $\sigma = 0.3$ and $\sigma = 0.5$ curves, moving it to a focus offset of more than 800 nm and an overlay shift of about 30 nm. For lithographic tools with variable σ , measuring the overlay error of the crossing point may provide a sensitive measurement of spherical aberration.

Conclusions

Lens aberrations have been examined through image simulations. A simple test based on the Strehl ratio can be used to verify the accuracy of such calculations. Aberrations cause a variety of problems in lithographic

imagery. Variation of Z_2 , Z_3 aberration across the lens field causes lens distortion, which results in lens matching overlay problems. Coma aberrations Z_7 , Z_8 cause image asymmetries and pattern-dependent overlay errors, which are seldom considered in standard lithographic practice. Variations in best focus Z_4 across the lens field and astigmatism Z_5 , Z_6 are well known to cause reduction in the usable depth of focus. The presence of spherical aberrations, such as Z_{11} , causes the "best focus position" to depend on the particular pattern being projected. Finally, the three-leaf-clover aberration Z_9 , Z_{10} can cause imaging artifacts with threefold symmetry. The present paper has concentrated on the effects of each individual member of the first eleven Zernike aberrations, but similar simulation methods can be applied to any other aberration or combination of aberrations.

The increasing use of advanced imaging techniques such as off-axis illumination or phase-shift masks will motivate tighter control of aberration. Such techniques can put more energy into the outer parts of the aperture, which can cause a greater sensitivity to aberrations. For example, in Figure 6 the ordinary chrome-on-glass contact hole was relatively insensitive to three-leaf-clover aberration compared with the attenuated PSM contact hole. Both coma and spherical aberration were found to cause larger image deviations when σ was small. This is probably due to increased averaging across the aperture when σ is large. In situations where σ is adjustable, one can choose small σ to measure aberrations and large σ for production use, though this certainly oversimplifies the trade-offs. It is hoped that new aberration-measurement techniques, in addition to the ones presented here, will be developed on the basis of simulations of aberrated test patterns.

Acknowledgments

The author thanks R. Ferguson and R. Martino for the software which calculated process windows in this paper. D. Cole and D. Samuels calculated some of the results used in the Strehl ratio accuracy test. J. Kirk and V. Pol of Motorola took time to review this manuscript. Finally, I would like to thank G. Gomba and J. Warlaumont for management support of this work.

PROLITH/2 is a trademark of FINLE Technologies.

References and note

1. H. Ishii and Y. Kawanobe, "A New Generation I-line Lens for 0.35 Micron Lithography," *Proc. SPIE* **2440**, 721-733 (1995).
2. M. D. Levenson, "Wavefront Engineering for Photolithography," *Phys. Today* **46**, pp. 28-36 (July 1993).
3. V. N. Mahajan, "Zernike Circle Polynomials and Optical Aberrations of Systems with Circular Pupils," *Eng. & Lab. Notes*, in *Opt. & Phot. News* **5**, S-21-S-24 (August 1994).

4. D. Flagello, "High NA Imaging in Homogeneous Thin Films," Ph.D. thesis, University of Arizona, Tucson, 1993. Portions of this work can be found in D. Flagello and T. Milster, "3D Modelling of High Numerical Aperture Imaging in Thin Films," *Proc. SPIE* **1625**, 246-261 (1992).
5. V. N. Mahajan, "Strehl Ratio for Primary Aberrations: Some Analytical Results for Circular and Annular Pupils," *J. Opt. Soc. Amer.* **72**, 1258-1266 (1982).
6. M. Yeung, D. Lee, R. Lee, and A. Neureuther, "Extension of the Hopkins Theory of Partially Coherent Imaging to Include Thin-Film Interference Effects," *Proc. SPIE* **1927**, 452-463 (1993).
7. D. Cole, E. Barouch, U. Hollerbach, and S. Orszag, "Calculation of Aerial Images for High NA Projection Systems," *J. Vac. Sci. Technol. B* **10**, 3037 (1992).
8. J. Gortych and D. Williamson, "Effects of Higher-Order Aberrations on the Process Window," *Proc. SPIE* **1463**, 368-381 (1991).
9. K. K. H. Toh and A. Neureuther, "Identifying and Monitoring Effects of Lens Aberrations in Projection Printing," *Proc. SPIE* **772**, 202-209 (1987).
10. B. J. Lin, "The Exposure-Defocus Forest," *Jpn. J. Appl. Phys.* **33**, No. 12B, 6756-6764 (1994).
11. T. Brunner and R. Ferguson, "Approximate Models for Resist Processing Effects," *Proc. SPIE* **2726** (1996).
12. C. P. Ausschnitt, "Rapid Optimization of the Lithographic Process Window," *Proc. SPIE* **1088**, 115-123 (1989).
13. J. P. Kirk, "Astigmatism and Field Curvature from Pin-Bars," *Proc. SPIE* **1463**, 282-291 (1991). J. P. Kirk, "Measurement of Astigmatism in Microlithography Lenses," presented at the OSA Annual Meeting, Toronto, Ont., Canada, October 4, 1993 (abstract only).
14. T. Yasuzato, S. Ishida, and K. Kasama, "Effect of Resist Surface Insoluble Layer in Attenuated PSM for Window Pattern Formation," *Proc. SPIE* **2440**, 804-815 (1995).
15. R. Grosso and R. Crane, "Precise Optical Evaluation Using Phase Measuring Interferometry," *Proc. SPIE* **192**, 65-74 (1979).
16. T. Brunner, S. Cheng, and A. Norton, "A Stepper Image Monitor for Precise Setup and Characterization," *Proc. SPIE* **922**, 366 (1988).
17. R. Mih, A. Martin, T. Brunner, D. Long, and D. Brown, "Using the Focus Monitor Test Mask to Characterize Lithographic Performance," *Proc. SPIE* **2440**, 657-666 (1995).
18. T. Brunner, A. Martin, R. Martino, C. Ausschnitt, T. Newman, and M. Hibbs, "Quantitative Stepper Metrology Using the Focus Monitor Test Mask," *Proc. SPIE* **2197**, 541-549 (1994).

Received February 9, 1996; accepted for publication August 9, 1996

Timothy A. Brunner IBM Microelectronics Division, East Fishkill facility, Route 52, Hopewell Junction, New York 12533 (tbrunner@vnet.ibm.com). Dr. Brunner has been working in the area of optical lithography since 1981, with particular interests in advanced image formation, simulation, and metrology techniques. He received his B.A. from Carleton College in 1975 and his doctorate from MIT in 1980, both in physics. He then worked on the measurement of lithographic tool performance at Perkin-Elmer Corporation and on amorphous silicon TFT processes at Xerox PARC. In 1988, he joined the research staff of IBM; he currently manages the Advanced Lithography and Metrology Department in the IBM SRDC. Dr. Brunner is a member of SPIE, IEEE, and the Optical Society of America. He holds several patents in the area of optical lithography.