INTERFACE '91

This paper was published in the proceedings of the
KTI Microlithography Seminar, Interface '91, pp. 23-35
It is made available as an electronic reprint with permission of KTI Chemicals, Inc.

Copyright 1991.
One print or electronic copy may be made for personal use only. Systematic or multiple reproduction, distribution to multiple locations via electronic or other means, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper are prohibited.

FUNDAMENTAL ISSUES IN PHASE-SHIFTING MASK TECHNOLOGY

by

Chris A. Mack
SEMATECH / DoD
Austin, TX

Chris A. Mack received Bachelor of Science degrees in Physics, Chemistry, Electrical Engineering and Chemical Engineering from Rose-Hulman Institute of Technology in 1982 and a Master of Science degree in Electrical Engineering from the University of Maryland in 1989. He joined the Microelectronics Research Laboratory of the Department of Defense in 1983 and began work in optical lithography research. He has authored numerous papers in the area of optical lithography, regularly teaches courses on this topic, and has developed the lithography simulation programs PROLITH and PROLITH/2. He is currently on assignment at SEMATECH.

ABSTRACT

The technical hurdles for achieving improved lithographic performance with phase-shifting mask (PSM) technology are numerous. Another more subtle problem is that PSM will change our fundamental intuition about lithography. Based on years of experience with binary mask lithography, we have developed a set of unwritten "laws" of lithography, mostly dealing with our notion of scaling: everything gets worse when trying to print smaller features. This simple observation has a profound impact on the way we approach lithography: one simply optimizes a process for the smallest feature of interest and the larger features can be safely ignored. Will this approach work for lithography with phase-shifting masks? This paper will investigate a number of phase-shifting mask techniques and examine their "scaling" properties, that is, how they behave lithographically as a function of feature size. It will be shown that several seemingly unrelated PSM methods, in fact, behave in nearly identical fashion. Thus, only a small number of PSM techniques need to be considered as fundamentally different. Based on modeling of aerial images and resist profiles, a first attempt at a new set of scaling "laws" for PSM lithography will be developed.

INTRODUCTION

The invention of phase-shifting masks for photolithography is generally attributed to Marc Levenson,[1] although mention of the technique is buried in the claims of a patent on x-ray lithography.[2] The so-called Levenson technique (also called the alternating aperture method) employs phase shifts of 0 and 180° in alternating apertures of a periodic structure (e.g., equal lines and spaces). The result is a decrease in the smallest printable feature size by up to a factor of two. This dramatic improvement in resolution is seen as a method of extending the useful life of optical lithography by one to two generations. Interest in phase-shifting masks (PSM) has increased greatly in the past year as several companies have demonstrated the fabrication of prototype circuits using PSM.[3,4]

As work proceeds, people are discovering that the use of phase information in the mask can greatly complicate the design of the mask, as well as our understanding of the imaging process.

This paper will attempt to give insight into the fundamental properties of phase-shifting mask lithography. Using this fundamental understanding,

some general rules of behavior will be developed for PSM which differ greatly in some respects from those for standard binary masks. Alternating aperture and chromeless masks will be discussed in some detail.

CONCEPT OF FREQUENCY DOUBLING

The original alternating aperture PSM technique of Levenson is an example of frequency doubling. To understand how this method works, consider the normal imaging behavior of equal lines and spaces. The simplest example of imaging, shown in Figure 1, uses normally incident monochromatic light from a coherent source impinging on a binary mask of equal lines and spaces. Light passing through the apertures will, of course, be diffracted. Far from the mask, a Fraunhofer diffraction pattern will be formed. This diffraction pattern, $M(v)$, can be easily calculated as the Fourier transform of the mask pattern, $M(x)$.

$$M(v) = \int_{-\infty}^{\infty} M(x)\, e^{-i2\pi vx}\, dx \qquad (1)$$

where x is the distance from the center of the feature of interest (in this case we will pick one of the spaces), and v is the spatial frequency, which is just the sine of the diffraction angle divided by the wavelength, λ. Figure 1 shows a diagram of the diffraction pattern (electric field) resulting from a mask pattern of equal lines and spaces. This diffraction pattern can be expressed mathematically as

$$M(v) = a_0 + 2\sum_{n=1}^{N} a_n\, \delta(v - \frac{n}{p}) \qquad (2)$$

where

$$a_n = \frac{sin(n\pi/2)}{n\pi}$$

and where p is the pitch, N is largest diffraction order possible (corresponding to a diffraction angle of not more than 90°), and δ is the dirac delta function. The nth diffraction order occurs at a spatial frequency of n/p.

The diffraction pattern is a picture of the light distribution as it enters the objective lens of the projection system. This lens can accept light as long as the light enters the lens at an angle such that the sine of this angle is less than or equal to the numerical aperture *(NA)* of the lens. Thus, the largest spatial frequency which can pass through the lens is $v = NA/\lambda$. In essence, the lens acts as a low-pass filter. Higher frequencies are lost, resulting in a loss of information and a reduction in the quality of the aerial image. The ultimate resolution limit of the imaging system depicted in Figure 1 occurs when the first diffraction order just makes it through the outer edge of the lens, i.e., when $p = \lambda/NA$.

The lens accepts the diffraction pattern and acts on it to produce an image. The ideal behavior of a lens is simply an inverse Fourier transform operation. If $M'(v)$ is the portion of the diffraction pattern which makes it through the objective lens, then the electric field of the aerial image is given by

$$E(x) = \int_{-\infty}^{\infty} M'(v)\, e^{i2\pi vx}\, dv \qquad (3)$$

Consider as a typical example feature sizes near the resolution limit. For this case, only the zero and first diffraction orders will pass through the lens and equation (3) will yield

$$E(x) = \frac{1}{2} + \frac{2}{\pi} cos(2\pi x/p) \qquad (4)$$

The intensity is simply the square of the magnitude of the electric field. Figure 1 graphically shows the results of equations (1) through (4). Recall that this figure and the above equations apply to the simple case of coherent, monochromatic light normally incident on a binary mask of small equal lines and spaces.

As the simplest example of frequency doubling, consider a darkfield microscope. The imaging behavior of a microscope is essentially the same as a lithographic projection tool, as described above. A dark field microscope adds an aperture at the entrance pupil of the objective lens which blocks the center of the lens. Thus, the zero order light is

blocked, but the higher order light makes it through the lens. Figure 2 compares the resulting diffraction patterns and aerial images for the normal and darkfield imaging systems. The effect of losing the zero order is to remove the DC bias term in the electric field expression for the aerial image. The aerial image for the darkfield case is

$$E(x) = \frac{2}{\pi} cos(2\pi x/p), \quad I(x) = \frac{4}{\pi^2} \cos^2(2\pi x/p) \qquad (5)$$

As can be seen in Figure 2, the loss of the zero order has a profound effect on the intensity pattern of the aerial image. Without the zero order term, the frequency of the intensity pattern is doubled. Thus, a two micron pitch on the mask will print as a one micron pitch when imaged through a darkfield system. As a result, the ultimate resolution of the darkfield imaging system is twice that of the conventional system due to frequency doubling.

We will now consider two PSM methods for printing periodic line/space patterns and show that their behavior is the same as that of darkfield imaging. As mentioned previously, the alternating aperture method takes a conventional array of lines and spaces and shifts the phase of every other space by 180°. The repeating unit of the alternating aperture mask contains two lines and two spaces and thus the actual pitch of the mask is twice the conventional line/space pitch which prints. Another PSM method for printing lines and spaces is the chromeless mask. In this PSM, alternating patterns of 0 and 180° phase glass are used without any opaque areas on the mask.

In order to qualitatively understand the effect of these two PSM methods, consider the undiffracted light passing through each mask. The zero order is the sum of all of the undiffracted light passing through the mask. In both masks, there will be an equal amount of undiffracted light with phase 0 and phase 180°. When the light is added together, the 0 and 180° light will interfere destructively, completely cancelling each other out. As a result, both phase-shifted masks will have no zero order light. A complete analysis of the diffraction patterns resulting from each PSM shows that both the alternating and chromeless equal lines and spaces give the same diffraction pattern as a darkfield imaging system, differing only in the magnitudes of the higher orders. The magnitude of the first diffraction order for each mask is shown below:

$$\begin{aligned} &darkfield: \quad a_1 = \frac{1}{\pi} \\ &alternating: a_1 = \frac{\sqrt{2}}{\pi} \\ &chromeless: a_1 = \frac{2}{\pi} \end{aligned} \qquad (6)$$

Thus, darkfield imaging, alternating aperture, and chromeless masks all produce frequency doubled images of equal lines and spaces, differing only in the total amount of energy which reaches the wafer.

CHROMELESS MASKS

A chromeless mask is a collection of 0 to 180° phase transitions. Since these types of transitions are commonly found in most if not all PSM schemes, it is useful to study the behavior of chromeless masks extensively. We will begin by examining the behavior of an isolated 0 to 180° phase transition edge.

Isolated Phase-Edge

Consider an isolated phase-edge with an instantaneous transition from 0 to 180° phase. Of course, such transitions will occur in chromeless masks, but they may also occur in alternating aperture, rim shifter, or subresolution type PSM methods. It is well known that a 180° phase transition results in a dark line in the aerial image. In fact, phase-edges have been used to print very high resolution lines in positive photoresist.[5] Some interesting questions arise as to the lithographic properties of such lines. Since the mask contains no information as the width of these lines, what determines linewidth for a phase-edge mask? What is the quality of the aerial image (i.e., what is its log-slope)? These questions can be answered by a straightforward analysis of the imaging of phase-edges.

The diffraction pattern for a perfect phase-edge mask feature can be obtained from the Fourier transform of the mask pattern.

$$M(V) = \frac{}{\pi v} \tag{7}$$

Using this diffraction pattern in equation (3) will give the aerial image for coherent illumination.

$$E(x) = -\frac{2}{\pi} Si\left(\frac{2\pi NAx}{\lambda}\right)$$

$$I(x) = \frac{4}{\pi^2} Si^2\left(\frac{2\pi NAx}{\lambda}\right) \tag{8}$$

where $$Si(\theta) = \int_0^{\theta} \frac{sinz}{z}\, dz$$

Figure 3 shows the resulting aerial image.

What will be the width of a line printed with the aerial image of equation (8)? Typically, the nominal width of a feature occurs where the aerial image intensity is in the range of 0.25-0.3 relative to the intensity in a large clear area. The "width" of the aerial image of a phase-edge can thus be estimated by the 0.25 intensity contour and, from equation (8), corresponds to a width of about

$$W = 0.25 \frac{\lambda}{NA} \tag{9}$$

(Note that for this width, the argument of the Sine integral is about 0.8.) Thus, the width of a feature printed using a phase-edge is determined by the wavelength and the numerical aperture. Note that the feature size defined by equation (9) is a factor of two less than what is generally considered to be the resolution limit of an imaging system. This value of the linewidth can be thought of as the *natural linewidth* of a phase-edge mask feature.

An interesting property of the Sine integral is that for small angles, $Si(\theta) \approx \theta$. The error in this approximation is about 3% when $\theta = 0.8$, and is considerably less for smaller angles. Thus, for the region of the aerial image within the nominal width defined by equation (9), the image becomes nearly quadratic.

$$I(x) \approx \left(\frac{4NAx}{\lambda}\right)^2 \tag{10}$$

This equation makes it very simple to calculate the normalized log-slope of the aerial image at the nominal line edge.

$$w \frac{\partial lnI}{\partial x} \approx 4 \tag{11}$$

Acceptable process latitude occurs when the normalized log-slope is in the range of 3.5 to 4.5 or greater for typical photoresists.

The above discussion describes the aerial image of a phase-edge mask pattern for coherent illumination. Practical projection systems employ partially coherent illumination. Figure 3 compares the coherent phase-edge image with an image generated with a partial coherence of 0.5. The "ringing" effect is reduced when the illumination is made less coherent; however, there is also a reduction in the image edge acuity. As in imaging with binary masks, the main effect of the coherence of the illumination is to change the focus behavior of the image. Figure 4 shows the degradation of the aerial image with defocus for three different illumination systems, $\sigma = 0$, 0.3, and 0.5. The behavior of the coherent image with defocus is dramatically different than that for binary masks. Going out of focus for this case results in a dramatic increase in the resulting linewidth without significant change in the quality of the image. The $\sigma = 0.5$ case shows a more typical behavior. Although the quality of the image decreases with defocus, the nominal linewidth should remain about the same.

The most useful way to investigate the change in linewidth with defocus is with a focus-exposure matrix. Figure 5 shows Bossung curves (linewidth versus focus for different values of exposure) for a phase-edge mask pattern with different values of the partial coherence factor. These curves were simulated with PROLITH/2 using the parameters

shown in Table I. Note that for the case of coherent illumination, the shape of the CD versus focus curves are all about the same, independent of exposure. There is no exposure at which linewidth is relatively independent of focus. The result is very small depth-of-focus. For partially coherent illumination, however, the situation is quite different. For each value of partial coherence there is an exposure at which the CD versus focus curve is virtually flat. The value of this "isofocal" CD and the exposure energy required to obtain it are a strong function of the partial coherence. Thus, the natural linewidth of a phase-edge transition is a function of the defocus-coherence interaction. For $\sigma = 0.45$, the isofocal linewidth is $0.25\lambda/NA$. The resulting process window for ±10% linewidth control is shown in Figure 6. An exposure latitude of ±10% is maintained over a focus range of 1.2μm.

Dark Grating

One issue for chromeless masks is the ability to print large dark areas. This issue can be resolved with a dark grating. As discussed previously, equal lines and spaces in a chromeless mask eliminate the zero order. If the pitch is made small enough, the first order will fall outside of the objective lens aperture resulting in no light being transmitted through the lens. Since the first order occurs at a spatial frequency of $1/p$, a grating of width $0.5\lambda/NA$ or less will print as a large dark area under the condition of coherent illumination. The effect of a partial coherence of σ is to widen the spatial extent of each diffraction order by $\pm\sigma NA/\lambda$. Thus, in order to completely eliminate the first diffraction order from the aperture of the objective lens, the grating width must be

$$w < \frac{0.5\lambda}{(1+\sigma)NA} \qquad (12)$$

What if the phase difference between lines in our dark grating is not exactly 180°? The magnitude of the zero order of the diffraction pattern for a phase error of $\Delta\phi$ from the desired 180° is given by

$$a_n = 0.5\,(1-e^{i\Delta\varphi}) \qquad (13)$$

Thus, if the phase error is zero, there is no zero order energy. If the grating meets the condition of equation (12), only the zero order will pass through the lens. Thus, the resulting intensity in the dark grating area is

$$I_{dark} = sin^2\left(\frac{\Delta\varphi}{2}\right) \qquad (14)$$

A phase error of 16° is required to raise the intensity in the dark grating to 2% of the incident light intensity. This should be well within the manufacturing capabilities of a chromeless mask, making dark gratings very practical.

Edge Superposition

Section A above described the imaging of an isolated phase-edge. The imaging behavior of chromeless features can be computed as the superposition of various edges. For example, consider an isolated line of width w and phase of 180° on a chromeless mask. This feature can be thought of as the superposition of two phase-edges at positions $x = \pm w/2$. Either the diffraction pattern or the electric field of the aerial image can be superimposed, but the intensity can not. Figure 7 shows the resulting aerial images for $\sigma = 0.5$ and different chromeless linewidths. The range of useful linewidths is limited by the center line intensity. For linewidths which are too small or too large, there is not enough destructive interference in the center of the line to keep the intensity low. The range of usable linewidths is about $0.2\lambda/NA$ to $0.35\lambda/NA$. Note that these are the mask dimensions (ignoring any reduction). The dimension of the feature printed on the wafer tends to be roughly twice the size of the mask feature. Larger features can be printed using the dark grating approach as discussed above. At least five lines must be used in a grating in order to get reasonable results.

Equal Lines and Spaces

As discussed previously, chromeless equal lines and spaces give the same result as the alternating aperture for equal lines and spaces with the exception of dose to size. The ultimate resolving

capability of these frequency doubling schemes is greatly improved over binary masks. What is not widely appreciated, however, is the strong bias that these systems produce. It is virtually impossible to print equal lines and spaces with a frequency doubling system. Figure 8 shows the exposure response of chromeless and binary 0.35 μm lines and spaces. At the nominal linewidth, the chromeless features have practically no exposure latitude. At higher doses, however, the chromeless lines tend toward their natural linewidth and exhibit very good process latitude.

CONCLUSIONS

Phase-shifting masks offer the potential to print $0.25\lambda/NA$ size features with acceptable process latitude in single layer resist. While this capability is quite exciting, it is not without problems. First, the use of phase-shifting masks requires a change in the way we approach lithography. With chromeless masks it is easier to print $0.25\lambda/NA$ features than $0.35\lambda/NA$ features and it may be impractical to print both size features at the same time. This simple fact completely destroys our implicit scaling law of lithography (i.e., bigger is easier). Thus, if one is willing to take into account the "natural" linewidth when designing a mask layout, the lithography can be pushed to its limits. If, however, one insists on scaling circuit designs without regard to the lithographic properties of PSM, the improvements in lithographic performance with PSM may not be as great as anticipated.

This paper represents an initial attempt by the author to determine new scaling relations for lithography with phase-shifting masks. The first and most important concept is that of the natural linewidth of a phase transition. In one sense, the natural linewidth is a detriment, inhibiting the ability of a lithographer to print arbitrary size features. In another sense, the natural linewidth offers great potential for dramatically improved resolution for optical lithography if we can learn to design circuits which can take advantage of this inherent lithographic property. Much work remains to be done.

1. M. D. Levenson, N. S. Viswanathan, and R. A. Simpson, "Improving Resolution with a Phase-Shifting Mask," *IEEE Transactions on Electron Devices*, Vol. ED-29, No. 12 (Dec., 1982) pp. 1828-1836.

2. D. C. Flanders and H. I. Smith, "Spatial Period Division Exposing," U.S. Patent 4,360,586 (Nov. 23, 1982). See claim #8.

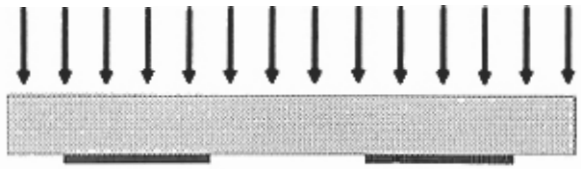
3. K. Nakagawa, M. Taguchi, and T. Ema, "Fabrication of 64MB DRAM with i-Line Phase-Shift Lithography," *IEDM, Proceedings* (1990) pp.817-820.

4. H. Ohtsuka, et al., "Conjugate Twin-Shifter for the New Phase Shift Method to High Resolution Lithography," *Optical/Laser Microlithography IV Proceedings*, SPIE Vol. 1463 (1991) pp. 112-123.

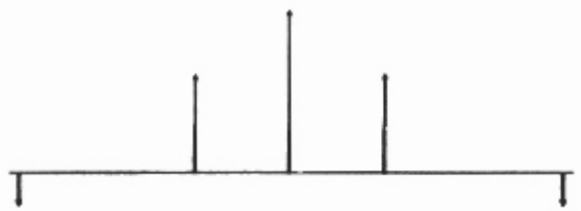
5. H. Jinbo and Y. Yamashita, "0.2 μm or Less i-Line Lithography by Phase-Shifting-Mask Technology," *IEDM, Proceedings* (1990) pp.825-828.

Table I- PROLITH/2 Input Parameters

Projection System:	Resist System: Positive
Wavelength = 365 nm	A = 0. 8 μm^{-1}
Numerical Aperture = 0.5	B = 0.03 μm^{-1}
Partial Coherence = 0.5	C = 0.016 cm^2/mJ
Defocus = 0.	R.I. = 1.70
Flare = 0.	
	R_{max} = 100 nm/s
Development Time = 75 sec	R_{min} = 0.1 nm/s
Non-reflecting substrate	m_{th} = –100
	n = 5
	no surface inhibition



Mask



Diffraction Pattern

$$M(\nu) = a_o + 2\sum_{n=1}^{N} a_n\,\delta(\nu - \frac{n}{p})$$

Aperture

Aerial Image (electric field)

$$E(x) = \frac{1}{2} + \frac{2}{\pi}\cos(2\pi x/p)$$

Aerial Image (intensity)

$$I(x) = \left[\frac{1}{2} + \frac{2}{\pi}\cos(2\pi x/p)\right]^2$$

Figure 1: Pictorial and mathematical description of the imaging process. The example shown is for coherent, monochromatic light normally incident on a binary mask of small equal lines and spaces.

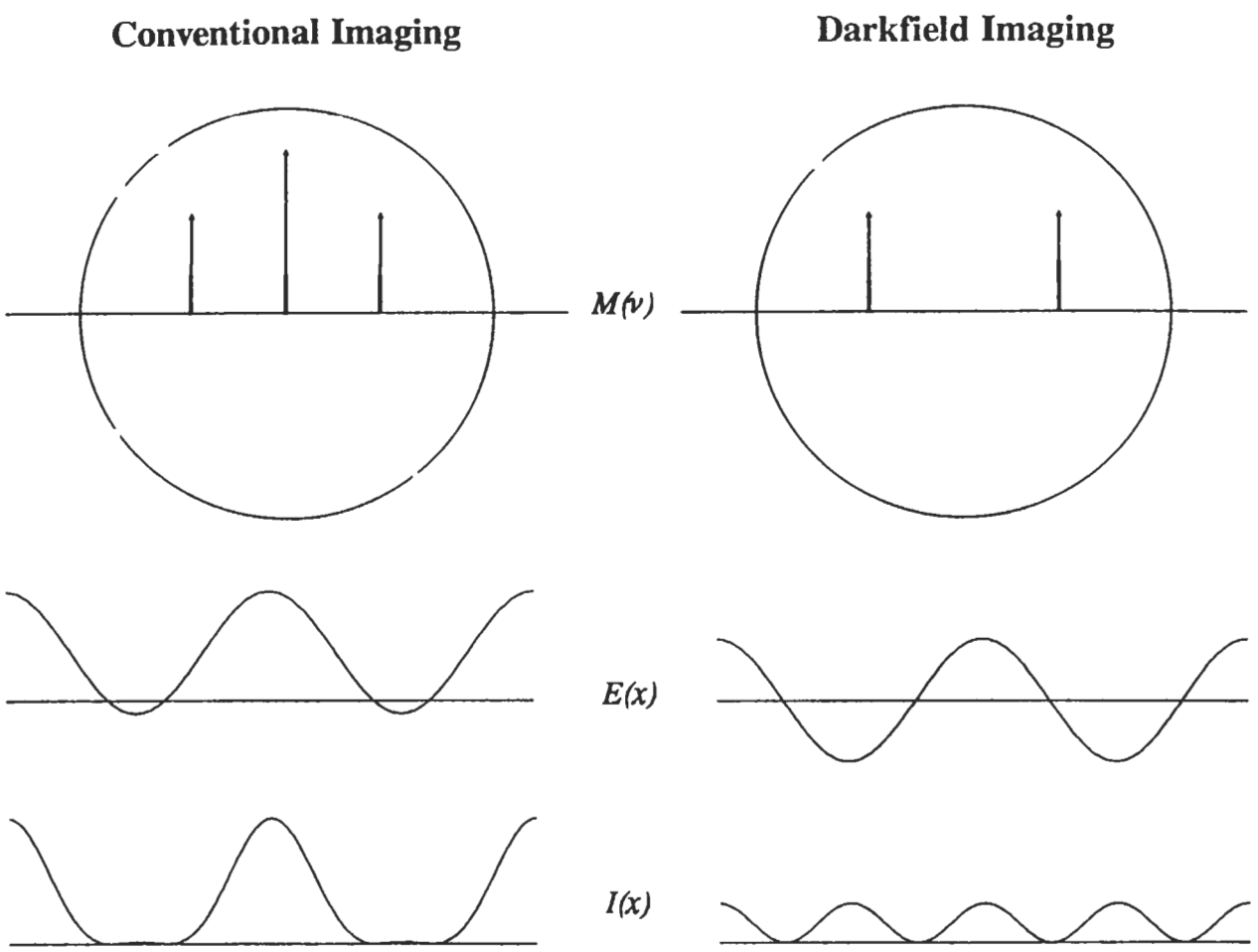


Figure 2: Comparison of normal and darkfield imaging.

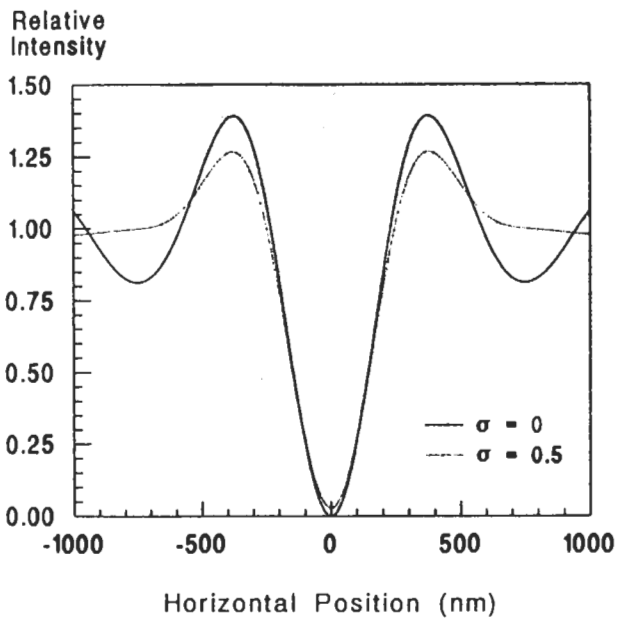


Figure 3: Aerial image resulting from an isolated 0 to 180° phase-edge transition.

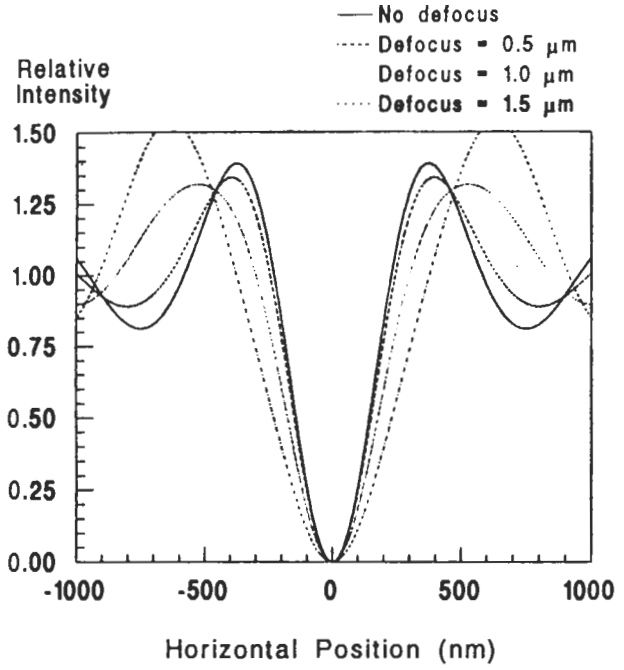


Figure 4a.

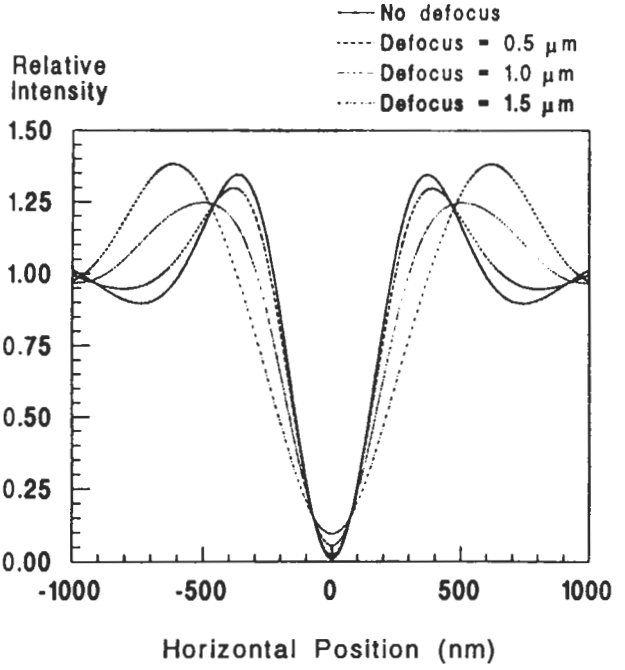


Figure 4b.

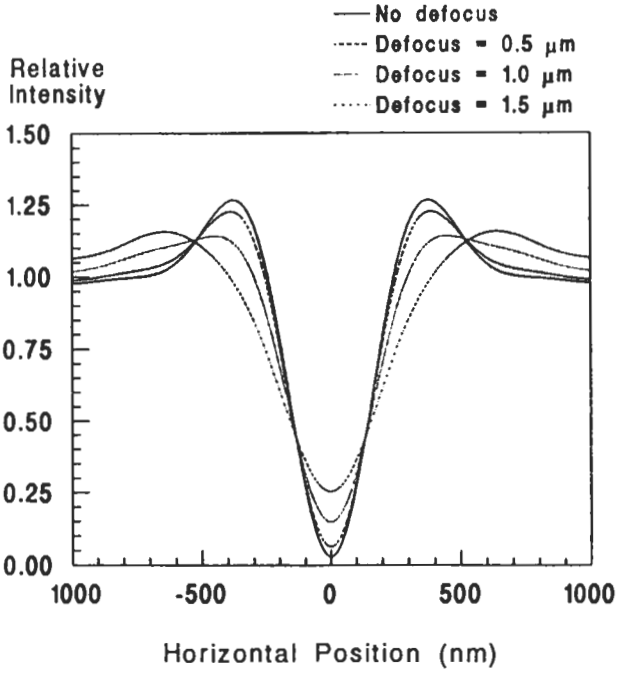


Figure 4c.

Figure 4: Degradation of the aerial image of a phase-edge pattern for a) coherent, b) $\sigma = 0.3$, and c) $\sigma = 0.5$ illumination.

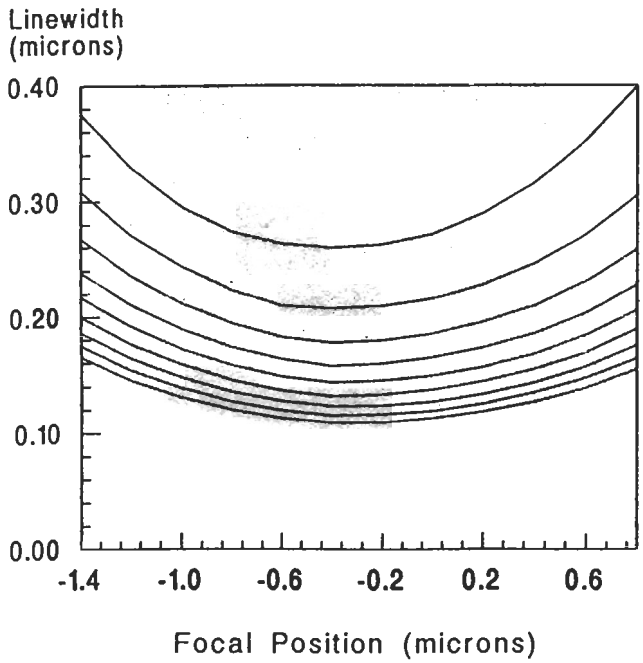


Figure 5a.

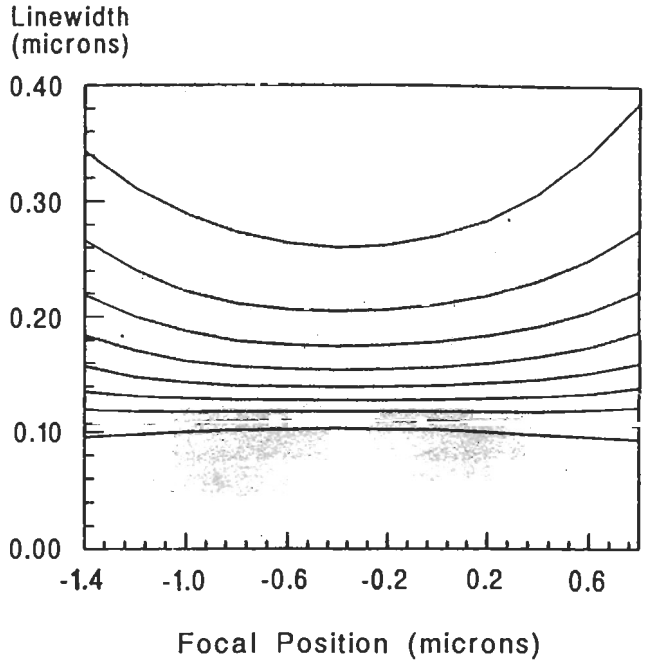


Figure 5b.

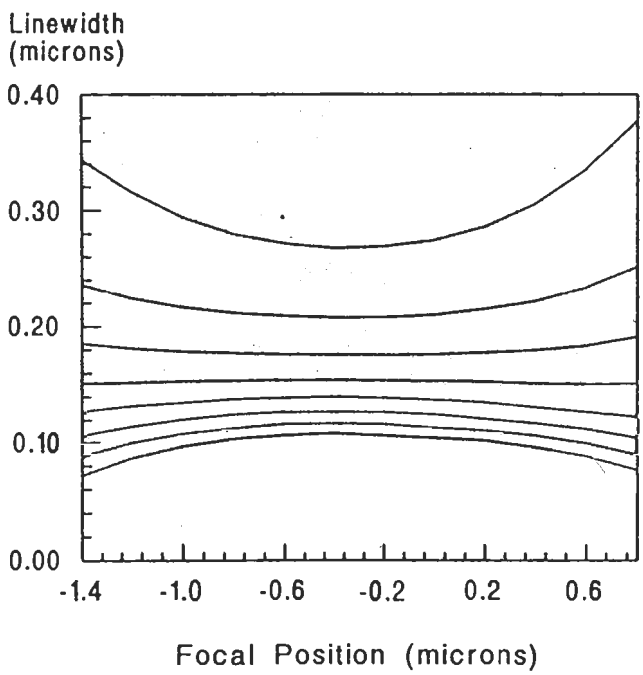


Figure 5c.

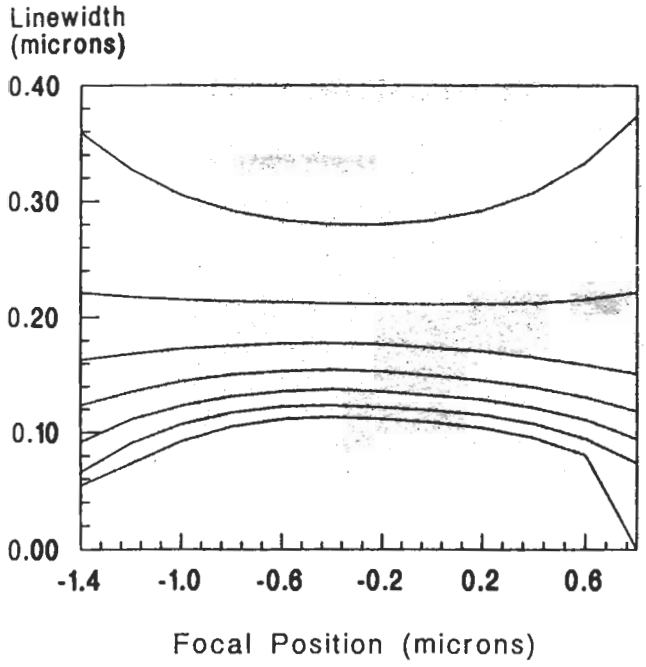


Figure 5d.

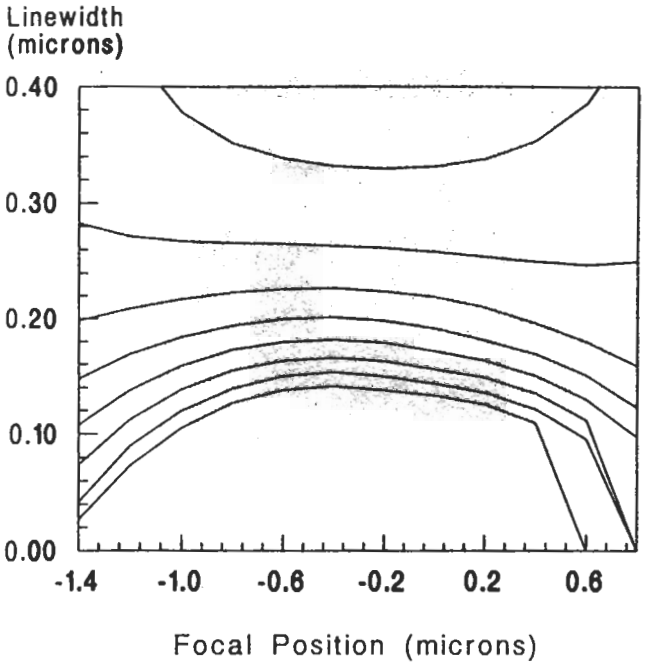


Figure 5e.

Figure 5: Bossung curves for an isolated phase-edge pattern simulated with PROLITH/2 using the parameters in Table I.

a) coherent illumination, exposures starting from 70mJ/cm^2 in increments of 20mJ/cm^2.

b) $\sigma = 0.3$, exposures starting from 70mJ/cm^2 in increments of 20mJ/cm^2.

c) $\sigma = 0.4$, exposures starting from 70mJ/cm^2 in increments of 20mJ/cm^2.

d) $\sigma = 0.5$, exposures starting from 70mJ/cm^2 in increments of 20mJ/cm^2.

e) $\sigma = 0.7$, exposures starting from 70mJ/cm^2 in increments of 10mJ/cm^2.

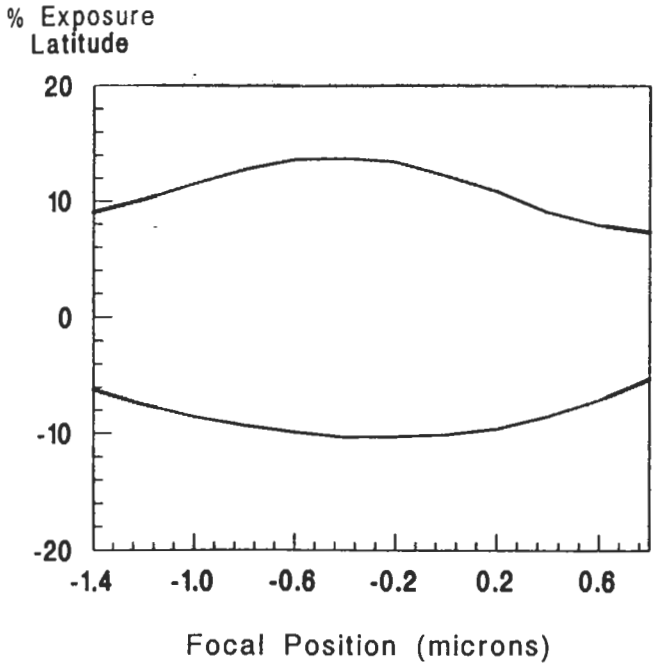


Figure 6: Process window for an isolated phase-edge pattern with $\sigma = 0.45$ and with ±10% linewidth specifications.

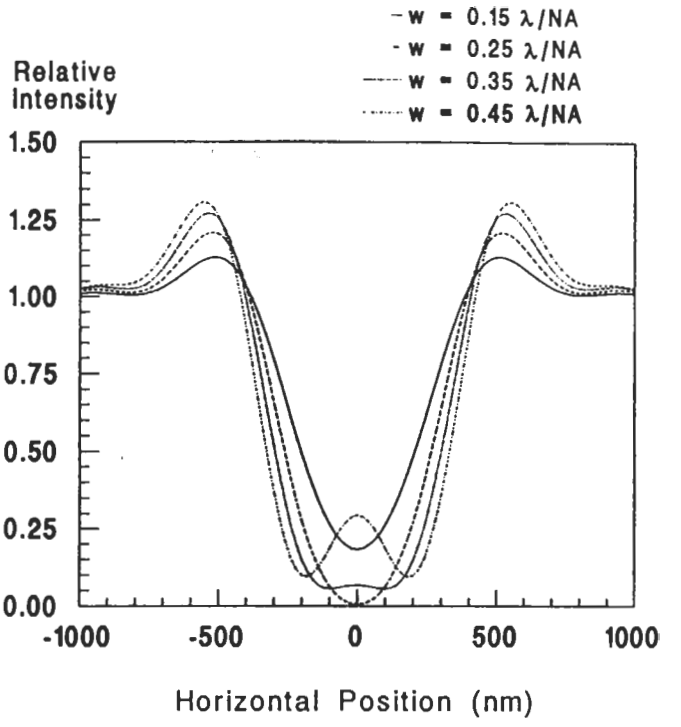


Figure 7: Isolated shifter lines of width $0.15\lambda/NA$ to $0.45\lambda/NA$ ($\sigma = 0.5$).

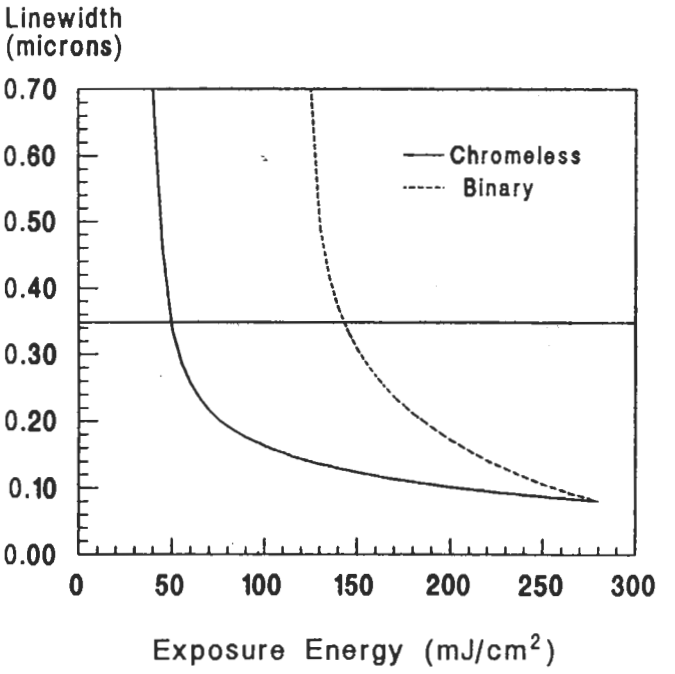
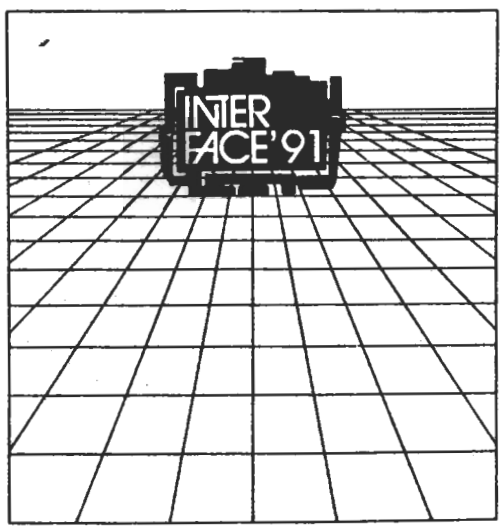


Figure 8: Exposure latitude for a pattern of 0.35 μm lines and spaces with both chromeless and binary masks.

KTI CHEMICALS INCORPORATED

$34.95



KTI MICROLITHOGRAPHY SEMINAR

PROCEEDINGS

Sponsored by KTI Chemicals, Inc.
October 14 - 15, 1991
San Jose, California