



HOKKAIDO UNIVERSITY

Title	Study on Optical Amplifying Aberration Measurement for Semiconductor Optical Aligner
Author(s)	諏訪, 恭一
Degree Grantor	北海道大学
Degree Name	博士(工学)
Dissertation Number	乙第6955号
Issue Date	2015-03-25
DOI	https://doi.org/10.14943/doctoral.r6955
Doc URL	https://hdl.handle.net/2115/63028
Type	doctoral thesis
File Information	Kyoichi_Suwa.pdf



Study on Optical Amplifying Aberration Measurement for Semiconductor Optical Aligner

Kyoichi Suwa

February 14, 2015

Contents

1	Introduction	1
1.1	Progress of Semiconductor	1
1.1.1	Planar Transistor Innovation	1
1.1.2	Line Shrink Trend	3
1.2	Progress of Lithography Optics	4
1.2.1	Imaging Theory	4
1.2.2	Lithography Optics	7
1.3	Lens Aberrations	9
1.3.1	Seidel Aberrations	9
1.4	Aberration Measurement	14
1.4.1	Star Image Test	14
1.4.2	Interferometer Test	15
1.4.3	Aerial Image Test	17
1.5	Aberration Measurement by Photoresist	17
1.5.1	Scanning Electron Microscope Issues	17
1.5.2	Non SEM Focus Measurement	18
1.6	Coma Aberration Problems	21
1.6.1	Shift Coma Aberration	21
1.6.2	Coma Aberration Expression	22
1.6.3	Typical Coma Measurement by Photoresist	22
1.7	Distortion and Registration Measurement Problems	23
1.8	Optical Image Amplification	26
1.8.1	Study Motivation	26
1.8.2	Optical Image Amplification of Aberration Measurement	27
1.8.3	Study Target	29
1.9	Overview of Chapters	30

2	Coma Aberration Measurement	31
2.1	Coma Intensity Amplification	31
2.1.1	Search for Coma Sensitive Image	31
2.1.2	Intensity Amplification Model	34
2.2	Coherent Imaging Analysis	38
2.2.1	Coherent Imaging Formula for Aberration	38
2.2.2	Sidelobe Peak Positions	40
2.2.3	Intensity Difference between Pair of Local Peaks	42
2.3	Aberration for Single Line	44
2.3.1	Single Line Positive Pattern	44
2.3.2	Single Line Negative Pattern	48
2.3.3	Simulation Results	49
2.4	Aberration for Plural Line	50
2.5	Effect of Even Wavefront Aberrations	54
2.6	Summary	55
2.7	Chapter Appendix	56
2.7.1	Nature of the Integral Sine Function	56
2.7.2	Sensitivity on Duty Ratio and Illumination Coherence	57
3	Position Measurement Metrology	58
3.1	Metrology Sensor used in this study	58
3.2	Position Measurement	58
4	Focus Measurement	61
4.1	Focus Wedge Mark	61
4.2	Defocus Image Analysis	63
4.2.1	One-Dimensional FWM Model	65
4.2.2	Focus Scalar Coherent Imaging Theory	66
4.2.3	FWM Metrology	69
4.2.4	Numerical Calculation	69
4.2.5	Verification of Focus Sensitivity	71
4.3	Best Focus Function	73
4.3.1	Defocus Dependence Specified by Cosine Curve	73
4.3.2	Future Application to Latest Exposure System	73
4.4	Summary	75

5	Linewidth Measurement	76
5.1	Double Exposure Magnification	76
5.2	Simulation and Experiments	77
5.2.1	Simulation Results	77
5.2.2	Experimental Results	79
5.3	Automated Optical Measurement	81
5.3.1	Offset in Measurement	81
5.3.2	Measurement Precision	82
5.4	Summary	82
6	Displacement Measurement	84
6.1	Moire Displacement Metrology	84
6.2	Alternating Direction Moire	85
6.2.1	Bar-in-Bar Method	85
6.2.2	Introduction to Alternating Direction Moire	87
6.3	Registration Measurement with Photoresist	87
6.3.1	Basic Concept in Practice	87
6.3.2	Error Analysis of Measurement	89
6.4	Experimental Results	91
6.4.1	Experimental Data	91
6.4.2	Verification of Assumptions	93
6.4.3	Future Precision Improvement	93
6.4.4	Non-Photoresist Application	95
6.5	Summary	95
7	Conclusion	97
7.1	Summary of Findings	97
7.2	Concluding Remarks	97

Chapter 1

Introduction

1.1 Progress of Semiconductor

1.1.1 Planar Transistor Innovation

In 1959, Kilby and Noyce independently filed the innovative patents that embedded the transistor in semiconductor surface [1, 2].

This invention implied a lot of transistor would be set in the unit area just beneath the surface of the silicon. This meant the huge downsizing was possible. Five years after the inventions, Moore said in a journal that the number of transistor integrated in the circuit had increased four times in every three years through past years' experience. He had a comment that; therefore, this trend would continue in the future [3, 4]. It is expressed as

$$p = 2^{n/1.5}, \quad (1.1.1)$$

where p is the relative value of the number of the transistor, and n is the number of years.

It is easily calculated that the number of the integrated transistor increases 10 times in five years, 102 times in ten years, and 10,321 times in 20 years, as shown in Fig. 1.1.1. Therefore, the Moore's Law has taken on a rule in the growth guide-line of semiconductor industries.

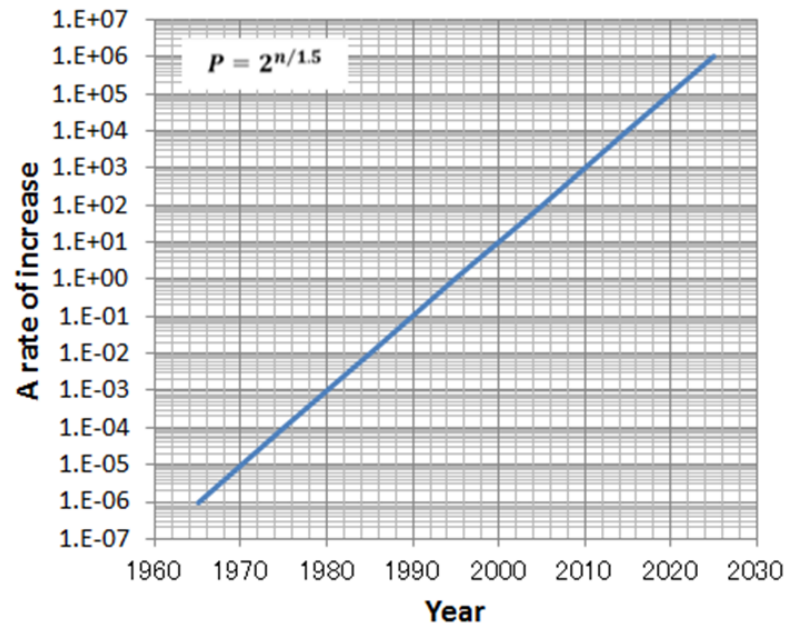


Figure 1.1.1: The Moore's Law: the abscissa shows the year and the ordinate shows the integrated transistor number as normalized as one in 1995.

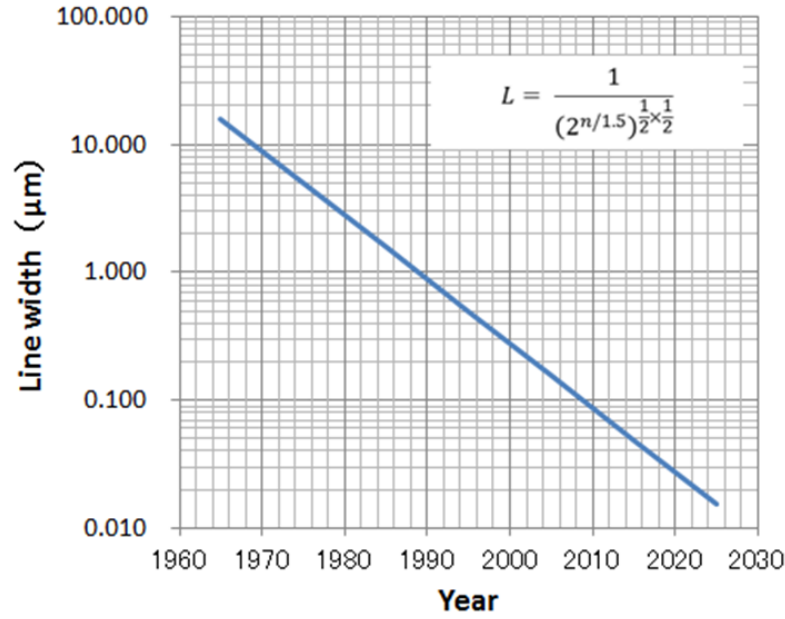


Figure 1.1.2: Linewidth derived from the Moore's Law: the abscissa shows the year and the ordinate shows the linewidth as $0.5 \mu\text{m}$ in 1995.

1.1.2 Line Shrink Trend

The progress of the semiconductor technology has been empirically tracing the Moore's Law. Among the shrink factor, the printed linewidth shrank almost $1/\sqrt{2}$ every three years in the past. The linewidth L is expressed for each x and y direction by

$$L = 1/p^{1/2 \times 1/2}, \quad (1.1.2)$$

where x and y linewidth contribution in the factor of Eq. (1.1.2) is $1/2$ and the other's bit increasing contribution is also $1/2$. Supposing in 1995 the leading edge linewidth is $0.5 \mu\text{m}$, it is calculated as $15.8 \mu\text{m}$ in 1965, $5.0 \mu\text{m}$ in 1975, $1.58 \mu\text{m}$ in 1985, 160 nm in 2005, 50 nm in 2015 and 16 nm in 2025. That means the linewidth has shrunk $1/31.55$ in thirty years, and this means the optical exposure system has achieved $1/31.55$ linewidth narrowing; therefore, optics has inevitably become the most key functional role in the semiconductor manufacturing field.

1.2 Progress of Lithography Optics

In this section, the history of the imaging theory and the lithography optics are described.

1.2.1 Imaging Theory

Willebroed Snell (1591-1626) empirically discovered the long hidden Law of Refraction in 1621; this was one of the great moment in optics [5]. Through the observation, he unveiled that the bending of the rays through different media of refractive index n showed the equation which was constant for all θ_i as

$$n_i \sin \theta_i = n_t \sin \theta_t, \quad (1.2.1)$$

where θ_i is the incident angle, and θ_t is the refractive angle at different media. In optics, “sin θ refraction” phenomenon became a key function for developing optical theories.

In the 19th century, L. J. Mande Daguerre innovated the practical photography “picture” system using immersion glass plate. In those days astronomical telescope, binocular, microscope, and camera were the trigger to develop the precise, wide field, and high speed optical lenses.

In 1850s Ludwig von Seidel (1821-1896) studied the ray refraction theory in detail for the first two terms of expansion as

$$\sin \theta = \theta - \theta^3/3! + \theta^5/5! - \theta^7/7! + \dots . \quad (1.2.2)$$

When the first two terms are retained as an improved approximation, we have called it third-order-theory that is embodied in five primary aberrations (spherical, coma, astigmatism, field curvature, and distortion). Accordingly, they are frequently referred to as the Seidel aberrations [6]. This third order theory is used basically to explain the whole aberrations; that is, the chromatic light ray refracts at the spherical surface to the flat focal plane or along the chief optical axis.

Although the theoretical resolution was derived by former scientists, Abbe was the first to reach the conclusion experimentally who credited the resolution limit of a microscope, and his formula (published in 1873) [7] is

$$R = \lambda/2NA, \quad (1.2.3)$$

where R is the resolution limit, λ is the wavelength, and the variable NA is numerical aperture defined by

$$\text{NA (Numerical Aperture)} = n \sin \theta. \quad (1.2.4)$$

The term numerical aperture is the sine of the half angle multiplied by the refractive index “ n ” of the media filling the space between the image plane and the front lens. Since his equation has been published, NA became a popular definition.

In 1950, Hopkins described in his paper [8] that the resolving separation of two points, R is given by

$$R = K_1 \cdot \lambda / n \sin \theta, \quad (1.2.5)$$

where K_1 , as known factor today, is a coefficient of R . K_1 is different in several illumination modes where he described the definition of σ as

$$\sigma = (\text{NA condenser}) / (\text{NA objective}). \quad (1.2.6)$$

The K_1 factor along with σ expanded the concept of Abbe in Eq. (1.2.3) to a general resolution criteria.

In 1950s, several works began to implement optics with the mathematical techniques and insights of communications theory. The tools of frequency analysis and linear systems theory have played important roles together with a wavefront aberration that represents the path difference of the ray as the phase difference from the ideal wave propagation.

It is useful to use Fourier transform in the imaging theory [5, 9], that can handle also the phase object which have not been handled by the ray optics; therefore, this is called Fourier optics.

When the pattern dimension of the circuit decreases to the wavelength scale, the imaging properties cannot be simulated in the geometrical optics. This is because, in wavelength scale region, the wave properties of the light, diffraction phenomena, play a dominant role in the imaging. Assuming the situation that the incident light is diffracted by the aperture, the image is formed by calculating the diffracted light at the far field. By applying the appropriate boundary conditions and approximations, we understand that the far field light is expressed by Fourier transformation of the aperture. Therefore, as long as the far field light, the imaging optical system is mathematically modeled by Fourier transformation [10].

As one dimensional equation of Fourier optics, the intensity profile of the image under coherent illumination can be written as

$$I(x) = \left| \int_{-\text{NA}}^{\text{NA}} d\xi \tilde{T}(\xi) G(\xi, 0) e^{ik\xi x} \right|^2, \quad (1.2.7)$$

where the spectrum $\tilde{T}(\xi)$ is given as Fourier transform of the amplitude of one-dimensional pattern $T(x)$

$$\tilde{T}(\xi) = \int dx T(x) e^{-ik\xi x}, \quad (1.2.8)$$

where $\xi(\eta)$ is the directional cosine on the $x(y)$ axis of the unit vector of the propagating wave direction. The optical pupil function $G(\xi, \eta)$ with wavefront aberration $W(\xi, \eta)$ can be represented as

$$G(\xi, \eta) = \exp[ikW(\xi, \eta)], \quad (1.2.9)$$

where $k = 2\pi/\lambda$ is the wavenumber.

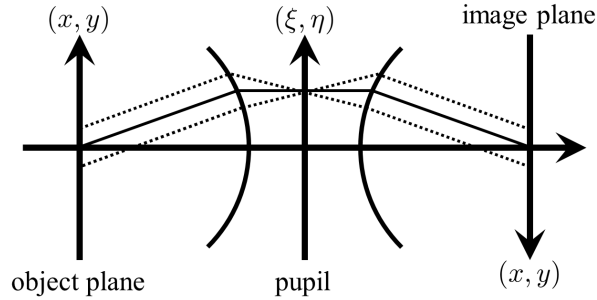


Figure 1.2.1: Coordinate diagram of the coherent image coordination corresponding to Eq. (1.2.7) [74].

The first laser was built in 1960; the great advantage for optics is the light source for interferometer of the lens surface curvature measuring metrology. Frequency continuously stabilized laser is suitable for position measurement tool, so that the precision metrology has been dramatically improved; therefore, it could build semiconductor optical aligners precisely.

One of the most useful laser application for the optical aligner is the excimer laser which have been developed for the exposure light sources at deep *UV* region, 248 nm (*KrF*) and 193 nm (*ArF*) individually.

Since the planar transistor was innovated, the precise optics was inevitably required for the projection lens of the semiconductor optical aligner which exposed the circuit patterns and aligned the patterns to the former fabricated patterns in semiconductor process.

The technologies developed in the precise optics which was used for microscope, astronomical telescope and military optics were very rapidly transferred to the optical aligner, at the same time, suitable projection lens and also illumination configuration were highly required.

1.2.2 Lithography Optics

Since 1970s, semiconductor device circuit patterns fabricated on the glass masks have been projected onto a wafer by high quality manufacturing lenses [11]. This lens is called as the projection lens which has well corrected aberration that is less than one-twentieth of the exposure wavelength. The projection lens is primarily illuminated by a monochromatic light source such as a mercury lamp or an excimer laser.

In semiconductor optical lithography field, the progress of lithography optics has been so rapid, and there are a few literature overviewing the optics and types of lens from 1960s to 2000s. Bruning had reviewed “Optical Lithography - Thirty years and three orders of magnitude” in which he described many types of optics and features in an excellent way [12].

Kingslake explained a Dyson Catadioptric System (1959), in which the optical configuration was made by one concave spherical mirror and one unique refractive lens block as co-axial arrangement [13]. This 1 : 1 magnification system had the limited image area with no Seidel aberrations. In 1975, an Offner Catoptric system was introduced similar to the Dyson arrangement for the same purpose; however, a refractive lens is replaced to a small convex mirror. Kidger described each type of lens from a lens ray tracing point of view [14].

From a point of view of resolution, there are several groups in order of increasing NA as follows:

1) NA is from 0.05 to 0.2: As described above, genius two persons found the Seidel aberrations free imaging concept of 1:1 magnification which were the major exposure optical aligners equal to or larger than 1.5 μm resolution

area.

2) NA is over 0.28: All refractive type lenses were dominant, that was called as a Dioptric which requested many refractive spherical lens to compensate residual aberrations, in which the chromatic aberration was only corrected so narrow bandwidth spectrum that the NA can increase larger than that of the former system.

Therefore, the lens configuration was so familiar to camera lenses that several companies could deliver the optical aligners with such projection lens, also in case of Nikon [15–17] as shown in Fig. 1.2.2.

3) NA is over 0.75: All refractive types with adding some aspheric surfaces lenses have been developed for KrF and ArF excimer light sources.

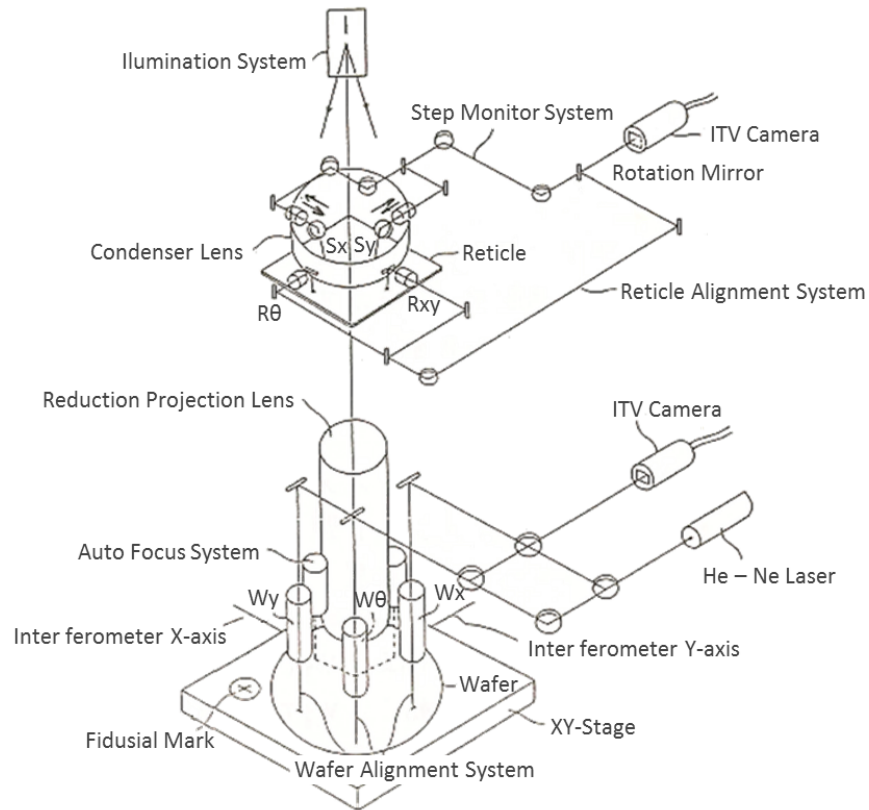


Figure 1.2.2: Optical Aligner configuration; NSR 1010G in 1980 [15].

1.3 Lens Aberrations

In order to understand the aberrations of lithography lens and measurement technologies, it is useful to describe the Seidel five key aberrations in polynomials.

1.3.1 Seidel Aberrations

Conrady described the Seidel five aberrations as the polynomials except the field curvature on the conditions that all aberrations come from the spherical aberration in his paper [18–20]¹. Kubota cited the expression for his optical textbook [21, 22] which is referring in this thesis, the Seidel four aberrations except the field curvature, are described as

$$x = m_3 r \theta \sin \psi + m_2 r^2 \sin 2\psi + m_1 r^3 \sin \psi, \quad (1.3.1)$$

and

$$y = m_4 \theta^3 + 3m_3 r \theta^2 \cos \psi + m_2 r^2 (2 + \cos 2\psi) + m_1 r^3 \cos \psi, \quad (1.3.2)$$

where x and y show the aberration of the blur from the ideal point on the image plane, m_1 is the coefficient for spherical aberration, m_2 is the coefficient for coma aberration, m_3 is the coefficient for astigmatism aberration, and m_4 is the coefficient for distortion aberration. The (r, ψ) represents the projected polar coordinate of the field aperture on the lens surface. The coordinate relation is described in the reference [21], in which the “ θ ” is an angle of the ray from the optical axis to the first entrance surface of the lens that represents the image height position on the image plane. The “ r ” represents the lens radius which directly represents the NA of the projection lens. If there is no aberration, the point spread image is shown as the star test in Fig. 1.3.1.

Supposing only m_1 exists, it represents the spherical aberration:

$$x = m_1 r^3 \sin \psi, \quad y = m_1 r^3 \cos \psi, \quad (1.3.3)$$

therefore:

$$x^2 + y^2 = (m_1 r^3)^2. \quad (1.3.4)$$

¹A given condition is that the spherical aberration on the optical axis is proportional to the power of 3 regarding the incident angle from the optical axis to the spherical surface.

The cause to reduce the peak contrast is the spherical aberration which is shown as the point image in Fig. 1.3.2.

Supposing only m_2 exists, it represents the coma aberration:

$$x = m_2 r^2 \sin 2\psi, \quad y = m_2 r^2 (2 + \cos 2\psi), \quad (1.3.5)$$

therefore:

$$x^2 + (y - 2m_2 r^2)^2 = (m_2 r^2)^2. \quad (1.3.6)$$

The coma aberration makes the point image be a comet shape towards the direction of the coma aberration in Eq. (1.3.5), in which the center weight point of image moves to the y direction in Eq. (1.3.6), that is, the coma aberration causes the shift of image position, and it also causes the linewidth abnormality of the group lines, therefore, the coma is definitely the worst aberration for lithography optics. It is shown as the point image in Fig. 1.3.3.

Supposing only m_3 exists, it represents the astigmatism aberration:

$$x = m_3 r \theta \sin \psi, \quad y = 3m_3 r \theta^2 \cos \psi, \quad (1.3.7)$$

therefore:

$$x^2 + \frac{y^2}{3^2} = (m_3 r \theta^2)^2. \quad (1.3.8)$$

where it causes the different focus points of the tangential and sagittal directions along the z axis. The astigmatism decreases the peak contrast as shown in Fig. 1.3.4; however, it does not cause the point image asymmetry so that it does not cause the shift of the image weight, therefore it is different from that of the coma aberration.

Supposing only m_4 exists, it represents distortion aberration:

$$y = m_4 \theta^3. \quad (1.3.9)$$

The distortion aberration is defined by the grid shift from the ideal square of the image as shown in Fig. 1.3.5. In the semiconductor optical aligner, the projection lenses need to be designed and to be adjusted as possible as they can be close to the ideal grids. Because the difference between the projected pattern grids by the distortion of the optical aligners causes the displacement error for the device process; however, the microscope objective

lenses do not care the distortion aberration, which is clearly different from that of the optical aligner.

The last Seidel coefficient for aberration is the field curvature which is represented by m_5 , where the rigorous definition of field curvature is the difference between the spherical image plane and the flat image plane as shown in Fig. 1.3.6; however, for practical lithography lenses, it shows several unexpected waving field curvatures by design and manufacturing results.

In the description of my study, m_1 , m_2 , m_3 , m_4 , and m_5 are used as the symbol of the spherical, the coma, the astigmatism, the distortion and the field curvature aberration, respectively.

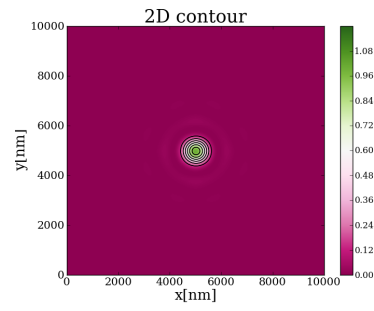


Figure 1.3.1: Star test without aberration.

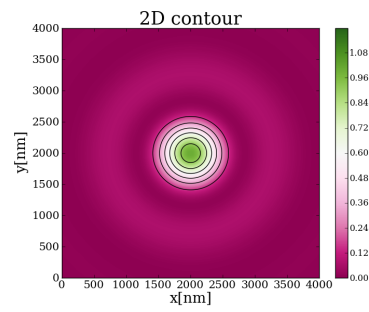


Figure 1.3.2: Star test with spherical aberration 250 mλ.

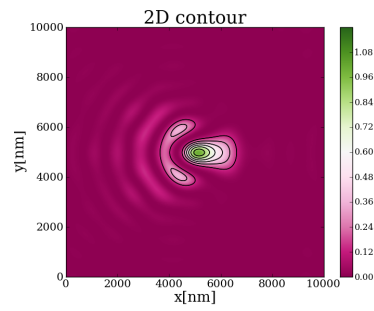


Figure 1.3.3: Star test with coma aberration 500 m λ .

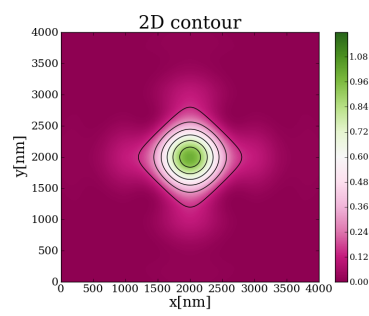


Figure 1.3.4: Star test with astigmatism aberration 250 m λ .

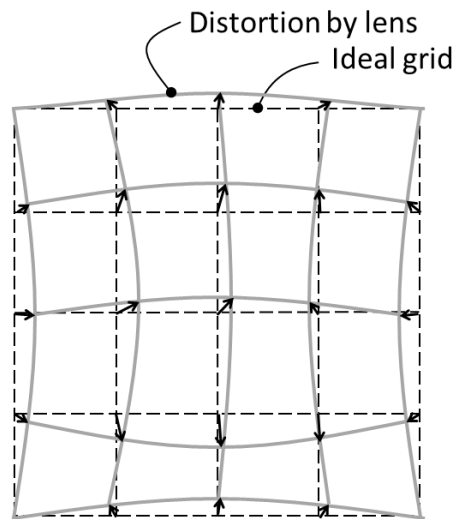


Figure 1.3.5: Distortion aberration.

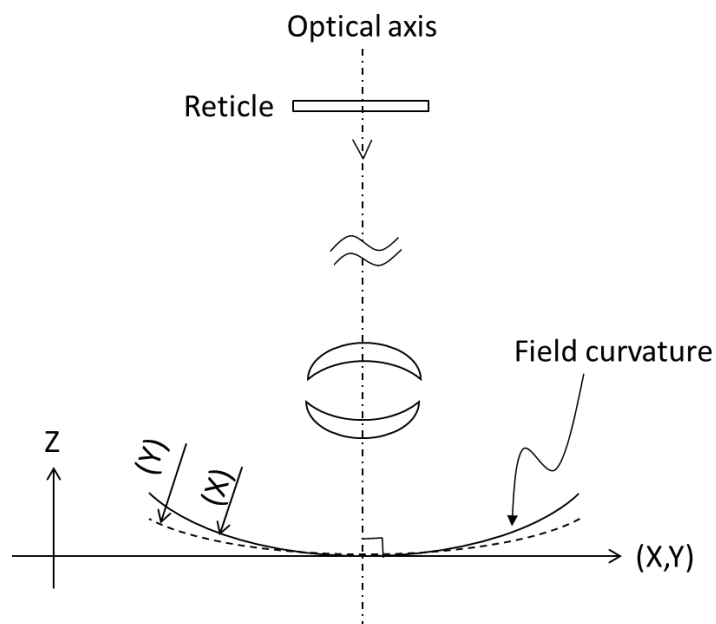


Figure 1.3.6: Field curvature aberration.

1.4 Aberration Measurement

The aberration measurement is very important for adjusting the lenses as shown in Fig. 1.4.1, where it represents the cycle of the lenses assemble, adjustment, measuring and judgment. In the lenses industry, the lens adjustment error is close to the precision that is three to five times over than the measurement error; therefore, the more precise measurement makes more easy and precise adjustment, which is the reason why the higher precision of aberration measurement is needed.

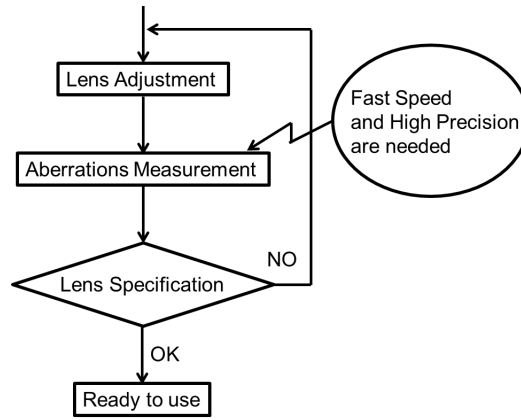


Figure 1.4.1: Lens adjustment flow.

1.4.1 Star Image Test

There are many lens aberration tests; the most common test is the star image test. Observation of a star image on the image plane is the fundamental lens test method for a long history [23, 24]. It was used for the lithography lens using a pinhole instead of a star. There are several pinhole coma aberration images ² as shown in Fig. 1.4.2 in which the aberration is clearly notified around $50\text{ m}\lambda$, however, the critical issue was the human eye damage problems by the short wavelength to observe images for a long period in mass production.

²The star test in this study is generated by the pinhole's far field image.

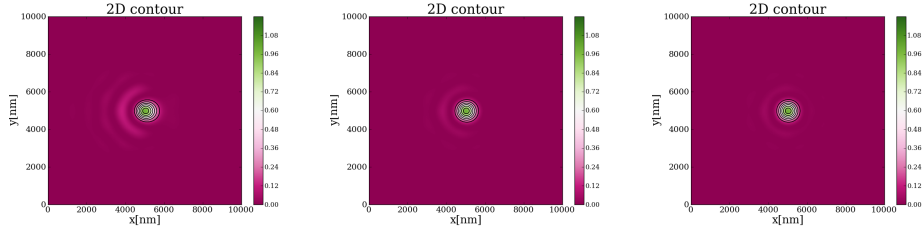


Figure 1.4.2: Star test with coma aberration; 250 mλ on the left, 100 mλ on the center, and 50 mλ on the right.

1.4.2 Interferometer Test

Figure 1.4.3 shows a Michelson Interferometer concept; the path difference between M_1 and M_2 is exactly compensated by G_2 [25]. Twyman replaced M_2 to the reference convex spherical mirror for measuring the concentric test lens which results in analyzing the aberrations by the interference fringe as shown in Fig. 1.4.4 [26–28].

Bruning firstly reported the laser source precise phase modulating Twyman and Green Interferometer that measured the lens aberrations as excellent results for the optical aligner [29]. It was the bench lens test, where the lens was away from the optical aligner.

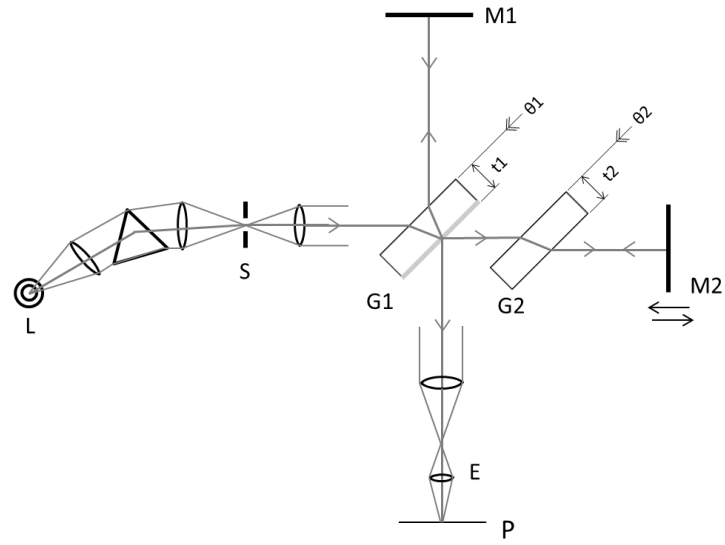


Figure 1.4.3: Michelson Interferometer.

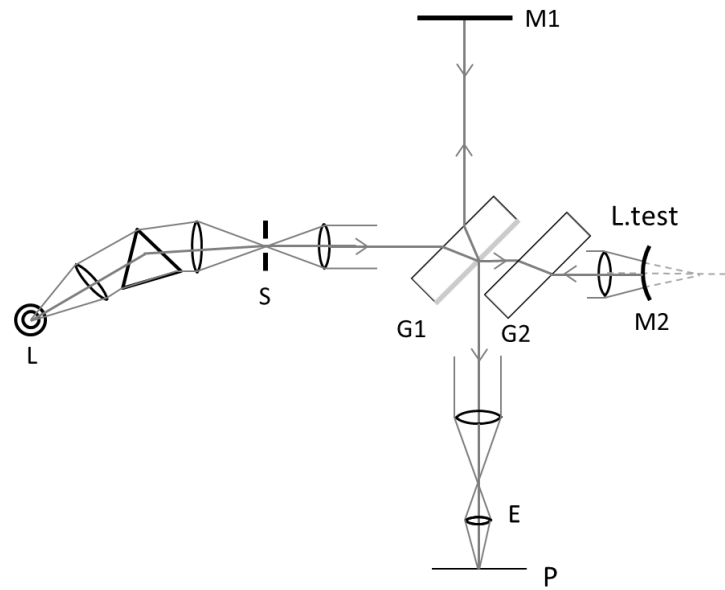


Figure 1.4.4: Twyman and Green Interferometer.

1.4.3 Aerial Image Test

To set the projection lens on the bench, the projected aerial image is scanned individual x and y axis on the image plane. One example is as shown in Fig. 1.4.5. It is useful to record the image to see the image spacial symmetry with regard to its center position and image contrast; however, it is not suitable to analyze each aberration.

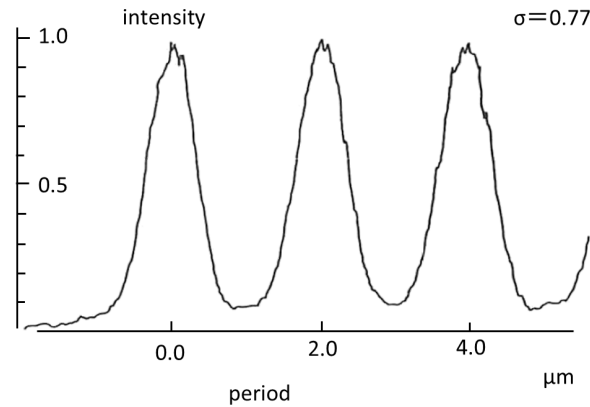


Figure 1.4.5: Aerial image with NSR 1010G optical aligner ($NA = 0.35$) [15].

1.5 Aberration Measurement by Photoresist

1.5.1 Scanning Electron Microscope Issues

The scanning electron microscope (SEM) has finer resolution than that of the optical microscope to measure the semiconductor linewidth; therefore, it has been only the standard instrument; however, the SEM has several difficulties in the field of use as follows;

1) In order to determine the best focus measurement, at least five to ten various top view linewidths at different focus position measurement data are needed, that means it takes time. Figure 1.5.1 shows the linewidth distribution data in the exposure field which requires many data.

2) The advanced photoresist maintains almost the same linewidth over a wide range of focus, especially when optimally exposed; therefore, even the

SEM has made it difficult to determine the best focus by the photoresist linewidth.

3) A wafer cannot be processed further due to electron beam induced damage in a photoresist. In addition, in case of destructive cross-section measurement the SEM method consumes a large amount of wafers.

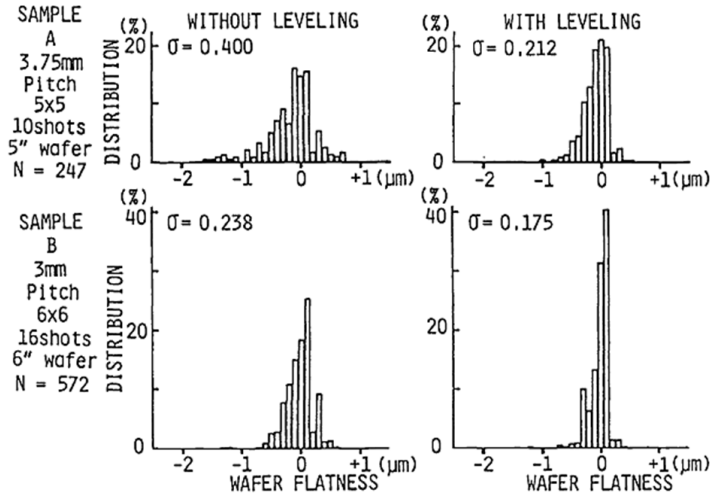


Figure 1.5.1: Linewidth distribution on an exposure wafer of a 15 mm x 15 mm lens field; the top 247 data are taken 3.75 mm x, y pitch and the bottom 572 data are taken also 3 mm x, y pitch by SEM measurement [30,31].

1.5.2 Non SEM Focus Measurement

In 1988, Nikon have innovated the field leveling function on the optical aligner which levels the curved wafer surface perpendicular to the axis of the projection lens by each exposure shot as shown in Fig. 1.5.2, [30,31]. The leveling can expand the usable depth of focus, however, in order to verify the leveling performance from exposed and developed wafer, it has been necessary to measure a lot of focus and field curvature data in the interfield; therefore, the development of new fast and precise focus measurement methodology is essential. Because, the SEM measurement involves a time-consuming manual procedure, for example, the data of cross section method is as shown in Fig.

1.5.3, and the data of CD/Focus method is as shown in Fig. 1.5.4.³

There are several former works which use a photoresist image to measure the aberration; one is star test on the resist [32], the other is using daggers on the resist, in which the wedge effect concept has already been developed by Murray named *Murray Daggers* that utilizes a narrow right triangle (dagger) marks to monitor the critical dimension on the wafer [33]; also there are pin-bars for astigmatism and field curvature on the resist image [34], and other unique methods [34–37].

Regarding the focus measurement, after invention of the phase shift mask technology [38, 39], one method uses a special phase shift mark to determine the best focus directly from the shift of image position [40, 41]. The advantage of this technique is that the focus measurement can be achieved with a single exposure; however, the illumination condition used for the measurement is not equivalent to that used for the actual device patterning.

Another approach utilizes light diffracted from the measurement pattern on the processed wafer [42–51]. The diffraction light is sensitive to tiny variations in the focus position when the dimensions of the measurement pattern are close to the resolution limit of the realistic manufacturing exposure process.

Despite this advantage, an estimation of the focus requires pattern analyzing software [42–44] or an iterative time-consuming electromagnetic calculation [45–48]. Moreover, special metrology tools are utilized, but it introduces additional costs and floor space requirements.

³This curve is called CD/Focus which means critical dimension focus position curve along defocus points.

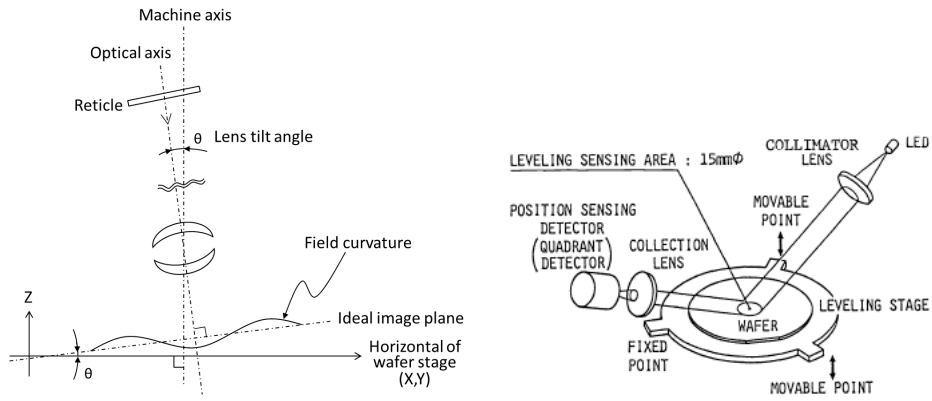


Figure 1.5.2: Left panel: Field leveling concept in which each exposure area on the wafer stage is continuously adjusted to the ideal image plane. Right panel: Field by field leveling function [30,31]; the LED (Light Emitted Diode) light is reflected by the wafer surface so that the angle is measured on the position sensor.

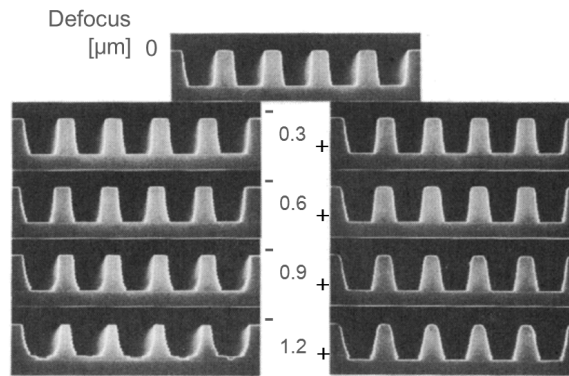


Figure 1.5.3: The 0.65 μm lines and spaces defocus cross section resist pattern; the angle of resist walls show the usable focus which is very accurate method, however, the cutting wafer for measurement is time consuming [30,31].

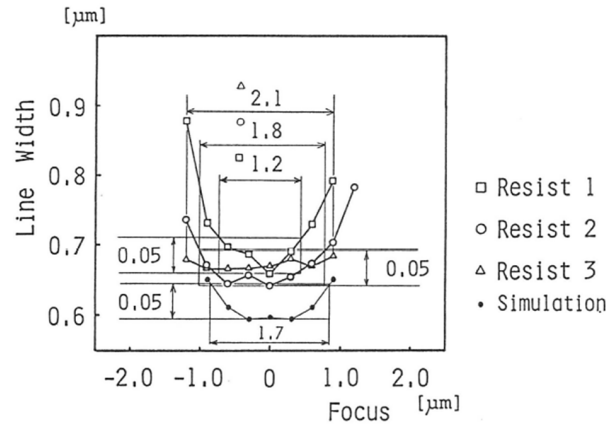


Figure 1.5.4: Linewidth measurement to determine the best focus along with the focus position by SEM metrology; however, there are different minimum linewidths among several types of photoresist [30, 31].

1.6 Coma Aberration Problems

1.6.1 Shift Coma Aberration

The coma aberration comes from not only the design residual aberration but also from the manufacturing or after manufacturing transportation process. Kubota described the lens lateral displacement [22] which means one lens shift from the optical axis of the lens group as shown in Fig. 1.6.1, using the references [19, 20].

The shift coma calculation induced by the lens lateral shift does not include the angle θ in Eq. (1.3.2); therefore, the lens shift coma amount is the same value from the image center to the peripheral image area. In the optomechanical work, how to mount lenses without shift is a very critical issue [52, 53].

This shift coma aberration gives the worst damage to lithography images; because it is caused by mechanical shift of which such amount is sometimes much larger than that of the lens design ray tracing results.

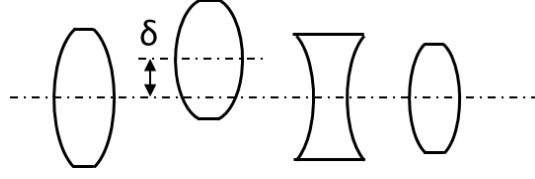


Figure 1.6.1: The cause of lens shift coma aberration; δ denotes the lateral shift of a lens in the lenses.

1.6.2 Coma Aberration Expression

The mathematical expression of the projection lens aberration has changed from the standard Seidel aberration formula, which originates from ray tracing, to the circular polynomial expansion, that is suitable for describing the wavefront propagation behavior from the light source to the image plane.

The aberration caused by decentering can be mathematically represented in the lowest-order coma aberration as the seventh and the eighth terms of the FRINGE Zernike polynomials [10, 54], Z_7 and Z_8 . It is essential to measure the amount of the decentering, or coma aberration, with high accuracy and to adjust the decentering accordingly.

Barakat and Houston developed the pioneering calculation for the positional shift or asymmetry of image profiles due to coma aberration [55–57]. This study quantitatively discussed the asymmetry of an image caused by coma aberration, wherein the target pattern was an isolated bright line in the dark field; however, the asymmetry caused by coma aberration for bright lines is not distinct, and the problem of accurate measurement remains unsolved. Therefore, a measurement method directly observing the asymmetry of aerial images remains undeveloped at that time.

1.6.3 Typical Coma Measurement by Photoresist

To measure the coma aberration, it is a well-known method to measure the linewidth of the most left and the most right lines of the group lines [58], and the subtraction between such two lines shows the coma aberration amount as shown in Fig. 1.6.2.

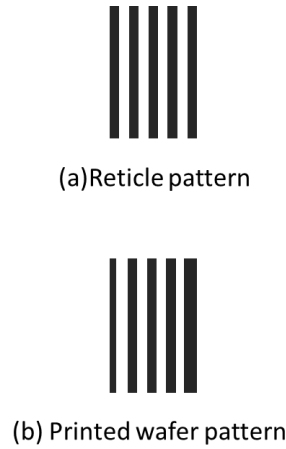


Figure 1.6.2: Linewidth abnormal between the right and the left side pattern with coma aberration; the top is the pattern on the reticle, and the bottom is the photoresist image through the projection lens that has the coma aberration, where the figures are conceptual drawings.

1.7 Distortion and Registration Measurement Problems

The Seidel aberration m_4 , the distortion, is defined and measured as the difference of the ideal full field grid and the actual grid on the x and y axes; where, both grids are exposed individually on the wafer as shown in Fig. 1.7.1.

The optical aligner aligns a wafer to a reticle for each exposure field. The wafer stage of the optical aligner moves field by field sequentially, and the optical aligner can expose a field with a series of alignment sequence steps until the entire wafer surface is exposed.

An optical registration metrology system (ORMS) measures the amount of exposure displacement error as shown in Fig. 1.7.2. ORMS measurement is possible during the alignment process using wafer alignment optics installed in the optical aligner. Occasionally, this measurement is done with an independent ORMS, which is a commercially available tool. Hereafter, we define ORMS to measure the photoresist marks on bare silicon wafers only. The ORMS will then be used mainly for analysis of machine performance, which is very useful for optical aligners.

The precision of ORMS sensors has two limitations, where the definition of precision is referred in these reference [59–62]. The first limitation relates to one dimensional laser spot sensors [63]. When a wafer must be moved beneath a fixed optical sensor [63], a scanning stage for providing object movement is required for one dimensional scanning optical sensor.

In usual case, both distortion and registration measurement are performed by the same method as the registration measurement which has the residual measurement error; where, the error in position of the laser metrology system for stage monitoring is generally larger than that of the optical sensor due to air temperature fluctuations and high frequency vibrations of machine components [64], which are beyond the frequency response of laser metrology, so the measurement precision can never be small, using conventional metrological methods.

The second limitation occurs when a two dimensional imaging sensor is used to measure a motionless alignment mark on the wafer [65–71]. It is difficult to improve measurement precision in this imaging mode due to the limits associated with optical image magnification and sensor pixel size reduction. Other types of difficulties also arise in case of ensuring adequate intensity to the solid state imaging pixels.

In a conventional approach, reducing the measurement error is to move the wafer stage very slowly and to take a large number of data points, N . However, the reduction factor of the error is proportional to $1/\sqrt{N}$, and so increasing N is time consuming.

An alternative approach to overlay measurements employs scatterometry [72], which is used to reconstruct the $3D$ shape of a registration mark from diffracted light. Although this method has the potential to meet the requirements for registration metrology, lengthy electromagnetic field calculations are required.

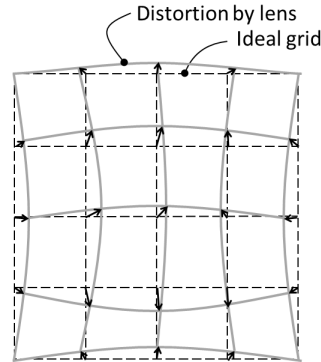


Figure 1.7.1: Schematics of distortion effects induced by imaging lens. The geometrical differences between the ideal grid and distorted grid are shown by vectors originated on the ideal grid points. The geometrical differences between the ideal grid and distorted grid are shown by vectors originated on the ideal grid points.

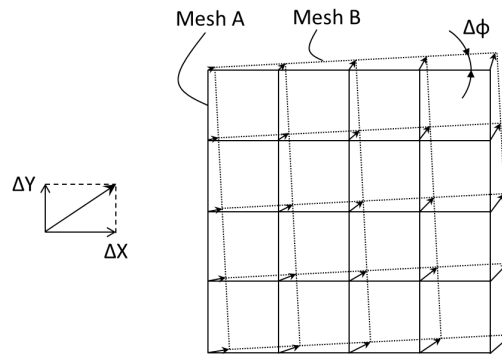


Figure 1.7.2: Definitions of coordinate system for registration metrology. There are two meshes: Mesh A for the first exposure and Mesh B for the second exposure. The difference between the two meshes are characterized by the vectors originated on each mesh point $(\Delta x, \Delta y)$ and the global inclination angle $\Delta\phi$.

1.8 Optical Image Amplification

1.8.1 Study Motivation

I denote my study motivation based on several explanations of this introduction; Table 1.8.1 summarizes the current metrological techniques and the associated problems which are described as follows;

Regarding the coma aberration measurement on the optical aligner, the SEM is considered the only tool for precise metrology; however, the SEM measurement is limited to around $50\text{ m}\lambda$ which can be a major limitation. Lens interferometry tests such as the Twyman and Green test which has around $5\text{m}\lambda$ precision cannot be installed in the actual optical aligner; therefore, a new metrological technique was required.

Regarding the focus measurement that means the field curvature measurement, the SEM is the only precise metrology; however, the SEM process requires at least 3 min per point measurement; therefore, a new faster optical metrology technique with precision equal to or better than that of the SEM was required.

Regarding the linewidth measurement, particularly for the linewidth variation of the multi-point interfield data, the SEM process again requires at least 3 min per point measurement. Consequently, the requirement for optical metrology technique with precision equal to or better than that of the SEM with measurement time of 1 to 5 s was required.

Regarding the registration and the distortion measurement, the measuring time of optical metrology was enough; however, the measurement precision was less than that required by a factor of 10. Consequently, the requirement for optical metrologies with higher precision was required.

Table 1.8.1: Motivation to be solved

My aberration measurement motivation underlying from 1980s to 1990s	
Seidel aberration	Metrology problems
Coma (m_2)	SEM resolution
Field curvature (m_5)	SEM lengthy duration
Linewidth (m_1, m_2, m_3, m_5)	SEM lengthy duration
Distortion (m_4)	Measurement precision
Best focus (m_1, m_2, m_3, m_5)	SEM lengthy duration
Registration	Measurement precision

1.8.2 Optical Image Amplification of Aberration Measurement

My study on technologies was consequently very focusing to amplify the optical aberration effects before the measurement, hereafter, the term amplification is used in this thesis.

Whatever type of aberrations, an optical image amplification in my context implies either the image intensity amplification or the image position amplification of the optical image as illustrated in Fig. 1.8.1.

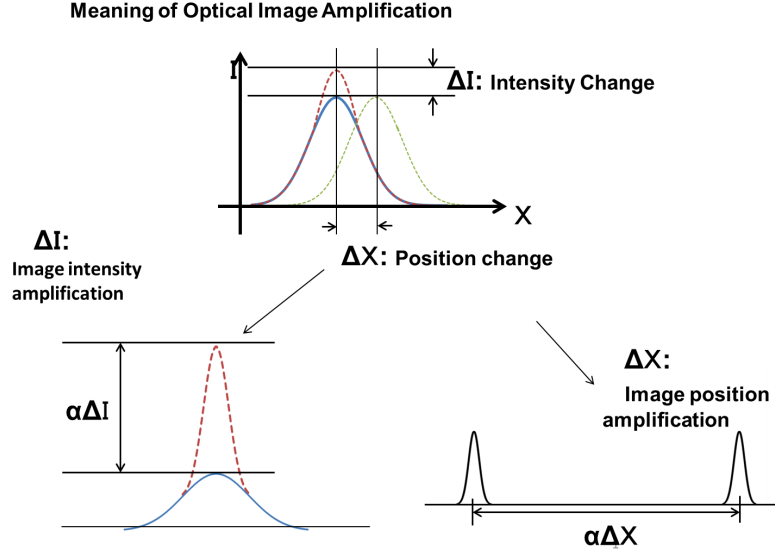


Figure 1.8.1: Two different definitions of optical image amplification, where $\alpha \Delta I$ denotes the image intensity amplification, and $\alpha \Delta X$ denotes the image position amplification, where α denotes the amplification factor.

At first; in Chapter 2, in order to amplify the image intensity which is a new concept of the aberration measurement, the principle of the heterodyne method is used as the reference concept. The expression for the heterodyne of two waves is given as

$$\sin \theta \sin \phi = \frac{1}{2}(\cos(\theta - \phi) - \cos(\theta + \phi)), \quad (1.8.1)$$

where θ is the carrier of the physical quantity, and ϕ is the signal of the physical quantity individually. When θ denotes the carrier frequency of light or alternating current, θ can be expressed as $\theta = \omega t$, and thus, we derive the following equation:

$$\sin(\omega t) \sin \phi = \frac{1}{2}(\cos(\omega t - \phi) - \cos(\omega t + \phi)). \quad (1.8.2)$$

In case of the heterodyne operation that involves trigonometric functions, the $\sin \phi$ component of the left term in the equation is the signal term which is observed independent of the carrier frequency; therefore, the key aspects of heterodyne involve the subtraction operation and the utilization of the

signal component $-\phi$ in the first term and $+\phi$ in the second term.

In the study of the image intensity amplification for coma aberration measurement, the amplitude $U(0)$ is the carrier amplitude, and the $U(1)$ is the signal amplitude in Eq. (2.1.10) which does not involve the use of trigonometric functions; however, the approach involves subtraction formality between $+U(1)$ in the first term and the $-U(1)$ in the second term in Eq. (2.1.10); this aspect exhibits similarity with the heterodyne approach.

At second; in Chapters 3, 4, 5, and 6, the image position amplification method involves three types of measurement, as shown in Fig. 1.8.2. All three measurements utilize the wedge effect that facilitates image position amplification.

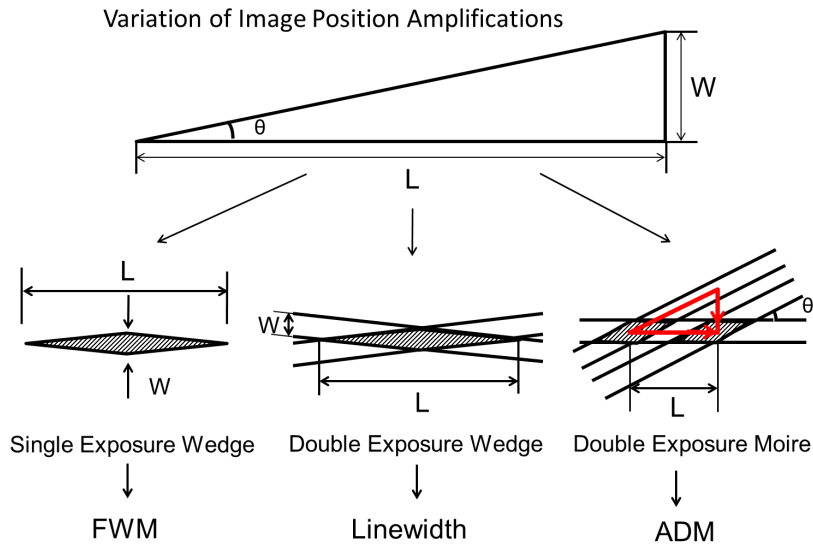


Figure 1.8.2: Three types of image position amplification; the focus wedge mark (FWM in Chapter 4), the linewidth measurement (in Chapter 5), and the displacement measurement (ADM in chapter 6), where the inclination angle θ is used for all three studies.

1.8.3 Study Target

I denote the study target at Table 1.8.2, in which a bar-in-bar is the displacement measurement method defined as shown in Fig. 6.2.1.

Table 1.8.2: Study target

Metrology target specifications for my study		
Derived metrology: type of amplification	precision	measuring time
(m_2) Sidelobe amplification : ΔI	x10 greater than SEM	less than SEM
(m_5) Focus wedge mark: ΔX	better than SEM	1 s to 5 s
(m_1, m_2, m_3, m_5) Linewidth: ΔX	equal to SEM	1 s to 5 s
Best focus (Focus wedge mark)	better than SEM	1 s to 5 s
(m_4) Alt. Dir. Moire: ΔX	x30 bar-in-bar	equal to bar-in-bar

1.9 Overview of Chapters

The following chapters describe the optical image and position amplification in detail; firstly, we start the optical image amplification measurement of aerial image in Chapter 2. In Chapters 4, 5, and 6, the image position amplification measurement is discussed.

The primary chapters of this thesis consist of four chapters;
 In Chapter 2: theoretical analysis of coma aberration measurement for aerial image sidelobe peaks [73].
 In Chapter 4: theoretical analysis of focus wedge mark for best-focus determination in photolithography [74].
 In Chapter 5: technique for optical characterization of exposure tool imaging performance down to 100 nm [75].
 In Chapter 6: optical displacement metrology using alternating direction moire [76].

Chapter 2

Coma Aberration Measurement

2.1 Coma Intensity Amplification

2.1.1 Search for Coma Sensitive Image

In order to determine the most coma sensitive aerial image test pattern, I have developed a measurement method of the intensity profile of aerial image which utilizes a photoresist top layer as the optical sensor; where, the photoresist has the exposure gate E_H as shown in Fig. 2.1.1, in which, the gate defines the threshold for incoming exposure energy.

The methodology is as follows:

If the exposure energy would continuously increase above and around the threshold of the gate, the surface of the photoresist begins to peel off at E_H , as shown in Fig. 2.1.1. The optical microscope observes this peeling area as the dark image with the interference at the resist thickness of $\lambda/4$ which is made constantly a little larger energy than E_H as shown in Fig. 2.1.2; therefore, it is the aerial image peak intensity measurement method which is derived from the resist peeling concept of E_H [77, 78].

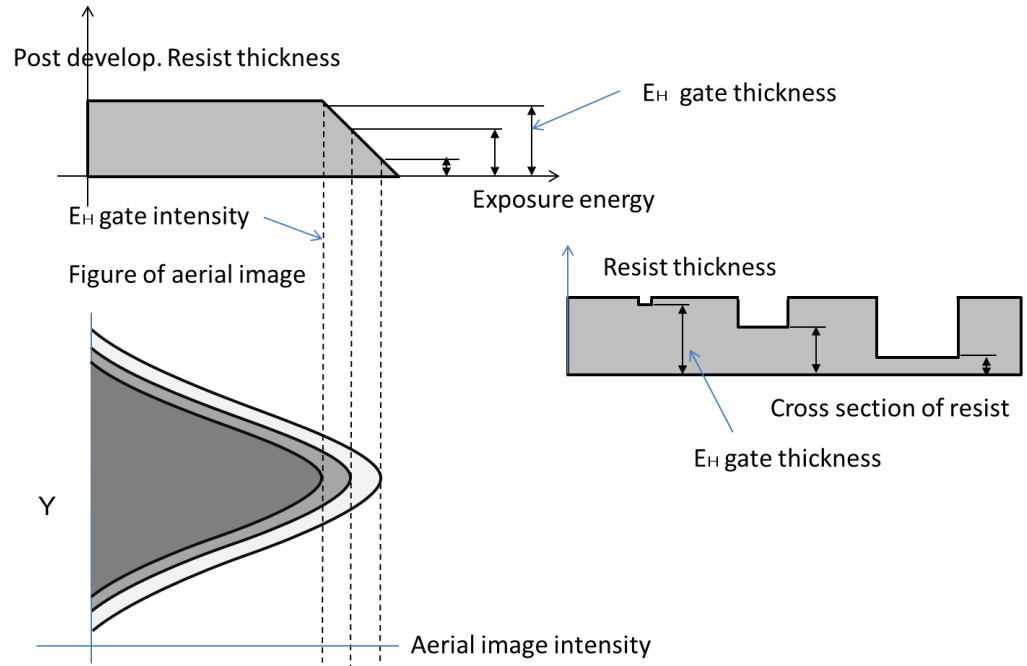


Figure 2.1.1: Aerial image peak intensity measurement method. Top panel: Post developed resist thickness is related to the exposure energy [77]. Bottom panel: Increasing the peak aerial image intensity corresponds to the resist thickness. Right panel: most small resist peeling phenomenon shows the peak image intensity defined by E_H gate.

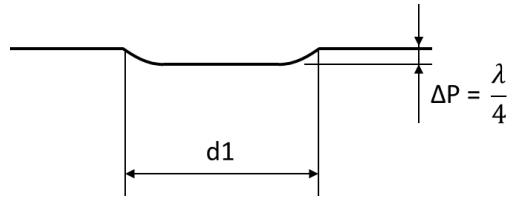


Figure 2.1.2: Optical inspection.

Using this method, I searched several sensitive patterns for aberration of coma under different illumination conditions, in which I could determine that the lines and spaces pattern (hereafter called the L&S) with bright-background is the most sensitive for the coma aberration [79]. Figure 2.1.3

shows the detailed simulation results of negative and positive patterns containing coma aberration. The followings are features of two types of patterns:

- 1) The side peak of the negative pattern is considerably larger than that of the positive pattern; therefore, this feature is used for coma measurement of the aerial image and the distinguished features of negative-pattern.

- 2) The linewidth difference for the coma aberration measurement is more precise by a factor of 40 with respect to the sidelobe aerial image as shown in the figure; this is why the SEM coma measurement is often low-sensitivity measurement.

- 3) The difference between the left side peak intensity and the right side peak intensity is proportional to the coma aberration amount which is calculated as shown in Fig. 2.1.6.

- 4) The coma aberration causes the shift of the image position.

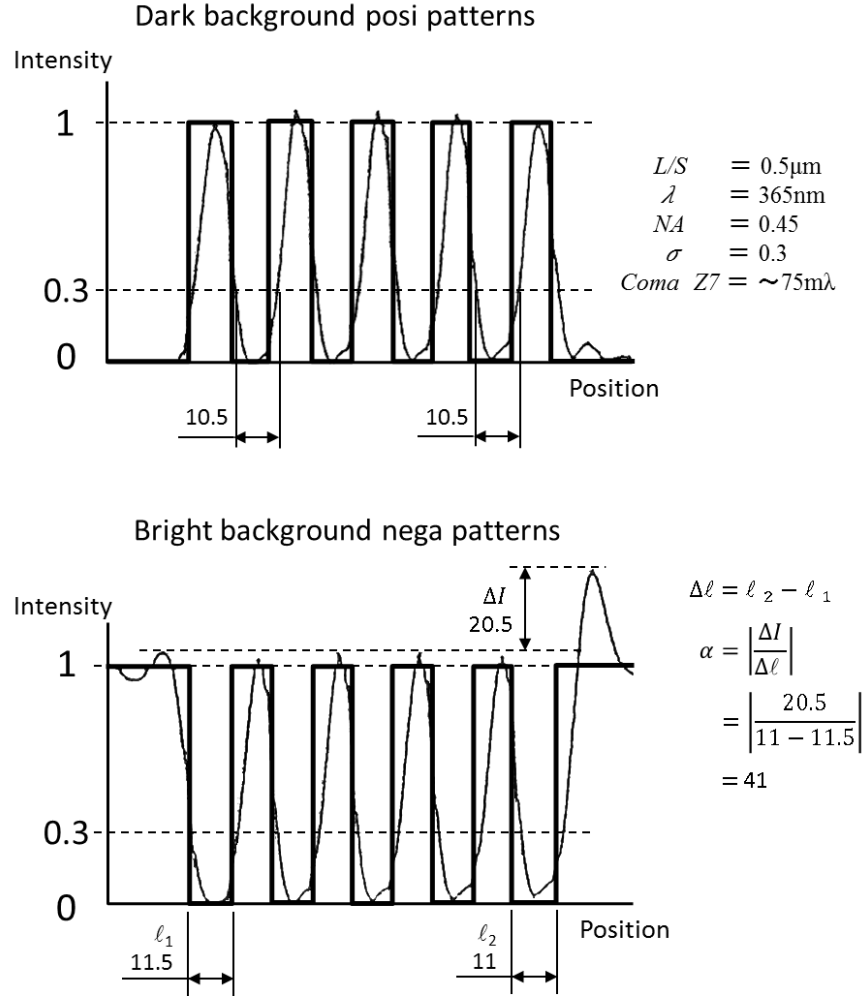


Figure 2.1.3: Aerial bars image under coma aberration; the top is dark background and the bottom is bright background.

2.1.2 Intensity Amplification Model

We are interested in the amplification of the background light under coherent illumination. The light amplitudes compose both the imaging amplitude with coma aberration and the background amplitude, as shown in Fig. 2.1.4, are expressed as follows:

Bright-background case, where there is $g(x)$:

$$I_B = (f(a) + g(x))^2, \quad (2.1.1)$$

$$= f(a)^2 + 2f(a)g(x) + g(x)^2. \quad (2.1.2)$$

Dark-background case, where there is no $g(x)$:

$$I_D = f(a)^2 = f(a)^2. \quad (2.1.3)$$

In addition, in case of fully incoherent illumination for bright-background amplitude:

$$I_{Inc.} = f(a)^2 + g(x)^2, \quad (2.1.4)$$

where $f(a)$ represents the sidelobe coma peak amplitude, and $g(x)$ represents the background constant amplitude. In the above cases, the bright-background image I_B exhibits the amplified coma sidelobe peak intensity as the 2-nd term in Eq. (2.1.2); that is what I looked for an optical image amplification effect in aberration measurement.

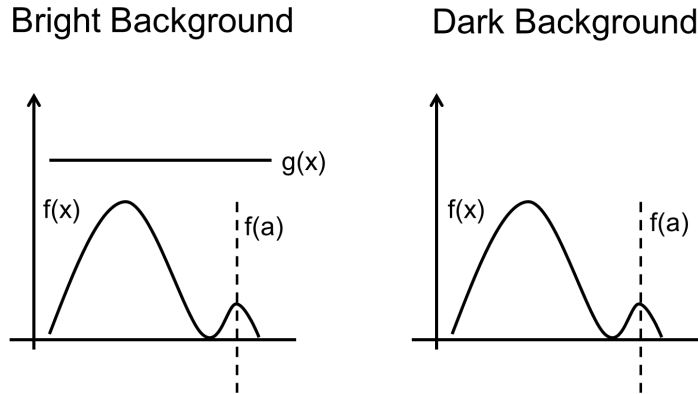


Figure 2.1.4: Simple amplitude amplification model.

In order to understand briefly the image intensity amplification by Fourier optics using “complex function”, we have derived the image amplitudes in Eqs. (2.1.5) and (2.1.6);

$$U(x_L) = U_0(x_L) + U_1(x_L) \quad (2.1.5)$$

$$U(x_R) = U_0(x_R) + U_1(x_R) \quad (2.1.6)$$

where those are composed of a fundamental amplitude U_0 (an even function) and the coma influenced amplitude U_1 (an odd function) as shown in Fig. 2.1.5. The fundamental amplitude is the even function:

$$U_0(x_L) = U_0(x_R) \quad (2.1.7)$$

The coma influenced function is the odd function:

$$U_1(x_L) = -U_1(x_R) \quad (2.1.8)$$

where the first sidelobe peak position:

$$x_L = -x_R \quad (2.1.9)$$

When the image intensity along the x axis is defined as $I(x)$, the intensity difference of the sidelobe peaks $\Delta I(x)$ is calculated as

$$\begin{aligned} \Delta I &= I(x_R) - I(x_L) \\ &= |U_0(x_R) + U_1(x_R)|^2 - |U_0(x_L) + U_1(x_L)|^2 \\ &= |U_0(x_R) + U_1(x_R)|^2 - |U_0(x_R) - U_1(x_R)|^2 \\ &= 4U_0U_1 \end{aligned} \quad (2.1.10)$$

where x_L and x_R denote the positions of the right and left sidelobe peaks, respectively.

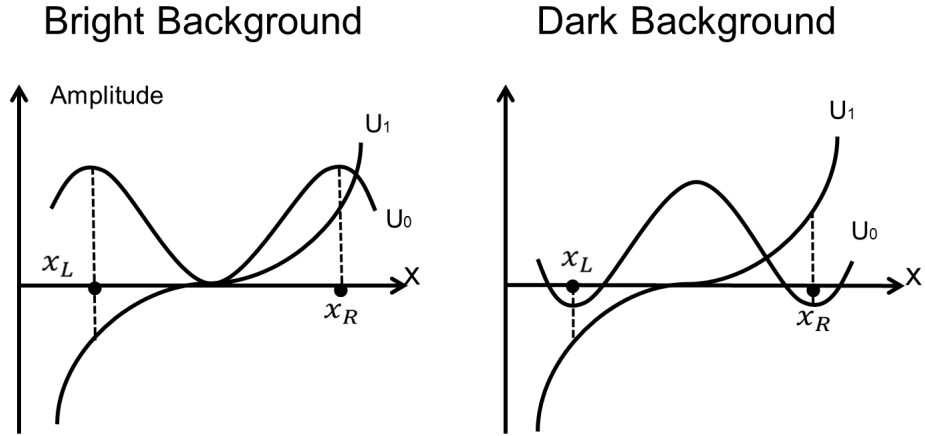


Figure 2.1.5: Image amplitude amplification model.

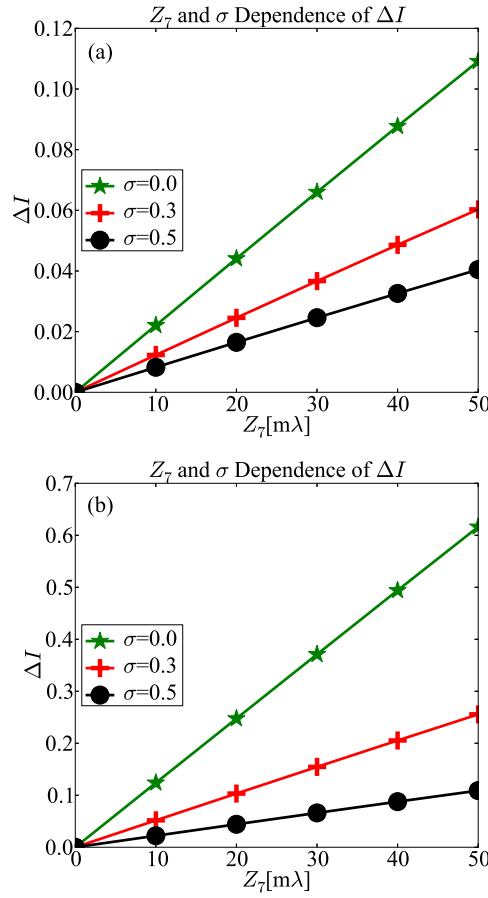


Figure 2.1.6: The intensity difference of the sidelobe defined in Eq. (2.1.10), plotted with respect to Z_7 and various σ values by the optical simulation software : (a) positive pattern (dark background) and (b) negative pattern (bright background) [74].

We have calculated $\Delta I(x)$ as the function of coma aberration amount by optical simulation as shown in Fig. 2.1.6, for both the positive and negative patterns for large illumination σ region.

It is very important that the intensity difference $\Delta I(x)$ is a linear function of the coma aberration amount. This is because the linearity is the key for any measurement technology.

2.2 Coherent Imaging Analysis

2.2.1 Coherent Imaging Formula for Aberration

In this section, we analyze precisely the effect of the lowest order coma aberration on the intensity profile by expanding the pupil function regarding the coma aberration into the second order of FRINGE Zernike polynomials.

The wavefront aberration $W(\xi, \eta)$ is written as

$$W(\xi, \eta) = Z_7 [3\xi^3 + (3\eta^2 - 2)\xi], \quad (2.2.1)$$

where the Zernike coefficient Z_7 indicates the residual amount of the lowest-order coma aberration. For the coma aberration, the odd functional relation $W(-\xi, \eta) = -W(\xi, \eta)$ holds. The optical pupil function $G(\xi, \eta)$ with wavefront aberration $W(\xi, \eta)$ can be represented as

$$G(\xi, \eta) = \exp[ikW(\xi, \eta)]. \quad (2.2.2)$$

The following discussion is also valid when the wavefront aberration consists of other Zernike coefficients, as described in Section 2.5. The intensity profile of the image under coherent illumination can be written as

$$\begin{aligned} I(x) &= \left| \int_{-\text{NA}}^{\text{NA}} d\xi \tilde{T}(\xi) G(\xi, 0) e^{ik\xi x} \right|^2 \\ &= \left| \int_{-\text{NA}}^{\text{NA}} d\xi \tilde{T}(\xi) e^{ikZ_7(3\xi^3)} e^{ik(x-2Z_7)\xi} \right|^2, \end{aligned} \quad (2.2.3)$$

where the magnification ratio is assumed to be unity. In Eq. (2.2.3), the second term of Eq. (2.2.1) corresponds to the uniform coordinate translation on the x -axis: $x \rightarrow x - 2Z_7$. Therefore, we calculate Eq. (2.2.3) after the coordinate translation. We show the coordinate diagram of the coherent image formula as shown in Fig. 1.2.1, where the wave is transmitted from the object plane to the image plane.

The pupil function expanded into a power series of the coma aberration up to the second order, hereafter called the “power series,” is given by

$$\begin{aligned} G(\xi, 0) &= \sum_{m=0}^{\infty} \frac{(ik)^m}{m!} \{(3\xi^3) Z_7\}^m \\ &\simeq 1 + ik(3\xi^3) Z_7 - \frac{k^2}{2} [(3\xi^3) Z_7]^2. \end{aligned} \quad (2.2.4)$$

Substituting Eq. (2.2.4) for Eq. (2.2.3), we derive the image intensity in terms of the power series of the coma aberration as

$$\begin{aligned} I(x) &\simeq |U'(x)|^2 \\ &\equiv |U_0(x) + U_1(x)Z_7 + U_2(x)(Z_7)^2|^2, \end{aligned} \quad (2.2.5)$$

where the amplitudes are given by

$$U'(x) = U_0(x) + U_1(x)Z_7 + U_2(x)(Z_7)^2, \quad (2.2.6)$$

$$U_0(x) = \int_{-NA}^{NA} d\xi \tilde{T}(\xi) e^{ik\xi x}, \quad (2.2.7)$$

$$U_1(x) = ik \int_{-NA}^{NA} d\xi \tilde{T}(\xi) (3\xi^3) e^{ik\xi x}, \quad (2.2.8)$$

$$U_2(x) = -\frac{k^2}{2} \int_{-NA}^{NA} d\xi \tilde{T}(\xi) (3\xi^3)^2 e^{ik\xi x}. \quad (2.2.9)$$

The amplitudes in Eqs. (2.2.7), (2.2.8), and (2.2.9) are guaranteed to be real functions, which are easily calculated in this theme than complex functions. The following equations represent the proofs of real function by the nature of the one-dimensional pattern symmetry for x axis and the oddness of the coma aberration for ξ axis:

$$\begin{aligned} U_0^*(x) &= \int_{-NA}^{NA} d\xi \tilde{T}^*(\xi) e^{-ik\xi x} \\ &= \int_{NA}^{-NA} d(-\xi) \tilde{T}(-\xi) e^{ik\xi x} \\ &= \int_{-NA}^{NA} d\xi \tilde{T}(\xi) e^{ik\xi x} \\ &= U_0(x), \end{aligned} \quad (2.2.10)$$

$$\begin{aligned} U_1^*(x) &= -ik \int_{-NA}^{NA} d\xi (3\xi^3) \tilde{T}(\xi) e^{-ik\xi x} \\ &= -ik \int_{+NA}^{-NA} d(-\xi) (-3\xi^3) \tilde{T}(-\xi) e^{ik\xi x} \\ &= U_1(x), \end{aligned} \quad (2.2.11)$$

$$\begin{aligned}
U_2^*(x) &= -\frac{k^2}{2} \int_{-NA}^{NA} d\xi (3\xi^3)^2 \tilde{T}(\xi) e^{-ik\xi x} \\
&= -\frac{k^2}{2} \int_{NA}^{-NA} d(-\xi) (-3\xi^3)^2 \tilde{T}(-\xi) e^{ik\xi x} \\
&= U_2(x).
\end{aligned} \tag{2.2.12}$$

The amplitudes $U_0(x)$ and $U_2(x)$ are verified to be even functions of x :

$$\begin{aligned}
U_0(x) &= \int_{-NA}^{NA} d\xi \tilde{T}(\xi) e^{ik\xi x} \\
&= \int_{NA}^{-NA} d(-\xi) \tilde{T}(-\xi) e^{-ik\xi x} \\
&= \int_{-NA}^{NA} d\xi \tilde{T}(\xi) e^{-ik\xi x} \\
&= U_0(-x),
\end{aligned} \tag{2.2.13}$$

$$\begin{aligned}
U_2(x) &= -\frac{k^2}{2} \int_{-NA}^{NA} d\xi (3\xi^3)^2 \tilde{T}(\xi) e^{ik\xi x} \\
&= -\frac{k^2}{2} \int_{NA}^{-NA} d(-\xi) (-3\xi^3)^2 \tilde{T}(-\xi) e^{-ik\xi x} \\
&= U_2(-x).
\end{aligned} \tag{2.2.14}$$

On the other hand, the amplitude $U_1(x)$ is an odd function of x :

$$\begin{aligned}
U_1(x) &= ik \int_{-NA}^{NA} d\xi (3\xi^3) \tilde{T}(\xi) e^{ik\xi x} \\
&= ik \int_{+NA}^{-NA} d(-\xi) (-3\xi^3) \tilde{T}(-\xi) e^{-ik\xi x} \\
&= -U_1(-x).
\end{aligned} \tag{2.2.15}$$

2.2.2 Sidelobe Peak Positions

The position of the sidelobe peak is shifted by the coma aberration. The positions of local peaks of the aerial image should be determined from the requirement that the first-order positional differentiation of the intensity profile must be zero. The following mathematical discussion can be applied to

any pair of peaks observed in the intensity profile, including the sidelobe peaks.

If the optics are aberration-free, the intensity profile $|U_0(x)|^2$ becomes an even function according to Eq. (2.2.13). Therefore, any pair of local peaks is observed in positions $\pm x_0$ that are equally distanced from the origin ($x = 0$). The positions $\pm x_0$ are defined as solutions of the equation:

$$\nabla U_0(x)|_{x=\pm x_0} = 0, \quad (2.2.16)$$

where the symbol ∇ is the derivative in one direction. The shift ζ_1 , caused by “the first-order aberration amplitude” $U_1(x)$ in Eq. (2.2.8), is calculated as the first derivative of the intensity profile that includes the zero- and first-order power series from Eq. (2.2.5), $|U_0(x) + U_1(x)|^2$;

$$0 = 2 \{U_0(x) + U_1(x) Z_7\} \times \nabla \{U_0(x) + U_1(x) Z_7\} |_{x=\pm x_0+\zeta_1}. \quad (2.2.17)$$

The first bracket in Eq. (2.2.17) is the amplitude at the peak position influenced by the first-order coma aberration. In general, there is a possibility that the sidelobe peak disappears when a large coma aberration is added. Our aberration measurement method inherently utilizes the sidelobe peak intensity detected under coma aberration, so the disappearance of the sidelobe peak is outside the scope of this method; therefore, the amplitude in the first bracket is finite, while the gradient of the amplitude in the second bracket should be zero. The peak position shift ζ_1 is written as

$$\zeta_1 = - \frac{\nabla U_1(x_0)}{\nabla^2 U_0(x_0)} Z_7, \quad (2.2.18)$$

where the detailed calculations are described in Appendix of [73].

Similarly, the peak position shift ζ_2 , caused by “the second-order aberration amplitude” $U_2(x)$ given by Eq. (2.2.9), is calculated as the first-order derivative of the intensity profile in Eq. (2.2.5):

$$0 = 2U'(x) \nabla U'(x)|_{x=\pm x_0+\zeta_1+\zeta_2}. \quad (2.2.19)$$

We rigorously calculate Eq. (2.2.19) within the second-order aberration:

$$\begin{aligned} \zeta_2 = \pm \frac{-1}{\nabla^2 U_0(x_0)} & \left[\frac{1}{2} \nabla^3 U_0(x_0) (\zeta_1)^2 \right. \\ & \left. + \nabla^2 U_1(x_0) \zeta_1 Z_7 + \nabla U_2(x_0) (Z_7)^2 \right], \end{aligned} \quad (2.2.20)$$

where the detailed calculations are given in Appendix of [73].

Finally, we derive the left sidelobe peak position x_L and right sidelobe peak position x_R including the influence up to the second-order aberration:

$$x_L = -x_0 + \zeta_1 - \zeta_2, \quad (2.2.21)$$

$$x_R = x_0 + \zeta_1 + \zeta_2. \quad (2.2.22)$$

2.2.3 Intensity Difference between Pair of Local Peaks

The calculated positional shifts of the peaks, given in Eqs. (2.2.21) and (2.2.22) are substituted for Eq. (2.2.5). Then, the intensity difference $\Delta I(x)$ in Eq. (2.1.10) can be calculated by subtracting the local peak intensity of left side from that of right side:

$$\begin{aligned} \Delta I &\simeq |U'(x^R)|^2 - |U'(x^L)|^2 \\ &\simeq \left| U_0(x_0) + \nabla U_0(x_0)(\zeta_1 + \zeta_2) \right. \\ &\quad \left. + \frac{1}{2} \nabla^2 U_0(x_0)(\zeta_1 + \zeta_2)^2 + U_1(x_0) Z_7 \right. \\ &\quad \left. + \nabla U_1(x_0)(\zeta_1 + \zeta_2) Z_7 + U_2(x_0)(Z_7)^2 \right|^2 \\ &\quad - \left| U_0(-x_0) + \nabla U_0(-x_0)(\zeta_1 - \zeta_2) \right. \\ &\quad \left. + \frac{1}{2} \nabla^2 U_0(-x_0)(\zeta_1 - \zeta_2)^2 + U_1(-x_0) Z_7 \right. \\ &\quad \left. + \nabla U_1(-x_0)(\zeta_1 - \zeta_2) + U_2(-x_0)(Z_7)^2 \right|^2 \\ &= |U_0(x_0) + \varepsilon + U_1(x_0) Z_7|^2 \\ &\quad - |U_0(x_0) + \varepsilon - U_1(x_0) Z_7|^2 \\ &= 2[U_0(x_0) + \varepsilon] \times 2U_1(x_0) Z_7 \\ &\simeq 4U_0(x_0) U_1(x_0) Z_7, \end{aligned} \quad (2.2.23)$$

where the contribution related to the second-order aberration ϵ is defined as follows:

$$\begin{aligned}
\varepsilon &\equiv \nabla U_0(x_0) (\zeta_1 + \zeta_2) \\
&\quad + \frac{1}{2} \nabla^2 U_0(x_0) (\zeta_1 + \zeta_2)^2 \\
&\quad + \nabla U_1(x_0) (\zeta_1 + \zeta_2) Z_7 + U_2(x_0) (Z_7)^2 \\
&\simeq \frac{1}{2} \nabla^2 U_0(x_0) [(\zeta_1)^2 + 2\zeta_1\zeta_2] \\
&\quad + \nabla U_1(x_0) \zeta_1 Z_7 + U_2(x_0) (Z_7)^2 \\
&= \left\{ -\frac{1}{2} \frac{[\nabla U_1(x_0)]^2}{\nabla^2 U_0(x_0)} + U_2(x_0) \right\} (Z_7)^2.
\end{aligned} \tag{2.2.24}$$

In this thesis, we ignore all terms higher than the third-order of coma Z_7 . The derivations of Eq. (2.2.23) are performed using the same technique as that of the ζ_1 and ζ_2 shifts in the reference [73]. Our method makes good use of the characteristics regarding the power series of coefficient Z_7 and the parity of Eqs. (2.2.13), (2.2.15) and (2.2.14).

From the above calculations, we see that as long as we treat the image intensity within the second-order aberration, the intensity difference ΔI given in Eq. (2.2.23) does not depend on the detail of ζ_2 in Eq. (2.2.20). The quadratic terms are canceled in the subtraction, and only the term linear to the aberration remains in the intensity difference ΔI . In Eq. (2.2.23), it is not necessary to calculate the positional shifts, ζ_1 and ζ_2 , caused by aberration.

We define the ‘‘aberration sensitivity,’’ which is a physical quantity to evaluate the magnitude of the intensity difference that can be measured under the lowest-order coma aberration Z_7 , expressed by

$$(\text{aberration sensitivity}) = \left. \frac{\Delta I}{Z_7} \right|_{Z_7 \rightarrow 0}. \tag{2.2.25}$$

In the case of coherent illumination, using Eq. (2.2.23), the aberration sensitivity is simply expressed as

$$(\text{aberration sensitivity}) = 4U_0(x_0)U_1(x_0). \tag{2.2.26}$$

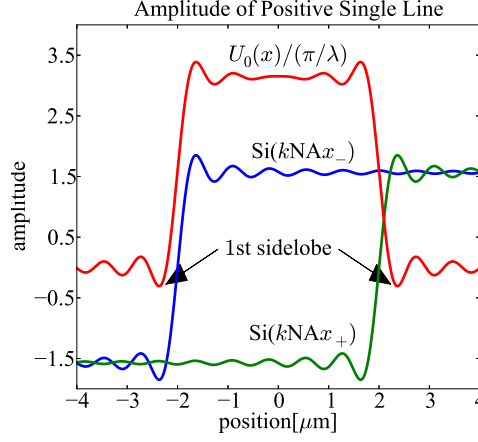


Figure 2.3.1: Positive single line in Eq. (2.3.3) for amplitude $U_0(x)/(\pi/\lambda)$, ringing waves $\text{Si}(k\text{NA}x_-)$ and $\text{Si}(k\text{NA}x_+)$, and $D = 4.0\mu\text{m}$. The calculation parameters are $\lambda = 0.365\mu\text{m}$ and $\text{NA} = 0.5$ [74].

2.3 Aberration for Single Line

2.3.1 Single Line Positive Pattern

In the case of a single line pattern, the intensity difference between the first sidelobe peaks can be calculated analytically using Eq. (2.2.23). Deriving the coma aberration sensitivity defined in Eq. (2.2.26) for both positive and negative single line patterns, we demonstrate that the aberration sensitivity of the negative line is much larger than that of the positive line pattern.

The transmittance $T(x)$ of the positive single line pattern with line width D and the spectrum $\tilde{T}(\xi)$ are given by

$$T(x) = \text{rect}\left[-\frac{D}{2}, \frac{D}{2}\right], \quad (2.3.1)$$

$$\tilde{T}(\xi) = D \text{sinc}\left(\frac{\pi D}{\lambda} \xi\right), \quad (2.3.2)$$

where the function $\text{sinc}(x) \equiv \sin(x)/x$. Substituting the spectrum in Eq. (2.3.2) for the amplitude $U_0(x)$ in Eq. (2.2.7), the aberration-free amplitude

is given as

$$\begin{aligned}
 U_0(x) &= \int_{-\text{NA}}^{\text{NA}} d\xi D \text{sinc} \left(\frac{\pi D}{\lambda} \xi \right) e^{ik\xi x} \\
 &= \frac{\lambda}{\pi} \int_0^{\text{NA}} d\xi \frac{1}{\xi} \left\{ \sin \left[k \left(x + \frac{D}{2} \right) \xi \right] \right. \\
 &\quad \left. - \sin \left[k \left(x - \frac{D}{2} \right) \xi \right] \right\} \\
 &= \frac{\lambda}{\pi} [\text{Si}(k\text{NA}x_-) - \text{Si}(k\text{NA}x_+)], \tag{2.3.3}
 \end{aligned}$$

where the integral sine function $\text{Si}(x)$ and the positions x_+ and x_- are defined as follows:

$$\text{Si}(x) = \int_0^x dt \frac{\sin t}{t}, \tag{2.3.4}$$

$$x_{\pm} = x \mp \frac{D}{2}. \tag{2.3.5}$$

The features of the integral sine function are shown in Chapter Appendix 2.7.1.

Fig. 2.3.1 shows that the aberration-free amplitude of the positive single line in Eq. (2.3.3) reproduces the Gibbs phenomenon at the boundaries of the line on the image plane, which reflects the nature of wave optics.

The two amplitudes from the different integral sine functions destructively interfere with each other to suppress the summed amplitude away from the boundary of the object, which corresponds to the characteristics of geometrical optics. On the other hand, around the origin, the two integral sine functions constructively interfere and emphasize the summed amplitude.

Similarly, the spectrum in Eq. (2.3.2) and the coma aberration in Eq. (2.2.1) are substituted for the amplitude $U_1(x)$ in Eq. (2.2.8). We obtain the following result:

$$\begin{aligned}
 U_1(x) &= ik \int_{-\text{NA}}^{\text{NA}} d\xi 3\xi^3 D \text{sinc} \left(\frac{\pi D}{\lambda} \xi \right) e^{ik\xi x} \\
 &= \frac{3k\lambda}{\pi} \left[\int_0^{k\text{NA}x_+} d\xi \frac{\xi^2 \cos \xi}{(kx_+)^3} - \int_0^{k\text{NA}x_-} d\xi \frac{\xi^2 \cos \xi}{(kx_-)^3} \right] \\
 &= \frac{3(\text{NA})^3 k\lambda}{\pi} [u_1(k\text{NA}x_-) - u_1(k\text{NA}x_+)], \tag{2.3.6}
 \end{aligned}$$

where the function $u_1(x)$ is defined as

$$u_1(x) = \frac{2x \cos(x) + (x^2 - 2) \sin(x)}{x^3}. \quad (2.3.7)$$

In both Eq. (2.3.3) and Eq. (2.3.6), the first and second terms represent the “ringing wave” from the left side boundary ($x = -D/2$) and from the right side boundary ($x = D/2$), respectively. The ringing wave originates from the virtual object centered at the left (right) boundary of the object $D/2$ ($-D/2$), which has a $+1/2$ transparent coefficient on the right side and a $-1/2$ transparent coefficient on the left side, respectively.

In estimating the first sidelobe peak for one side, we introduce an approximation that neglects the ringing wave from the boundary of the opposite side. Therefore, in deriving the position of the first sidelobe peak on the right side, the calculation is approximated as follows:

$$\begin{aligned} 0 &= \nabla |U_0(x)|^2 = 2U_0(x) \nabla U_0(x) \propto \nabla U_0(x) \\ &\propto \nabla [\text{Si}(k\text{NA}x_-) - \text{Si}(k\text{NA}x_+)] \\ &\simeq \nabla [\text{Si}(\infty) - \text{Si}(k\text{NA}x_+)] \\ &= -\nabla \text{Si}(k\text{NA}x_+) = -\frac{\sin(k\text{NA}x_+)}{k\text{NA}x_+}. \end{aligned} \quad (2.3.8)$$

The position of the first sidelobe peak x_0 is given by

$$k\text{NA}x_+|_{x=x_0} = \pi, \quad (2.3.9)$$

$$\therefore x_0 = \frac{D}{2} + \frac{\pi}{k\text{NA}}. \quad (2.3.10)$$

In the above approximation, we assume the following condition:

$$k\text{NAD} = \frac{2\pi}{\lambda} \cdot \text{NA} \cdot D \gg 1, \quad (2.3.11)$$

where the ringing wave from one side is assumed to be negligible so that it can be treated as a constant when it reaches the opposite side. With the given wavelength λ and numerical aperture NA, this condition can be met if an adequately large line width D is applied.

The intensity difference ΔI_{posi} in Eq. (2.2.23) is obtained using Eqs.

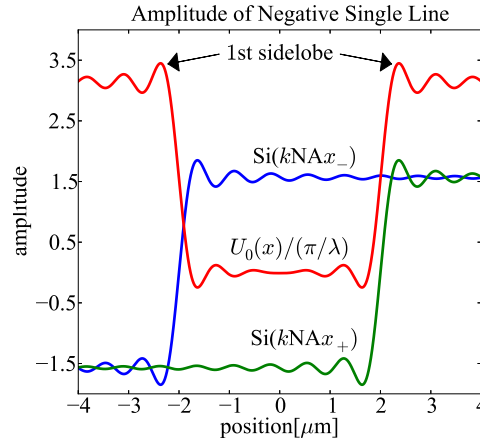


Figure 2.3.2: Negative single line in Eq. (2.3.15) for amplitude $U_0(x)/(\pi/\lambda)$, ringing waves $\text{Si}(k\text{NA}x_-)$ and $\text{Si}(k\text{NA}x_+)$, and $D = 4.0\mu m$. The calculation parameters are $\lambda = 0.365\mu m$ and $\text{NA}=0.5$ [74].

(2.3.3), (2.3.6), (2.3.7), and (2.3.10) as shown below:

$$\begin{aligned}
 \Delta I_{\text{posi}} &\simeq 4U_0(x_0)U_1(x_0)Z_7 \\
 &\simeq 4\frac{\lambda}{\pi}[\text{Si}(\infty) - \text{Si}(\pi)] \\
 &\quad \times \frac{3(\text{NA})^3 k\lambda 2(\pi) \cos(\pi)}{\pi \pi^3} Z_7 \\
 &= 24\frac{\lambda}{\pi} \left[\text{Si}(\pi) - \frac{\pi}{2} \right] \frac{(\text{NA})^3 k\lambda}{\pi^3} Z_7.
 \end{aligned} \tag{2.3.12}$$

The second term of Eq. (2.3.6), originating in the ringing wave from the left side boundary $(-D/2)$, is neglected because it is a contribution of the ringing wave from the opposite edge and is inversely proportional to the factor $k\text{NA}D$.

2.3.2 Single Line Negative Pattern

In the negative single line pattern case, the transmittance of the object and the spectrum are given by

$$T(x) = 1 - \text{rect} \left[-\frac{D}{2}, \frac{D}{2} \right], \quad (2.3.13)$$

$$\tilde{T}(\xi) = \lambda \delta(\xi) - D \text{sinc} \left(\frac{\pi D}{\lambda} \xi \right). \quad (2.3.14)$$

Similarly to the derivation of Eq. (2.3.3), we calculate the aberration-free amplitude as

$$\begin{aligned} U_0(x) &= \int_{-\text{NA}}^{\text{NA}} d\xi \left[\lambda \delta(\xi) - D \text{sinc} \left(\frac{\pi D}{\lambda} \xi \right) \right] e^{ik\xi x} \\ &= \frac{\lambda}{\pi} [\pi - \text{Si}(k\text{NA}x_-) + \text{Si}(k\text{NA}x_+)]. \end{aligned} \quad (2.3.15)$$

In Fig. 2.3.2, we show the amplitude profile of the wave aberration-free image expressed in Eq. (2.3.15). Considering the condition given in Eq. (2.3.11), we conclude that the first sidelobe peak position for the negative single line pattern is the same as that of the positive single line pattern given in Eq. (2.3.10). Therefore, the intensity difference of the first sidelobe peaks is obtained by

$$\begin{aligned} \Delta I_{\text{nega}} &\simeq 4U_0(x_0)U_1(x_0)Z_7 \\ &\simeq 4\frac{\lambda}{\pi} [\pi - \text{Si}(\infty) + \text{Si}(\pi)] \\ &\quad \times \frac{-3(\text{NA})^3 k\lambda 2(\pi) \cos(\pi)}{\pi \pi^3} Z_7 \\ &= 24\frac{\lambda}{\pi} \left[\frac{\pi}{2} + \text{Si}(\pi) \right] \frac{(\text{NA})^3 k\lambda}{\pi^3} Z_7. \end{aligned} \quad (2.3.16)$$

Finally, using the analytic results of our calculation based on the coherent ($\sigma = 0$) imaging theory given in Eqs. (2.3.12) and (2.3.16), the negative/positive ratio of aberration sensitivity of the intensity difference can be calculated as

$$\left. \frac{\Delta I_{\text{nega}}/Z_7}{\Delta I_{\text{posi}}/Z_7} \right|_{Z_7 \rightarrow 0} = \frac{\text{Si}(\pi) + \frac{\pi}{2}}{\text{Si}(\pi) - \frac{\pi}{2}} \simeq 12.17. \quad (2.3.17)$$

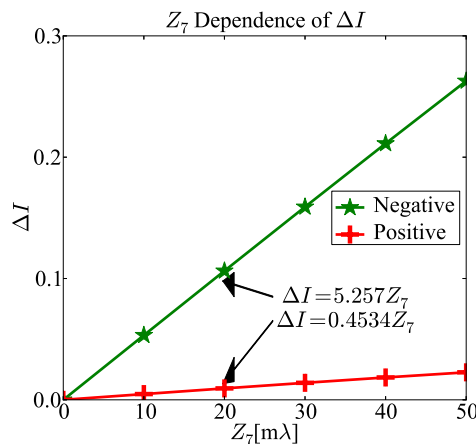


Figure 2.3.3: Coherent image simulation results for the intensity difference ΔI in Eq. (2.2.23), showing the linear dependency of coma aberration Z_7 for positive and negative single-line patterns, where $D = 4.0\mu m$. The simulation parameters are $\lambda = 0.365\mu m$, $NA = 0.5$, and coherence factor $\sigma = 0.0$ [74].

2.3.3 Simulation Results

Using scalar diffraction theory, Eqs. (2.3.12) and (2.3.16) predict a linear dependence of the left/right intensity difference on coma. In order to confirm the linearity between the intensity difference of the first sidelobe peaks and the aberration, we perform a numerical simulation using an in-house photolithographic simulator.

The imaging parameter is assumed to be the *i*-line exposure tool with $NA = 0.5$. The width of the single line is $4.0\mu m$ so that we may neglect the ringing wave from the opposite boundary. The simulated correlation between the intensity difference of the first sidelobe peaks and the seventh-coefficient of Zernike polynomials Z_7 is as shown in Fig. 2.3.3.

The intensity difference has excellent linearity with the amount of aberration. In addition, the aberration sensitivity of the negative pattern is 11 times higher than that of positive pattern; this is consistent with the analytic result, which is 12 times higher as given by Eq. (2.3.17).

2.4 Aberration for Plural Line

This section describes the correlation of the aberration sensitivity with the ratio of the line width to the pitch (hereafter called the “duty ratio”) where the pattern is assumed to be negative/positive L&S patterns under coherent illumination $\sigma = 0.0$. The aberration sensitivity dependence on duty ratio, which is derived from linear curve fitting explained below, is as shown in Figs. 2.4.1(a) and (b). The negative and positive patterns both show their maximum sensitivity near the duty ratio 0.5; however, they show some differences around duty ratios 0.0 and 1.0.

Additionally, we discuss the dependency of both the circle illumination coherency σ and the duty ratio in Chapter Appendix 2.7.2. The coherent illumination condition ($\sigma = 0.0$) produces the highest aberration sensitivity regardless of the duty ratio.

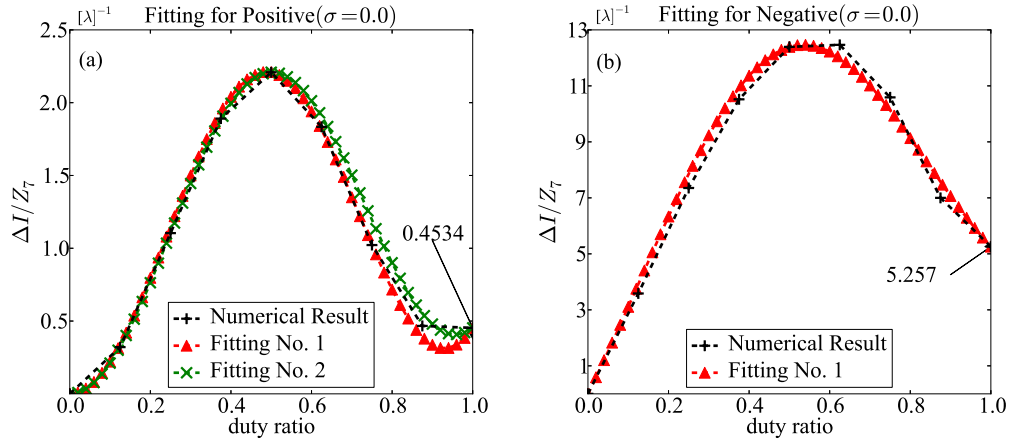


Figure 2.4.1: Duty ratio dependency of the intensity difference ΔI_{posi} under the coherent illumination $\sigma = 0.0$. In addition to the numerical simulation result, the fitting curves using the parameters of Fitting No.1 and Fitting No. 2 are plotted in part (a). Similarly, the numerical simulation result and the fitting curves using Fitting No.1 are plotted in part (b). These imaging parameters are exactly same as the coherent case ($\sigma = 0.0$) in Fig. 2.7.2 [74].

If we assume a positive N-line pattern with pitch P and line width D ,

the spectrum function is given as

$$\begin{aligned} \tilde{T}(\xi) &= \sum_{m=-\infty}^{\infty} (-1)^{(N+1)m} N D \\ &\times \operatorname{sinc}\left(\frac{\pi D}{P} m\right) \operatorname{sinc}\left[\frac{\pi N P}{\lambda} \left(\xi - \frac{m \lambda}{P}\right)\right]. \end{aligned} \quad (2.4.1)$$

Because the numerical validation is performed under the condition $\lambda/P < \text{NA} < 2\lambda/P$, it is adequate to consider the diffracted light up to the first-order. After substituting of the spectrum $\tilde{T}(\xi)$ given in Eq. (2.4.1) for the amplitude $U_0(x)$ in Eq. (2.2.7), we can derive the amplitude including the terms up to diffraction order $|m| = 1$.

In order to explicitly indicate the correlation of the intensity difference ΔI with the duty ratio ($\equiv D/P$), we introduce the following representation:

$$U_0(x) = \frac{\lambda D}{\pi P} \left[C_0^{0\text{th}}(x) + C_0^{1\text{st}}(x) \operatorname{sinc}\left(\frac{\pi D}{P}\right) \right], \quad (2.4.2)$$

where the coefficients $C_0^{0\text{th}}(x)$ and $C_0^{1\text{st}}(x)$ are independent of duty ratio. The $C_0^{0\text{th}}(x)$ represents the contribution from the 0th diffracted light, while the coefficient $C_0^{1\text{st}}(x)$ represents the interference between the +1st order and -1st order diffracted lights. In the same manner, the amplitude $U_1(x)$ in Eq. (2.2.8) is the contribution linear to the aberration:

$$U_1(x_0) = \frac{\lambda D}{\pi P} \left[C_1^{0\text{th}}(x) + C_1^{1\text{st}}(x) \operatorname{sinc}\left(\frac{\pi D}{P}\right) \right], \quad (2.4.3)$$

where the coefficients $C_1^{0\text{th}}(x)$ and $C_1^{1\text{st}}(x)$ are independent of duty ratio as before.

After the substitutions of Eqs. (2.4.2) and (2.4.3) for Eq. (2.2.23), the intensity difference ΔI_{posi} is obtained as follows:

$$\begin{aligned} \Delta I_{\text{posi}} &= \left(\frac{\lambda D}{\pi P}\right)^2 \left[a_0 + a_1 \operatorname{sinc}\left(\frac{\pi D}{P}\right) \right. \\ &\quad \left. + a_2 \left\{ \operatorname{sinc}\left(\frac{\pi D}{P}\right) \right\}^2 \right] Z_7, \end{aligned} \quad (2.4.4)$$

Table 2.4.1: Result of fitting parameter for Positive L&S

data	sampling positions of duty ratio	a_0	a_1	a_2
No. 1	(0.50, 0.75, 1.0)	0.4534	-3.145	25.62
No. 2	(0.25, 0.50, 1.0)	0.4534	-1.202	22.57

Table 2.4.2: Result of fitting parameter for Negative L&S

data	sampling positions of duty ratio	b_0	b_1
No. 1	(0.50, 1.0)	4.804	24.45

$$a_0 = C_0^{0th}(x_0) C_1^{0th}(x_0), \quad (2.4.5)$$

$$a_1 = C_0^{0th}(x_0) C_1^{1st}(x_0) + C_0^{1st}(x_0) C_1^{0th}(x_0), \quad (2.4.6)$$

$$a_2 = C_0^{1st}(x_0) C_1^{1st}(x_0). \quad (2.4.7)$$

Because the position of the first sidelobe peak x_0 , which is derived by substituting Eq. (2.4.2) for Eq. (2.2.16), is dependent on the duty ratio, the three coefficients a_0 , a_1 , and a_2 are also indirectly dependent on duty ratio.

Since the rigorous derivation of the relationship between x_0 and the duty ratio is outside the scope of this thesis, we adopt an approximation that neglects the duty ratio dependency of the coefficients a_0 , a_1 , and a_2 . Therefore, under this approximation, the correlation between the aberration sensitivity and the duty ratio can be represented with three undetermined parameters a_0 , a_1 , and a_2 . We treat the coefficients a_0 , a_1 and a_2 as fitting parameters, hereafter called “parameter.” In what follows, the three undetermined parameters a_0 , a_1 , and a_2 are determined by curve fitting, targeting the numerical calculation results as shown in Fig. 2.4.1.

We sample three different duty ratios (abscissa axis) and their corresponding aberration sensitivities (ordinate axis) assuming a five-bar pattern. Table 2.4.1 shows the curve fitting results of the three parameters a_0 , a_1 , and a_2 for two data sets (No. 1 and No. 2) in Eq. (2.4.4).

Both curves using different sampling point ranges show excellent agreement with the numerical simulation as shown in Fig. 2.4.1. As shown in Table 2.4.1, the most dominant parameter in Eq. (2.4.4) is the parameter a_2 , which is significantly larger than a_1 .

Therefore, the duty ratio that produces the highest aberration sensitivity

should be determined by the behavior of a function proportional to the fitting parameter a_2 :

$$\left(\frac{\lambda D}{\pi P}\right)^2 \left\{ \text{sinc}\left(\frac{\pi D}{P}\right) \right\}^2 \propto \sin^2\left(\frac{\pi D}{P}\right). \quad (2.4.8)$$

Because all terms in Eq. (2.4.4) are normalized to unity, we can determine the dominant parameter.

We conclude that when Eq. (2.4.8) reaches its maximum near the duty ratio of 0.5, the sensitivity also reaches its maximum because the sine function reaches its maximum at $\pi D/P = \pi/2$.

The equivalent approximation, neglecting the duty ratio dependency of the position of the first sidelobe peak, is applied to the negative pattern case. We assume that the negative N-line pattern with pitch P and line width D and its spectrum $\tilde{T}(\xi)$ are represented as

$$\begin{aligned} \tilde{T}(\xi) &= \lambda \delta(\xi) - \sum_{m=-\infty}^{\infty} (-1)^{(N+1)m} N D \\ &\times \text{sinc}\left(\frac{\pi D}{P} m\right) \text{sinc}\left[\frac{\pi N P}{\lambda} \left(\xi - \frac{m \lambda}{P}\right)\right]. \end{aligned} \quad (2.4.9)$$

Comparing Eqs. (2.4.1) and (2.4.9), we see that only difference is the existence of the delta function. The amplitudes $U_0(x)$ and $U_1(x)$ are obtained as

$$U_0(x) = \lambda - \frac{\lambda D}{\pi P} \left\{ C_0^{\text{0th}}(x) + C_0^{\text{1st}}(x) \text{sinc}\left(\frac{\pi D}{P}\right) \right\}, \quad (2.4.10)$$

$$U_1(x) = - \frac{\lambda D}{\pi P} \left[C_1^{\text{0th}}(x) + C_1^{\text{1st}}(x) \text{sinc}\left(\frac{\pi D}{P}\right) \right]. \quad (2.4.11)$$

Therefore, the intensity difference in the negative pattern case is given by

$$\Delta I_{\text{nega}} = \frac{\lambda D}{\pi P} \left[b_0 + b_1 \text{sinc}\left(\frac{\pi D}{P}\right) \right] Z_7 + \Delta I_{\text{posi}}, \quad (2.4.12)$$

where the coefficients b_0 and b_1 depend on the duty ratio through the position of the first sidelobe peaks x_0 . Because the term ΔI_{posi} is nearly ten-time smaller than the term ΔI_{nega} , it is negligible.

Similarly to the positive pattern case, the coefficients b_0 and b_1 are assumed to be independent of duty ratio. Table 2.4.2 shows the fitted two

parameters from the results of the negative five-bar pattern under coherent illumination $\sigma = 0.0$ as shown in Fig. 2.4.1(b). The curve using the fitting parameters shows excellent agreement with the numerical simulation.

In Table 2.4.2, the parameter b_1 is nearly five-time larger than b_0 , so the dominant function term in Eq. (2.4.12) is written as

$$\frac{\lambda D}{\pi P} \text{sinc} \left(\frac{\lambda D}{\pi P} \right) \propto \sin \left(\frac{\lambda D}{\pi P} \right). \quad (2.4.13)$$

Therefore, the sensitivity will reach its maximum near a duty ratio of 0.5 where the function also reaches its maximum.

The aberration sensitivity, defined by Eq. (2.2.25), of the negative plural line pattern at duty ratio 1.0 is as shown in Fig. 2.7.2(b) as $5.257/\lambda$, which corresponds to the coefficient of the intensity difference as shown in Fig. 2.3.3. This is to say that the derived function for plural line patterns, Eq. (2.4.12), inherently includes the single line pattern for the duty ratio of 1.0. It is clear from Fig. 2.4.1 that the aberration sensitivity of plural line patterns is larger than that of single line patterns. In the case of plural line patterns, this occurs because there is constructive interference between the +1st and -1st diffracted light waves that increases the imaging amplitude.

We have demonstrated, using curve fitting targeting the numerical simulation results, that the derived Eqs. (2.4.4) and (2.4.12) reproduce the duty ratio dependence of aberration sensitivity very well. The sensitivity reaches its maximum around the duty ratio of 0.5 for both negative and positive patterns. Furthermore, the sensitivity of the negative pattern at duty ratio 0.5 is over five-time larger than positive pattern, as shown in Fig. 2.4.1(a) and 2.4.1(b).

We conclude that negative L&S patterns with duty ratio 0.5 are best suited for aberration measurement applications. Incidentally, the difference in the sensitivity between the negative and positive patterns is distinct in small duty ratio regions.

2.5 Effect of Even Wavefront Aberrations

The method we have proposed is effective for the projection lens that also have spherical aberration (Z_9), defocus (Z_4) and astigmatism (Z_5, Z_6). These aberrations are even wavefront aberrations.

Even wavefront aberrations always contribute to the image intensity only in the quadratic term, not in the linear term under the assumptions introduced above. By using the similar discussion, we can verify that the terms regarding the even wavefront aberration cancel in the subtraction between the corresponding two sidelobe intensities; therefore, the residual aberration terms in the intensity difference are expressed as a linear combination of only odd wavefront aberrations.

2.6 Summary

A new coma aberration measurement method has been achieved, which utilizes the asymmetry of the sidelobe peaks caused by the lowest-order coma aberration.

We draw attention to the pair of symmetric local peaks, and we have derived an equation representing the intensity difference of the peaks, taking into account the positional shift and intensity change of the peaks due to aberration.

Furthermore, we have rigorously proven that the intensity difference of the peaks is linear with respect to the coma aberration, and that the second-order aberration terms are canceled by the subtraction of the intensity difference based on scalar diffraction theory.

Using the aberration sensitivity formula derived for the intensity difference, we have shown that the negative pattern has an aberration sensitivity that is about ten-time higher than that of the positive pattern in the case of the first sidelobe peaks of single line pattern.

We make clear the fact that the intensity difference is created by interference between the amplitude of aberration-free imaging and the amplitude modified by aberration; therefore, the existence of the bright background light enhances the aberration sensitivity. We have studied the correlation between the aberration sensitivity and the duty ratio of plural L&S patterns. The numerical simulation results for negative and positive plural-line patterns suggest that the aberration sensitivity is the highest near the 0.5 duty ratio.

Finally, the most advantageous condition for our coma aberration measurement is the negative plural line pattern when the duty ratio is approximately 0.5.

2.7 Chapter Appendix

2.7.1 Nature of the Integral Sine Function

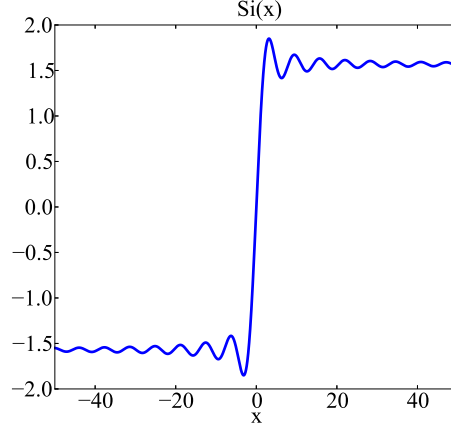


Figure 2.7.1: Integral sine function defined in Eq. (2.3.4) [74].

The integral sine function $\text{Si}(x)$ is not expressed in elementary mathematics. Fig. 2.7.1 shows the function $\text{Si}(x)$. It is possible to consider the behavior around $x = 0$ and far away $x \gg 1$, which is used in the power series expansion of the function given as follows:

$$\text{Si}(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)(2n+1)!}, \quad (2.7.1)$$

$$\begin{aligned} \text{Si}(x) = \frac{\pi}{2} - \cos x \sum_{n=0}^{\infty} \frac{(-1)^n (2n)!}{x^{2n+1}} \\ - \sin x \sum_{n=0}^{\infty} \frac{(-1)^n (2n+1)!}{x^{2n+2}}. \end{aligned} \quad (2.7.2)$$

Followed by the behavior around the origin $x = 0$, $\text{Si}(x)$ shows a first-order precipitous slope. The first local peak on both the left and right side is the result of the cubed term, and it starts the damped oscillation, which finally converges to the constant $\pm\pi/2$.

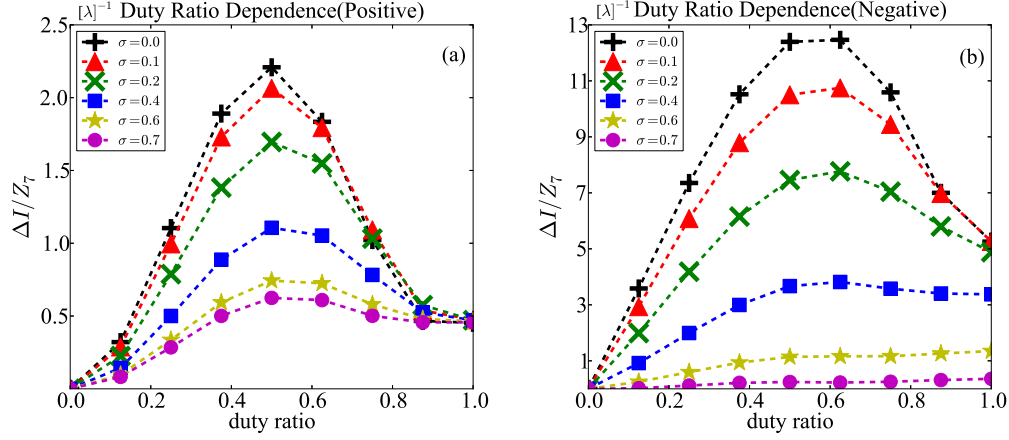


Figure 2.7.2: Duty ratio dependence of the aberration sensitivity under a wide range of illumination coherency $\sigma = 0.0 - 0.7$ for plural L&S patterns with pitch $P = 0.8\mu m$ and the duty ratio between 0.0 – 1.0. The simulation parameters are $\lambda = 0.365\mu m$ and $NA = 0.5$ [74].

2.7.2 Sensitivity on Duty Ratio and Illumination Coherence

We discuss the aberration sensitivity of the positive and negative L&S pattern under partial coherent illumination. The shape of the illumination source is assumed to be a circle, represented by the coherence factor σ . We calculate the aberration sensitivity in Eq. (2.2.25) with our imaging simulator, as showing the numerical results in Fig. 2.7.2. The result at $\sigma = 0.0$ is exactly same as shown in Fig. 2.4.1.

The characteristic, which determines the maximum aberration sensitivity occurred around a duty ratio of 0.5 for both the negative and positive pattern, is also valid for small partial coherent illumination. Regarding correlation with the factor σ , the aberration sensitivity in both Figs. 2.7.2(a) and 2.7.2(b) monotonically decreases as the factor σ increases at any duty ratio. In particular, the negative pattern shows extremely low sensitivity around $\sigma = 0.7$.

Chapter 3

Position Measurement Metrology

3.1 Metrology Sensor used in this study

In my study of all image position amplification measurement, the Laser Step Alignment (LSA) is used; wherein the laser spot can produce a significant amount of scattered light even for small wedge shapes of photoresist images to validate the metrology precision [63, 81].

The LSA measurement system has been installed as one dimensional alignment optics in a Nikon, for example, NSR-1755iA (NA = 0.50, Illumination NA = 0.30), which is an *i*-line exposure apparatus; in which, a normally incident laser sheet beam creates an oblong spot on the wafer.

As shown in Fig. 3.2.1, a periodic array of die-shaped marks is set on the wafer. When the laser spot is positioned on the marks during the stage scan, the laser beam is diffracted by the marks and detected by an optical device located at an angle $\phi = \arcsin(\lambda/d)$ relative to the wafer surface. The angle is determined by the period of the marks and the wavelength of the laser light, the top view is as shown in Fig. 3.2.2.

3.2 Position Measurement

The one wedge in one group array should have the distance “d” to the one wedge mark of the neighborhood group so that they make the interference relation as shown in Fig. 3.2.3.

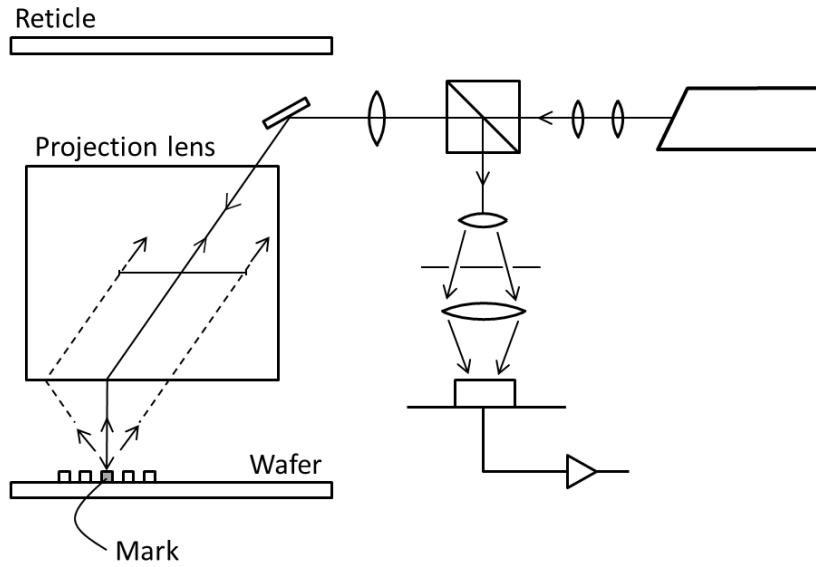


Figure 3.2.1: Laser step alignment configuration.

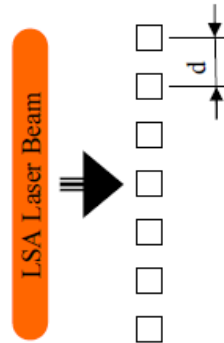


Figure 3.2.2: Basic concept of LSA; the laser oblong spot and the periodic pitch-d mark of LSA [63].

In the study, all of the length of the wedge marks or the position of the wedge marks are measured by LSA; a matrix of 3 by 17 points can be measured in both x and y in only 5 min.

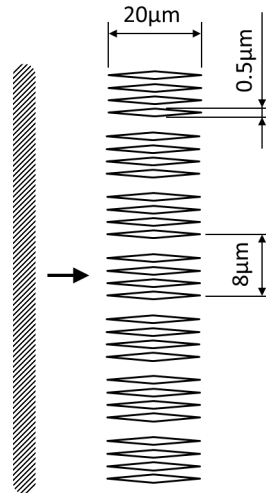


Figure 3.2.3: Basic concept of wedge mark detection; group of $0.5\text{ }\mu\text{m}$ width wedges figures are shown by $8\text{ }\mu\text{m}$ pitch-d. The laser oblong spot scans in the direction as shown by the arrow [81].

Chapter 4

Focus Measurement

4.1 Focus Wedge Mark

We have invented a focus wedge mark concept (hereafter named FWM) [80] which is suitable for the LSA measurement [81]. In Fig.3.2.3, each dense group of $0.5\ \mu\text{m}$ wedge marks is located at the position of the $8\ \mu\text{m}$ period of the grating that diffracts light into the LSA sensors, both +1st and -1st diffraction orders. The procedure is to print wedge marks on the bare silicon wafer at different locations in which the focus is changed as shown in Fig. 4.1.1.

In comparison, the L&S marks and FWM marks are printed at the same time on the same wafer along the different focus positions, wherein, the lengths of FWM are measured by LSA in the scan direction. We plot the lengths of FWM and the linewidths by SEM measurement, in which, in order to determine the best focus, the linewidths of L&S and wedge lengths are fitted to polynomial functions, individually as shown in Fig.4.1.2.

Figure 4.1.3 compares the best focus points of the linewidth by SEM and the wedge mark by LSA in several areas; where the horizontal and vertical direction focus point is averaged. The data points spread around the same straight line so that the best point of two method agree with a sufficient accuracy; therefore, the image position amplification of FWM has been successfully used for the best focus determination instead of SEM method.

However, the physical model of FWM still has been an unknown issue; therefore it is worth to analyze it in this thesis.

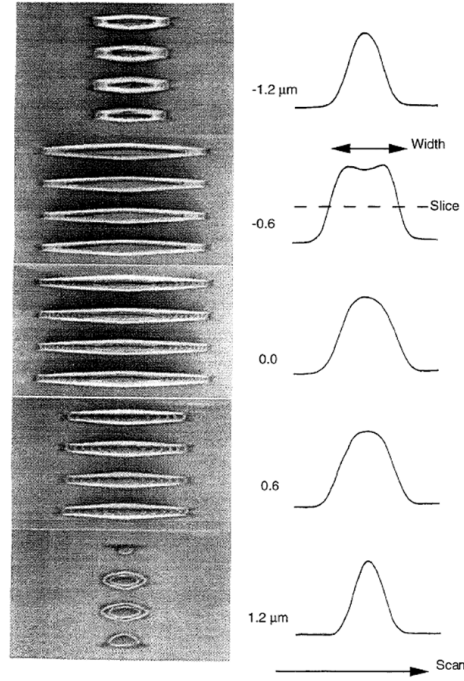


Figure 4.1.1: Left panel: Top view SEM images of the processed FWM for the five focus positions with an interval $0.6\mu\text{m}$. The length of FWM becomes rapidly shorter along out of focus. Right panel: The signal profile of LSA measurement [81].

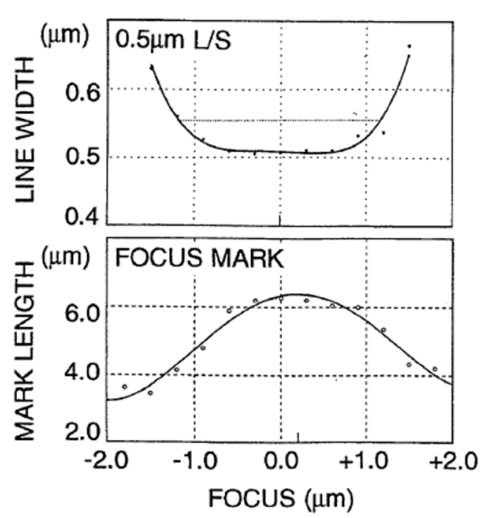


Figure 4.1.2: Top panel: Linewidth measurement by SEM for achieving the best focus. Bottom panel: Length measurement of FWM by LSA for achieving the best focus [81].

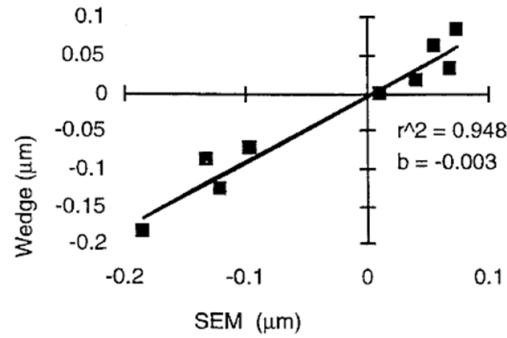


Figure 4.1.3: Best focus determined accuracy comparison between the linewidth by SEM measurement and the FWM by LSA [81].

4.2 Defocus Image Analysis

At first, to understand the image position amplification for the case of FWM, we make a physical model as shown in Fig. 4.2.1; the wedge tip contrast is

composed of the transmitted amplitudes through the projection lens which is restricted via the cut-off spatial frequency of the lens aperture; therefore, the exposure resist position of the wedge tip is "shorter" than that of the mark on the glass reticle.

The tip position has been rapidly shrunk according to the small transmittance amplitude of the defocus, which is the image position amplification of the defocus.

Object (lower half part of a single FWM)

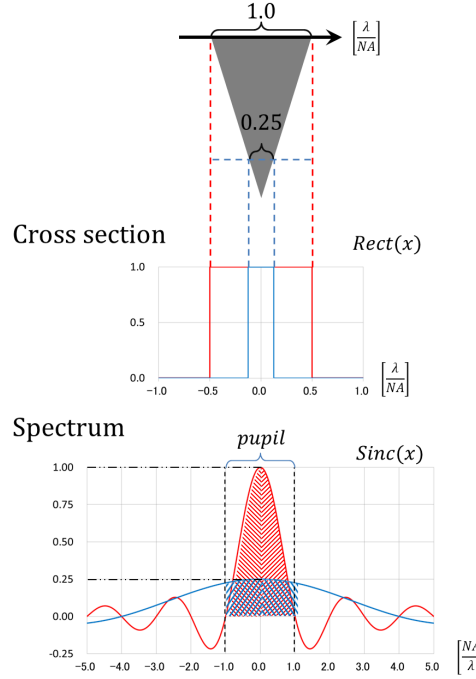


Figure 4.2.1: Top panel: Schematic of the lower half part of a single FWM on the reticle. The width of each wedge depends on the positions of the horizontal cross sections. Middle panel: The cross section profile of FWM with normalized widths 0.25 and 1.0. Bottom panel: Corresponding Fourier spectra of the cross sections as shown in the middle panel. The spectra of the rectangular function are given by sinc function. The red (blue) line is the spectra for the rectangular with width 1.0 (0.25). The pupil of imaging optics works as the optical low-pass filter so that it restricts Fourier spectra, which are drawn by the red (blue) slash region; therefore, the lens low-pass function contributes to the decreasing of image contrast at the tips of FWM.

4.2.1 One-Dimensional FWM Model

We calculate the focus characteristic of FWM in detail;

Figure 4.2.2 presents the physical model of FWM on the two-dimensional mask and defines the coordinate system. The diamond wedge marks have a width of w in the horizontal direction and a length L in the vertical direction such that $L > w$, and the marks are separated by a distance P . The length of the wedge mark will be determined by the extent of the residual resist. Because the wedge array is arranged on the mask symmetrically about the x axis, the analytic calculation is performed for $x = 0$ and $y > 0$. Here, we analyze the focus dependence of FWM model in the case of opaque wedges (wedges fabricated with chrome). Note that the reverse (positive) case can be derived easily from this negative case.

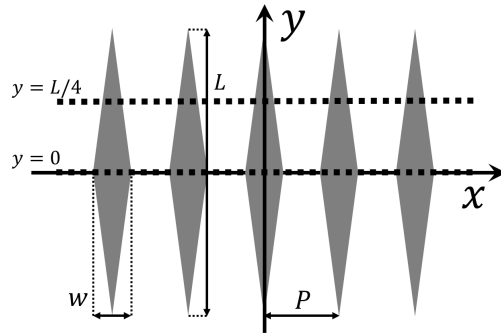


Figure 4.2.2: Two-dimensional model of FWM pattern. The wedges are arranged periodically along the x axis separated by a distance P . Each wedge is defined by a diamond of width w and length L . In the negative pattern, the grey diamond regions are opaque to the incident light [74].

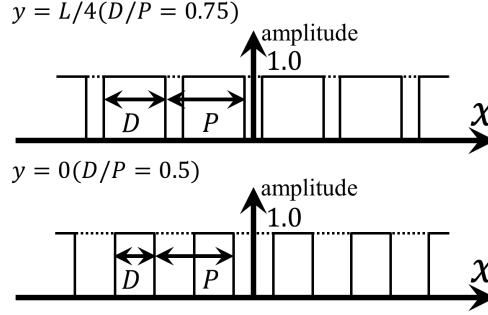


Figure 4.2.3: Profiles of FWM as shown in Fig. 4.2.2 along the lines $y = 0$ and $L/4$. As the wedges are periodic along the x axis, the profiles correspond to a one-dimensional L&S pattern [74].

Assuming that the mask has an infinite number of wedges along the x axis, we ignore the effects of diffraction along the y axis. This is a key approximation in our analysis, which in the case of FWM is likely to be valid because $L > w$, and comparisons to numerical simulations as shown in section 4.2.4, confirm the validity of the approximation. Thus, under this approximation, the intensity at each y position (> 0) is determined only by the characteristic of the x axis diffraction that reflects the horizontal cross section of FWM.

The effective pattern along the x axis can be regarded as an L&S pattern. The ratio of the pitch and linewidth, i.e., the duty ratio, of the L&S pattern depends on the position along y . For example, the cross section of FWM at $y = 0$ is equivalent to the L&S pattern with a duty ratio of 0.5, while that at $y = L/4$ is equivalent to the L&S pattern with a duty ratio of 0.75, as shown in Fig. 4.2.3. In this manner, the FWM behaves approximately as the L&S pattern with a duty ratio dependent on the distance from the center of the wedge mark.

4.2.2 Focus Scalar Coherent Imaging Theory

Assuming imaging is done under conditions of normal incidence and coherent illumination, the amplitude in the image plane, $U(x, y)$ is expressed as

$$U(x, y) = \iint_{\text{pupil}} d\xi d\eta \tilde{T}(\xi, \eta) e^{ik\phi(\xi)} e^{-ik(\xi x + \eta y)}, \quad (4.2.1)$$

where $\tilde{T}(\xi, \eta)$ is the mask spectrum function, k is the wave number ($= 2\pi/\lambda$), and $\phi(\xi)$ is a function describing the wavefront aberrations of the imaging optics including the defocus of the stage. The integral region is restricted in the pupil which is assumed to be a circle with a radius equal to the numerical aperture (NA).

The coherent imaging is as shown schematically in Fig. 1.2.1. We set the parameter region for the pitch P of the L&S to be

$$\frac{\lambda}{P} < \text{NA} < 2\frac{\lambda}{P}, \quad (4.2.2)$$

which results in the angle of the first-order diffraction being close to the aperture angle.

The spectrum $\tilde{T}(\xi, \eta)$ of the one-dimensional periodic L&S of opaque linewidth D and pitch P can be written as

$$\tilde{T}(\xi, \eta) = \left[a_0 \delta(\xi) + \sum_{m \neq 0} a_m \delta\left(\xi - \frac{m\lambda}{P}\right) \right] \delta(\eta), \quad (4.2.3)$$

where $\xi(\eta)$ is the directional cosine along the x (y) axis with a unit vector in the direction of the propagating wave. The amplitude of the zeroth-order diffraction a_0 and that of the m -th-order diffraction a_m are defined as

$$a_0 = \frac{D(y)}{P}, \quad (4.2.4)$$

$$a_m = (-1)^m \frac{D(y)}{P} \text{sinc}\left(\pi m \frac{D(y)}{P}\right) = a_{-m}. \quad (4.2.5)$$

where $D(y)$ is a linear function of y :

$$D(y) = \frac{P}{2} - \frac{P}{L}y. \quad (4.2.6)$$

For the one-dimensional pattern, the integral in Eq. (4.2.1) is performed for all spatial frequencies within the NA.

The path of the diffracted light depends on the order of diffraction, and the path difference between the zeroth- and first-order diffracted light is given by

$$\phi(\xi) = [1 - \cos \theta(\xi)] z \equiv cz, \quad (4.2.7)$$

where z is the defocus from the image plane, and $\theta(\xi)$ is the angle of the first-order diffraction:

$$\sin \theta(\xi) = \frac{\lambda}{P}. \quad (4.2.8)$$

Because the factor c in Eq. (4.2.7) is a function of λ and P , we can regard it as a constant when P and λ are fixed. To highlight the defocus characteristic of FWM, we set the other aberrations, such as spherical, astigmatism and coma, to zero.

In general, aberrations are odd or even functions with respect to the optical axis direction. The primary effect of an odd aberration is a positional shift in the image with little effect on the image intensity. Even aberrations can effectively be considered together with the defocus z . Considering the effect of even wavefront aberrations such as astigmatism, which have angular dependency about the optical axis, the best focus position at each measurement position should be determined along both the horizontal and the vertical axes. The FWM measurement can be applied to this case by rotating its array direction in the plane of the mask, and our one-dimensional physical model can also be applied after a rotation of the coordinate system.

Substituting the spectrum in Eq. (4.2.3) into Eq. (4.2.1) gives the defocus-dependent image intensity $I(x, y)$ as

$$\begin{aligned} I(x, y) &= |U(x, y)|^2 \\ &= \left| a_0 + a_1 e^{ikcz} e^{-ik\lambda/Px} + a_{-1} e^{ikcz} e^{ik\lambda/Px} \right|^2 \\ &= \left[a_0 + 2a_1 \cos(kcz) \cos\left(k \frac{\lambda}{P} x\right) \right]^2 + \left[2a_1 \sin(kcz) \cos\left(k \frac{\lambda}{P} x\right) \right]^2 \\ &= [a_0]^2 + 4a_0 a_1 \cos(kcz) \cos\left(k \frac{\lambda}{P} x\right) + 4[a_1]^2 \cos^2\left(k \frac{\lambda}{P} x\right). \end{aligned} \quad (4.2.9)$$

The intensity profile of the wedge mark along y corresponds to setting $x = 0$ in Eq. (4.2.9):

$$I(y) \equiv I(0, y) = [a_0]^2 + 4a_0 a_1 \cos(kcz) + 4[a_1]^2, \quad (4.2.10)$$

where the image intensity for a fixed y is a minimum at $x = 0$. By using

Eqs. (4.2.4) and (4.2.5), the intensity in terms of the duty ratio is given by

$$I(y) = \left[\frac{D(y)}{P} \right]^2 - \frac{4}{\pi} \frac{D(y)}{P} \sin \left[\frac{D(y)}{P} \pi \right] \cos(kcz) + \left\{ \frac{2}{\pi} \sin \left[\frac{D(y)}{P} \pi \right] \right\}^2. \quad (4.2.11)$$

We define the simplest resist developing model in which the resist reacts chemically when the incident intensity is above a predefined threshold, and the length of the wedge mark at any given threshold value $I(0, y)$ is determined by Eq. (4.2.11) through $D(y)$.

The solution to this nonlinear equation is combined with Eq. (4.2.11) to obtain the length of the wedge. The second term on the right-hand side of Eq. (4.2.11), which is linear with respect to $\cos(kcz)$, shows the essential focus dependence of the wedge length and corresponds to the interference of three beams diffracted from the one-dimensional L&S pattern.

4.2.3 FWM Metrology

When the pattern is projected on the wafer, the dimensions are appropriate for the resolution and focus characteristic of an *i*-line exposure apparatus. The FWM reticle is exposed at different wafer locations and under different focus conditions which consist of typically up to 10 settings within the depth of focus. After the wafer is developed, the wedge lengths of FWM located in different directions are measured by an optical sensor equipped with the exposure apparatus, such as a laser spot scanning sensor or a two-dimensional image sensor. Then, we obtain the focus curves of FWM at different wafer locations and calculate the best focus position by fitting them with a quadratic polynomial function.

4.2.4 Numerical Calculation

Figure 4.2.4 shows the focus dependence of the wedge length for various threshold intensities. The analytical solutions of Eq. (4.2.11) and the numerical results of a lithography simulation for an exact two-dimensional wedge model under coherent illumination conditions are shown as the dashed lines and symbols, respectively. The simulations are performed with the lithograph simulation software Dr. Litho [82] assuming an *i*-line exposure apparatus ($NA = 0.5$, $\lambda = 0.365\mu m$, $\sigma = 0.0$) with a reduction magnification of

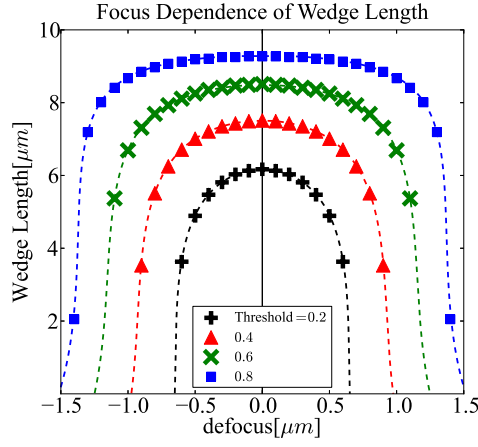


Figure 4.2.4: Focus dependence of FWM wedge length for various threshold intensities calculated using Eq. (4.2.11) (dashed lines) and numerical lithography simulations (symbols) [74].

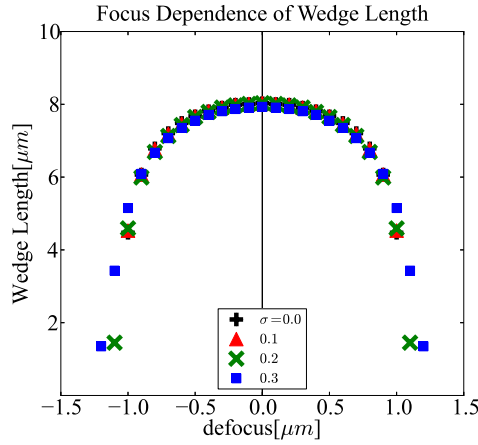


Figure 4.2.5: Illumination coherence dependence of FWM focus characteristic; the figure shows that our theory for coherent illumination conditions is valid for coherent factors of $\sigma < 0.3$ [74].

$\beta = 1/4$. Furthermore, the wedge parameters were $w = 0.5 \mu\text{m}$, $L = 10.0 \mu\text{m}$, and $P = 1.0 \mu\text{m}$. We see that the analytical results are in excellent agreement with the numerical simulations.

Since ideal coherent illumination conditions are impossible to achieve in

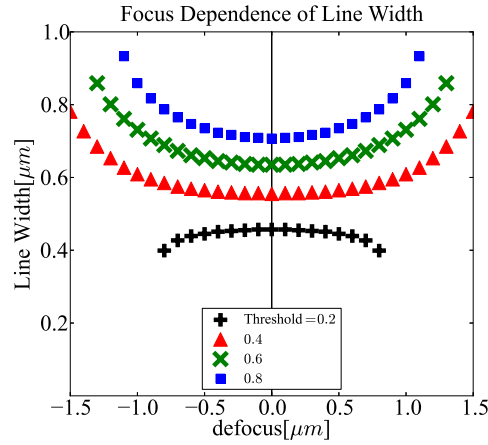


Figure 4.2.6: Defocus dependence of the linewidth of the L&S pattern for various threshold intensities calculated using numerical lithography simulations [74].

reality on the optical aligner, the focus dependence of FWM under finite coherent illumination conditions needs to be considered. We thus perform simulations for circular illumination defined by σ for $0.0 < \sigma < 0.3$. The wedge length as a function of the defocus under these conditions is as shown in Fig. 4.2.5. We find that the defocus dependence of FWM is almost the same around the best focus even for different σ . The effects of the illumination appear at large defocus values and are more apparent for $z > 1.0 \mu\text{m}$.

4.2.5 Verification of Focus Sensitivity

The notable advantage of FWM measurement method is that the physical quantity is amplified relative to the focus variation of L&S pattern; therefore, the following measuring issues are solved as follows:

Regarding the desired precision for the linewidth measurement, it is about 2.5% of the CD linewidth. For example, when we employ typical L&S pattern, e.g., one with a 500-nm width, for the focus measurement, the CD variation by the focus change must be measured about 1.25 nm.

The precision of the optical sensor utilizing a wavelength λ may approach $\lambda/100$. When the optical sensor using a 500-nm wavelength has the precision close to 5 nm, the optical sensor cannot be used for the usual focus measurement of L&S because of this resolution limit.

In contrast, the acceptable precision in the length of wedge is less than or equal to 25 nm because the focus-error-induced variation of FWM is geometrically enhanced; for example, the magnification of 20 times is easily possible in comparison with that of L&S pattern.

The following comparison reveals the focus sensitivity difference between the L&S pattern and the FWM.

The best-focus measurement using the L&S pattern involves measuring the linewidth on the processed wafer, and the focus dependence of the linewidth is as shown in Fig. 4.2.6. The simulated results are obtained with the same basic parameters used in the coherent illumination simulations in section 4.2.4. The dimensions of the L&S pattern are a linewidth of $0.5 \mu\text{m}$ and a pitch of $1.0 \mu\text{m}$. This geometry is equivalent to that as shown in Fig. 4.2.3 for the FWM at $y = 0$.

To compare the focus sensitivities of FWM and L&S pattern, we introduce a fitting function for the curves as shown in Figs. 4.2.4 and 4.2.6. Because the influence of the focus is the same as an even aberration, the focus dependence of the wedge length (FWM) and the linewidth (L&S) can be expressed as a quadratic function of z :

$$f(z) = c_0(z - z_0)^2 + c_1, \quad (4.2.12)$$

where c_0 , c_1 , and the best-focus position z_0 are determined by the fit. The focus sensitivity around the best focus can then be defined by

$$(\text{focus sensitivity}) \equiv \frac{1}{2} \frac{\partial^2}{\partial z^2} f(z) = c_0. \quad (4.2.13)$$

The focus sensitivities of FWM and L&S pattern are as shown in Fig. 4.2.7 as a function of the thresholds. We find that the sensitivity of FWM is up to ten times larger than that of the L&S pattern. The ratio of the absolute values of the focus sensitivity, i.e.,

$$(\text{focus sensitivity ratio}) \equiv \frac{|c_0|_{(\text{FWM})}}{|c_0|_{(\text{L\&S})}}, \quad (4.2.14)$$

where it defines the amplification factor of the sensitivity. In Fig. 4.2.7(b), the largest ratio, indicated by the arrow, is that calculated using the two sensitivities indicated by the arrows in (a). Figure 4.2.7(b) shows the case

for threshold intensities above 0.5 in which the focus sensitivity ratio is close to 20, which corresponds to the geometrical magnification.

Including the practical process considerations, such as resist development, the proper dose energy depends strongly on the exposure conditions [83–85], the illumination pupil, whether the pattern is isolated or periodic, and whether the resist is negative or positive. However, the merit of the sensitivity amplification is still valid in the large dose range, as observed in Fig. 4.2.7(a).

Furthermore, we note that the FWM measurement enables us to realize a much higher focus sensitivity than that of L&S pattern, around exposure energies (20–40%) where the L&S has little focus sensitivity.

4.3 Best Focus Function

4.3.1 Defocus Dependence Specified by Cosine Curve

The imaging formula for coherent illumination, given in Eq. (4.2.10), reveals that the intensity depends on the cosine of the defocus, and the period of the cosine function will depend on the positions of the diffracted beams in the pupil plane.

Therefore, the period may be modified under partially coherent illumination conditions to changes in the positions of the diffraction beams. An analytical treatment under partial coherent illumination conditions is known to be difficult, and investigating the influence of the illumination on the FWM focus dependence is a topic for future work.

4.3.2 Future Application to Latest Exposure System

In the case of the latest immersion ArF exposure apparatus using high NA optics, the inevitable decrease in the depth of focus means that it has become important to measure the best focus to ensure the desired pattern printing onto the wafer.

We consider that, even in the latest ArF exposure apparatus, the physically basic behaviors of FWM can be explained by our one dimensional model. However, it should be noted that such ArF exposure apparatus utilizes various illumination conditions [86–90]. The focus determination equation derived in this thesis is assumed to be for small σ illumination cases, not far from coherent illumination. Therefore, to extend the range of appli-

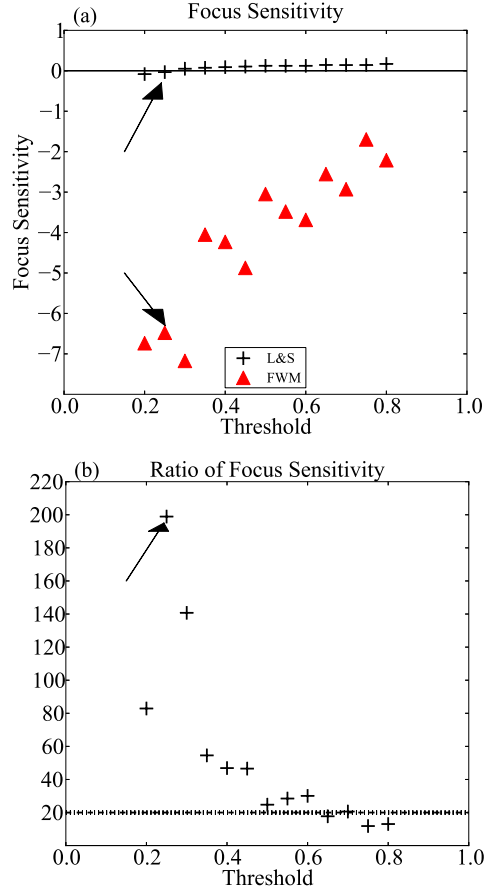


Figure 4.2.7: (a) Threshold intensity dependence of the focus sensitivity of FWM and L&S pattern. (b) Focus sensitivity ratio as defined by Eq. (4.2.14) as a function of the threshold. The dashed line represents the geometrical magnification defined by the ratio of the length and width of a wedge. The largest ratio indicated by the arrow is calculated using the two sensitivities indicated by the arrows in (a) [74].

cation of our theory, we will need to consider the effects of the illumination condition more carefully and may need to modify the focus determination equation.

4.4 Summary

We have outlined the concept of FWM for the best-focus measurement and analyzed the defocus characteristic. An analytical equation for coherent illumination conditions has been derived using scalar diffraction theory, and because the wedge has a length which is much larger than its width, we are able to construct a one-dimensional model corresponding to the L&S pattern with variable duty ratio.

The derived FWM focus equation which is characterized by a cosine function provides us the best-focus position with a high accuracy.

We have demonstrated that the focus sensitivity of FWM is much higher than that of L&S pattern, and the FWM remains highly sensitive to the defocus variation even at the dose energy region where the L&S pattern has little focus sensitivity.

Chapter 5

Linewidth Measurement

5.1 Double Exposure Magnification

We have created a new linewidth measurement that is an idea of the double exposure incline magnification [75,91–94], which utilizes the wedge conversion ratio as

$$W = L \sin \alpha, \quad (5.1.1)$$

where W denotes the linewidth of the pattern and L denotes the length of the double exposed incline pattern, in which the method uses two inclination angle lines by double exposure in order to amplify tiny variations in the linewidth to a large amount of an exact ratio defined by the geometrical angle.

The most important linewidth is the critical linewidth; therefore, the linewidth is usually called the critical dimension (CD) which is also used in Chapter 4, moreover, when the CD becomes smaller, the corresponding angle also becomes smaller, thereby ensuring sufficient measurement precision. The concept is illustrated as (a), (b), and (c) as shown in Fig. 5.1.1.

We discuss the test reticle and exposure pattern, that is, in each block on the reticle lines are arranged in seven groups with a periodicity of $8 \mu\text{m}$ which is designed to diffract the incident light as per the LSA operation in a manner similar to that of the FWM experiments.

Figure 5.1.2 shows SEM images of the test marks printed on the wafer in which the marks have cigar-type shapes and are oriented along the long diagonal of the overlapping diamond area of two lines.

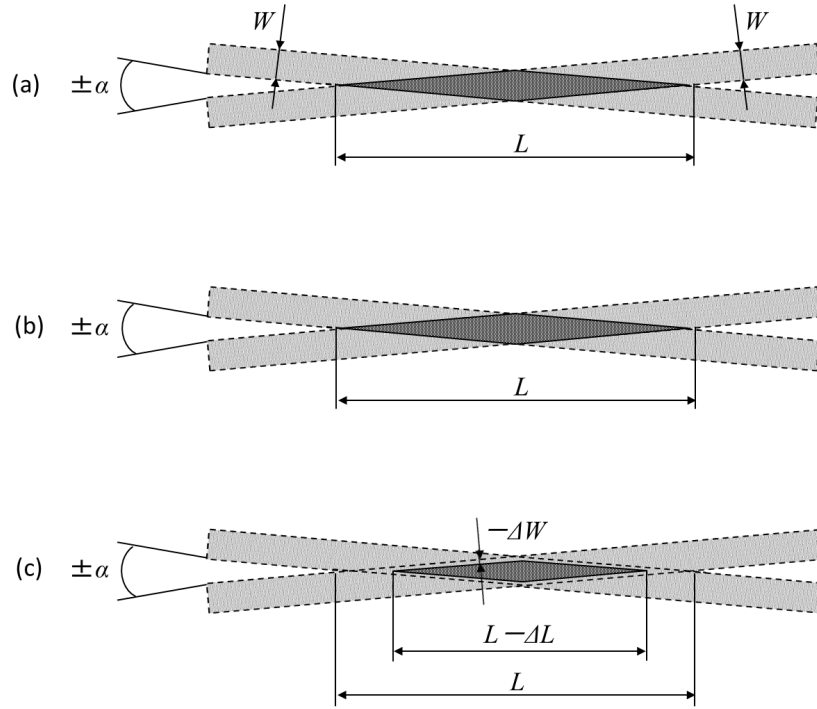


Figure 5.1.1: (a): Geometry underlying double exposure incline concept, L and W are related as in Eq. (5.1.1). (b): Double exposure at best focus; the linewidth is the same as the line. (c): Double exposure at defocus; the linewidth is tiny narrower than that of the best focus, wherein, the linewidth narrowing, $-\Delta W$, is magnified to the line length, $L - \Delta L$.

5.2 Simulation and Experiments

5.2.1 Simulation Results

Theoretical modeling is based on simulation of aerial images using VENUS, an in-house photoresist image simulation tool, wherein the intensity level corresponding to the desired linewidth of the image is calibrated in a regular exposure process for the real photoresist.

The double exposure projection image makes the round top shape as shown in Fig. 5.2.1, at left panel, which differs from a diamond shape of an overlapping area of two lines on the reticles as shown at right panel; therefore, the left panel shows cigar-type length that is same as shown in Fig. 5.1.2.

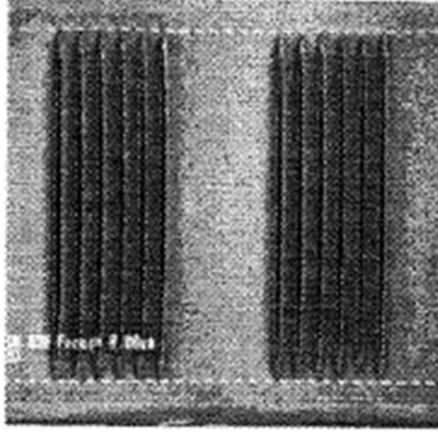


Figure 5.1.2: Top view SEM images of test marks which are doubly exposed on the photoresist. The dotted lines (above and below trace signals) show the vertical length of the marks determined by SEM measurement [75].

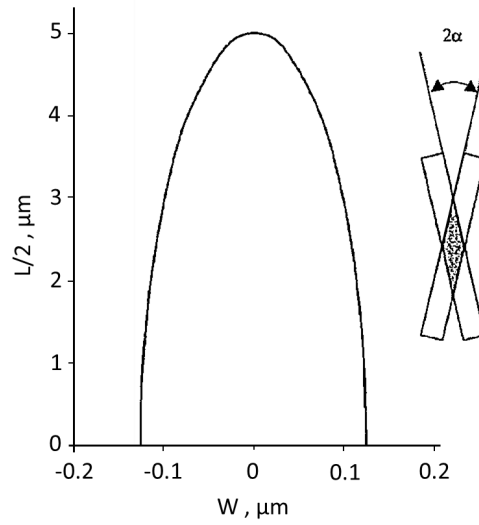


Figure 5.2.1: Right panel: Double exposure of the original two reticle lines, linewidth $W = 0.25 \mu\text{m}$, angle of inclination $= \pm 1.43$ degrees. Left panel: Lithography simulation result for W and $L/2$ which top is round shape; where, $\text{NA} = 0.6$, $\sigma = 0.7$ and $\lambda = 248 \text{ nm}$ [75].

5.2.2 Experimental Results

To study new technique, we use the specially designed binary reticles with reduction ratio of 5: 1 and 4: 1, where a basic reticle cell comprises two blocks with lines of the same CD with different orientations in different blocks.

Figures 5.2.2(a) and 5.2.2(b) present certain important results of comparison between the behaviors of L and W as follows:

1) The data demonstrate L and W behavior with change in focus position, where the lines as shown in Fig. 5.2.2(a) have a nominal CD of $0.25\ \mu\text{m}$ and $\alpha = \pm 1.43$ degrees, and those as shown in Fig. 5.2.2(b) have a nominal CD of $0.20\ \mu\text{m}$ and $\alpha = \pm 1.15$ degrees, in which the marks as shown in Fig. 5.2.2(b) are slightly overexposed.

2) We note that the length and width of the test marks correlate well with each other over the entire range of focus, where the correlation coefficient = 0.96 and 0.97 as shown in Figs. 5.2.2(a) and 5.2.2(b), respectively. Further, from the insets as shown in Figs. 5.2.2(a) and 5.2.2(b) we observe that the mark length varies linearly with its width as $L = k_w W$. Using the aerial image simulation, we recall that the coefficient k_w equals $\sin^{-1}(\alpha)$, and therefore, we have $k_{w1} = 40.07$ and $k_{w2} = 49.83$ for the marks as shown in Figs. 5.2.2(a) and 5.2.2(b) respectively.

3) On the contrary, in the printed image we use k_{w1} and k_{w2} values of 55.32 and 70.58 respectively, so that the marks printed in the photoresist are longer than the simulated images. Despite the difference in the length, it is significant that the experimental value rate, k_{w2}/k_{w1} , of 1.276 is very close to the simulated value rate of 1.243, where this consistency clearly reflects the $\sin(\alpha)$ relationship between L and W .

4) The several linewidth double exposure results for the photoresist are obtained via an optical microscope with a CCD (charge coupled device) camera measurement, in which, the normal CD value and the inclination angle are as shown in Fig. 5.2.3, where the exposure condition is the same as that corresponding to Fig. 5.2.1.

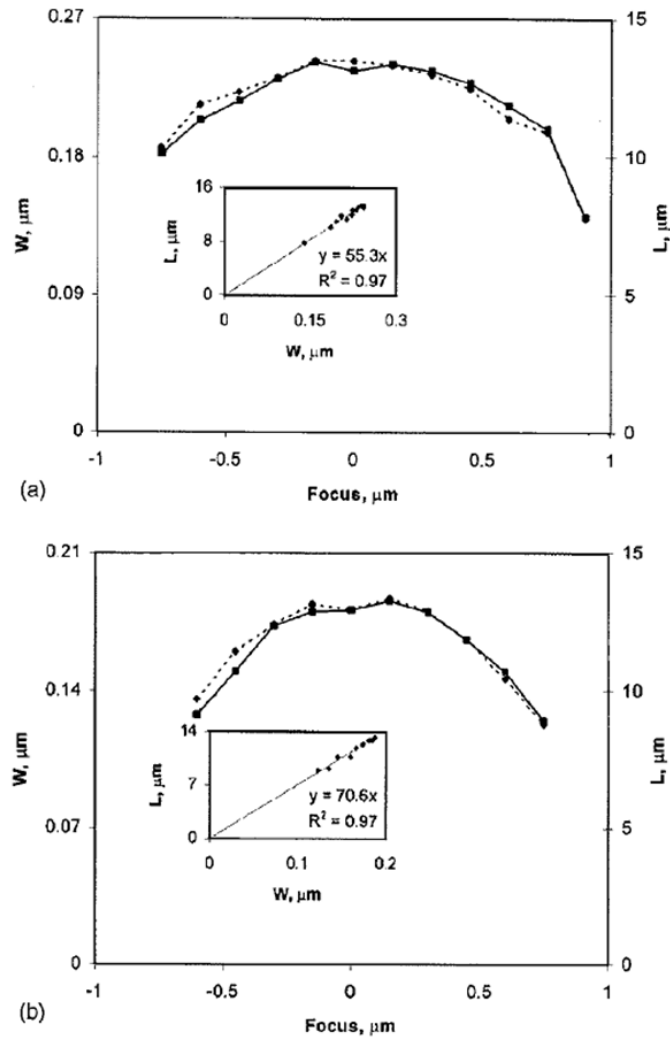


Figure 5.2.2: (a) Results of measurement of the width (dotted line) and length (solid line) of marks printed with the photoresist at various focus positions. The left y axis values represent the calculated W , and the right y axis values represent the measured length of L . The normal CD is 0.25 μm and $\alpha = \pm 1.43$ degrees. The exposure conditions are the same as Fig. 5.2.1. The inset in the graph shows the correlation between W and L with a fitting linear function. (b) The normal CD is 0.20 μm , and the conditions are identical to those for Fig. 5.2.2(a) [75].

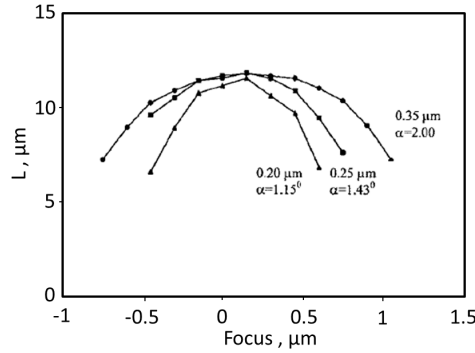


Figure 5.2.3: Focus curves for various angles formed by the intersection of two lines and spaces patterns. The ordinate represents the length of the doubly exposed incline patterns. It is confirmed that the smaller intersection angles realize the higher measurement magnifications for the narrow line width. Considering the practical exposure processes, the excessive magnification would not be suitable. Therefore, the intersection angle should be optimized [75].

5.3 Automated Optical Measurement

5.3.1 Offset in Measurement

We use LSA for the length measurement, where the signal duration is proportional to the mark length, in which the diffraction efficiency of the marks can vary at line-end regions due to decreasing photoresist thickness that distorts the proportionality between mark length and the diffracted signal.

In order to avoid this effect, we measure the signal width at a non-zero slice level of the signal waveform; therefore, we find a slice level of 20 % or more which are practically negligible through the resists studied. As a result, the “optical” length is smaller than the length by SEM which can be observed from Fig. 5.3.1, where the fitting curve as shown in Fig. 5.3.1 represents a linear function with a constant shift C which is expressed as

$$L_{LSA} \sin \alpha = CD_{SEM} - C, \quad (5.3.1)$$

where L_{LSA} is the length measured by LSA and CD_{SEM} is that by SEM.

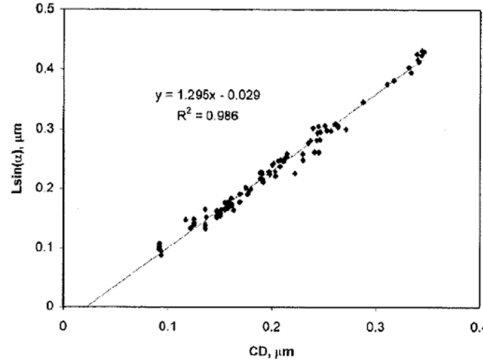


Figure 5.3.1: Comparison of the automatic line length measurement results (LSA) and the SEM-CD results [75].

5.3.2 Measurement Precision

One of the main advantages of the scanning optical method is its measurement precision. The precision (3 standard deviations) has been reported to be less than 25 nm at best focus and less than 50 nm at the ends of the focus range, in the 1990s.

At $\alpha = \pm 1.43$ degrees, the corresponding values range from 0.5 to 1.0 nm for CD values in the range of 0.13 - 0.25 μm . Another benefit of using the LSA is its high-speed operation. The process takes only 0.5 s including traveling time to measure one site.

5.4 Summary

A new technique for characterization of exposure tool imaging performance has been developed that is based on using the LSA, to measure the large length of test marks printed in a photoresist in a two-step process. In the first step, the line of the specific CD is exposed with energy one half of the nominal energy for a fully formed image. In the second step, the line of the same CD with a slightly different inclination angle is exposed with the same energy as in the first step.

The length of the resulting test mark is linearly proportional to the width of a line printed in a regular exposure process and is inversely proportional

to $\sin \alpha$, where 2α is the angle between two inclinations. The mark lengths are then converted into CD values using the calibration relationship between the length and CD of a given process. Measurement precision of 0.5 nm has been achieved for sub-150 nm CD.

This methodology is applicable both to the optical measurement and SEM measurement, and also to the future narrow linewidth measurement, this is because the double exposure inclination technique has the general feature of the scaling rule on measurement.

Chapter 6

Displacement Measurement

6.1 Moire Displacement Metrology

In the field of optical measurement, the moire effect is a well-studied phenomenon [95, 96]. The moire method has a particularly unique advantage in which it can significantly improve precision with no further demands in the sensitivity of the optical sensor. For this reason, the moire method is widely used for optical displacement measurement in optical encoders [64, 97, 98], and curved surface metrology [99–102]. We attempt here to implement moire technologies into Optical Registration Metrology System (ORMS) in order to improve the measurement precision. There are many previous researches in moire, in particular for measuring 3D topology. However, in the context of ORMS applications, our implementation to photolithography belongs to D moire measurement, similar to that found in optical encoders.

With regard to optical encoder, Liu [98] have proposed the electrical signal process using two independent moire fringe signals with same phase. Many types of moire measurement utilize a two dimensional monitor camera to observe real-time topological deformation [102, 103]. Kafri has studied a double exposure technique [95, 101] in which the camera film realizes a moire fringe image by exposing a shifted grating on a previously exposed grating as a reference pattern.

In registration metrology, measurement of the displacement between two points by photoresist images through double exposure is required. The photoresist image can indicate several problems in the exposure of a wafer aligner, such as registration error, machine vibration, and exposure stage error. Such

information is valuable because one can perform a calibration for the wafer aligner.

The large advantage of using moire technology for the ORMS is that the physical quantity to be measured is highly magnified. For a double exposed wafer, we measure the distance between two marks exposed in sequence. The distance is magnified by the moire phenomena and calculated by dividing the moire magnification factor from the measured distance.

In this case, theoretically, the measurement error is not magnified. This is because the division operation reduces the measurement error with the inverse of the moire magnification factor. As compared with conventional approaches for reducing the error, our method does not require slow machine movements to record a large number of data points.

6.2 Alternating Direction Moire

6.2.1 Bar-in-Bar Method

The ORMS measurement precision is expressed by $\sigma_{metrology}$. A first mark or set of marks is created on a wafer by a first exposure. A second mark or a set of marks is created by a second exposure. Assume that σ denotes the measurement precision of any one of the photoresist marks obtained with metrology optics.

Then, σ is the measurement error of the sensor defined as the statistical error distribution from numerous data points [76]. Assuming that a conventional ORMS measures the distance between two marks, $\sigma_{metrology}$ is given by

$$\sigma_{metrology} = \sqrt{2}\sigma. \quad (6.2.1)$$

In our case, the marks are photoresist marks so that there is less waveform error for the optical signal. In conventional systems, the fundamental mark arrangements for the ORMS are the box-in-box [104,105] or bar-in-bar [105]. In Fig. 6.2.1, the sequential process of the double exposure for the bar-in-bar system is shown.

In this process, two types of reticles, R1 and R2, are used for the photo mask. After the reticle R1 is exposed on the wafer, the reticle R2 is also exposed with the stage in motion to make the bar-in-bar pattern on the photoresist. In case of the bar-in-bar system, a line is exposed between two

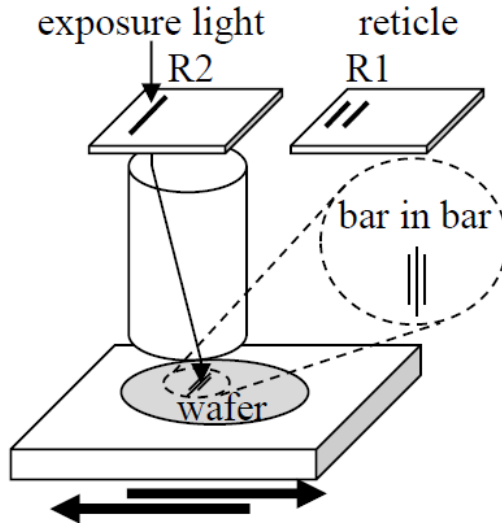


Figure 6.2.1: Double exposure to print bar-in-bar marks on the wafer using reticle R1 and reticle R2 [76].

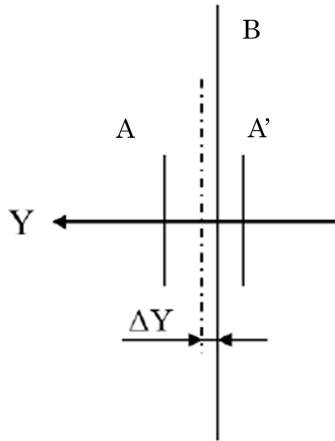


Figure 6.2.2: Bar-in-bar marks on the wafer: the shift is ΔY , where A , and A' are exposed by reticle R1, and B is by reticle R2 as shown in Fig. 6.2.1.

previously exposed lines. The relative position of the center line from the two outside lines corresponds to both the polarity of direction and the amount of displacement. Bar in bar relation is as shown in Fig. 6.2.2. This bar-in-bar system is also known as electrical micro-measurement [106].

6.2.2 Introduction to Alternating Direction Moire

To implement moire metrology into the bar-in-bar method for the ORMS, we develop a new concept using two types of moire, $+\theta$ and $-\theta$, with opposite phases. Fig. 6.2.3 shows the schematics of the general principle of Alternating Direction Moire (ADM) where the $-\theta$ moire is assigned to A and A', and the $+\theta$ moire is assigned to B. The moire A and A' correspond to the two outside lines in the bar-in-bar system. Moire B corresponds to the center line of the bar-in-bar system.

As in the bar-in-bar system, the offset term is cancelled by the differentiation operation without loss of the precision as

$$\mathcal{ORMS} \equiv [(A - B) - (B - A')]/2. \quad (6.2.2)$$

This equation shows the displacement error between gratings G1 and G2 as shown in Fig. 6.2.3. It is notable that there is a large magnification factor for the ADM in comparison to the bar-in-bar mark, which strongly reduces measurement errors.

6.3 Registration Measurement with Photore-sist

6.3.1 Basic Concept in Practice

We propose the ADM method for ORMS. The first and the second exposures are initially performed for two different mask patterns consisting of arrays of lines. Resist development leads to marks on the wafer with positions that are very sensitive to stage alignment errors. Chromium lines on one mask have an inclination angle, θ , with respect to the lines on the other mask. The overlapped areas of the chromium lines of the two masks lead to wedge-shaped unexposed areas on the wafer, which emerge as a positive photoresist pattern after development. The resulting set of unexposed patterns is in the form of a moire pattern.

Assuming the stage moves in the Y axis in order to overlap the two patterns, the resulting marks are displaced in the X axis by an amount related to the inclination angle and the displacement. In particular, a change in displacement ΔY causes a change of $\Delta X = \Delta Y \cdot (1/\tan \theta)$ in the X axis.

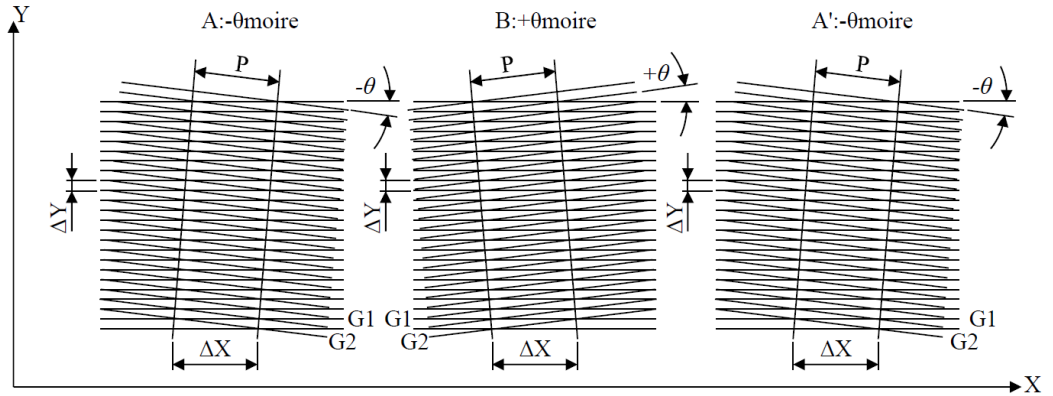


Figure 6.2.3: General principle of ADM; the notation is different from that of the subsection 6.3.2. The moiré pattern A, observed when two grating images G1 and G2 are tilted with an angle $-\theta$; if the grating G1 is moved a distance ΔY in the Y axis, the moiré pattern will move a distance $\Delta X = -\Delta Y / (\tan \theta)$ in the X axis. The moiré pattern B, tilted with an angle $+\theta$; if the grating G1 is moved a distance ΔY in the Y axis, the moiré pattern will move a distance $\Delta X = \Delta Y / (\tan \theta)$ in the X axis. Pattern A and pattern B have opposite phase. Pattern A' is identical to Pattern A, but the position is the opposite of pattern A [76].

The factor $(1/\tan \theta)$ is referred to as the moiré multiplication ratio. This procedure is used to measure distances between adjacent marks and then to divide the measured distances by the moiré multiplication ratio. It should be noted that all measurement errors are reduced by a factor of the moiré multiplication ratio.

In creating moiré photoresist images, the marks should meet the requirements for the ADM method, a zero point of the position should exist and the amount of its shift from the zero point should include directional information concerning the positional error. Figure 6.3.1 shows how moiré photoresist images are created by the double exposure process. A Y axis shift of the wafer, Δ , causes an X axis shift of the mark, Λ , in which Δ is multiplied by the ratio, $1/\tan \theta$, to create the large shift through the double exposure of a zigzag mark and a straight line mark. θ is angle between the zigzag and straight lines. At least three points of moiré photoresist images are needed to meet the requirement for the ADM method.

The changes in the coordinates of X1 to X2 and that of X2 to X3 are

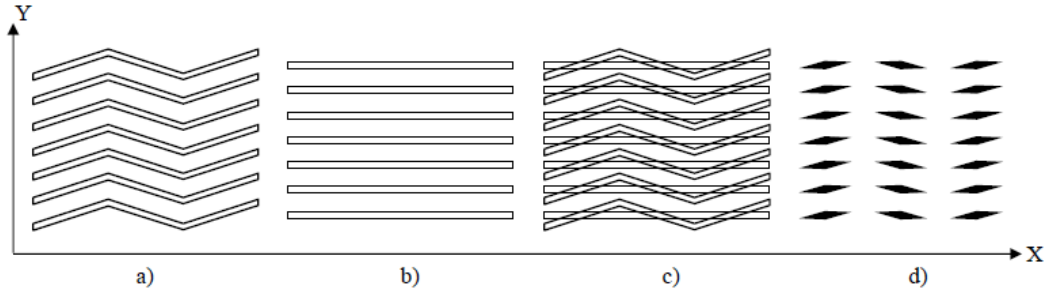


Figure 6.3.1: Combination of moire photoresist patterns. (a) The first mask pattern. (b) The second mask pattern. (c) Double exposed image. (d) Patterns after development used for the ADM. There are seven pairs of zigzag and straight lines [76].

related to alignment errors through the moire effect. Figure 6.3.2 shows moire photoresist images when Δ is zero. The zigzag line in the figure is the centerline of the zigzag pattern of the second mask pattern. The finite Δ case is as shown in Fig. 6.3.3.

6.3.2 Error Analysis of Measurement

We discuss the estimation of ADM measurement errors, where the following analysis is applicable to any type of ORMS sensor. We analyze the ADM measurement error in ORMS assuming measurement error follows a normalized Gaussian distribution, hereafter called the normal distribution in this thesis, without loss of generality. The measurement error is reduced by a factor equal to the inverse moire multiplication ratio. The coordinates of three moire photoresist marks are given by

$$X1 = -L + \Lambda + \epsilon_1 \quad (6.3.1)$$

$$X2 = -\Lambda + \epsilon_2, \quad (6.3.2)$$

$$X3 = L + \Lambda + \epsilon_3, \quad (6.3.3)$$

$$\Lambda = \Delta \cdot (1 / \tan \theta). \quad (6.3.4)$$

The measurement error of the detection for each mark is written as ϵ_i . The physical quantity measured in this method $\Lambda_{measured}$ is defined as the difference between the distance of X1 and X2 and that of X2 and X3, as depicted in Fig. 6.3.2 and Fig. 6.3.3.

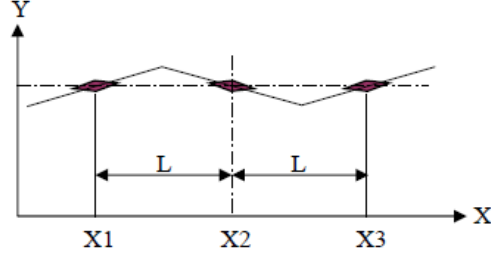


Figure 6.3.2: Example of double exposed patterns. Dot-dashed line represents the case where the Y axis shift Δ is zero [76].

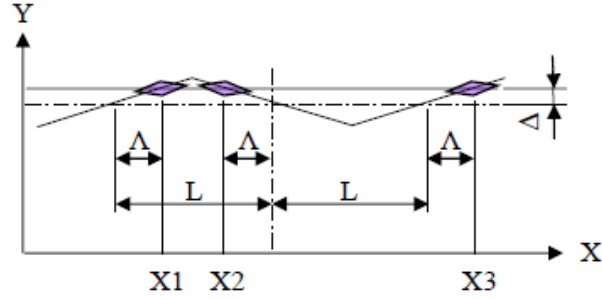


Figure 6.3.3: Solid line represents the case where the Y axis shift is Δ . X1, X2, and X3 are the coordinates of the marks shifted by the multiplied shift, Λ [76].

$\Lambda_{measured}$ can be expressed in terms of the true value Λ :

$$\begin{aligned}\Lambda_{measured} &= [(X3 - X2)/2 - (X2 - X1)/2]/2 \\ &= \Lambda + [(\epsilon_1 + \epsilon_3 - 2\epsilon_2)/2]/2.\end{aligned}\tag{6.3.5}$$

Using Eqs. (6.3.4) and (6.3.5), Δ can be expressed as

$$\begin{aligned}\Delta_{measured} &= \tan \theta \cdot [(X3 - X2)/4 - (X2 - X1)/4] \\ &= \Delta + \tan \theta \cdot (\epsilon_1 + \epsilon_3 - 2\epsilon_2)/4 \\ &\equiv \Delta + \epsilon.\end{aligned}\tag{6.3.6}$$

Finally, $\Delta_{measured}$ is defined from $\Lambda_{measured}$, and ϵ is the total measurement error.

The measurement error of each mark ϵ_i is assumed to have a constant variance regardless of θ , although the shape of the moire photoresist mark

appears different depending on the angle. This assumption is discussed in later Section in detail. In order to discuss the θ dependence of ϵ , we introduce a new variable ρ described by

$$\epsilon = (\tan \theta / 4) \cdot \rho. \quad (6.3.7)$$

We can derive the following equation:

$$\rho \sim N(0, 6\sigma^2), \quad (6.3.8)$$

in which the derivation is given in Appendix of the reference [76]. Referring to Eq. (6.3.8) and Eq. (6.3.7), ϵ is expressed as

$$\epsilon \sim N(0, (\tan \theta / 4)^2 \cdot 6\sigma^2). \quad (6.3.9)$$

From Eq. (6.3.9), the ORMS precision metrology of ADM is:

$$\sigma_{metrology} = \sqrt{(\tan \theta / 4)^2 \cdot 6\sigma^2}. \quad (6.3.10)$$

A factor K is defined as

$$K = 4 / \tan \theta. \quad (6.3.11)$$

The moire sensitivity and reduction ratio of the error can be adjusted by the appropriate selection of θ . K is proportional to the moire multiplication ratio.

6.4 Experimental Results

6.4.1 Experimental Data

The LSA for the ADM measurement is as shown in Fig. 6.4.1. The ADM mark arrangement is displayed as three columns of marks along the X axis in seven rows with $d = 4 \mu\text{m}$ pitch. The LSA diffraction signal along the X axis is as shown in Fig. 6.4.2. In the process of the stage scans, the light signal consists mostly of diffracted light by the columns of marks which is the same as other position amplification methods [74, 75]. In the discussion below, the ORMS precision σ is multiplied by a factor of 3. In the 3σ expression, all measured values lie within 3σ , approximately 99.7 %.

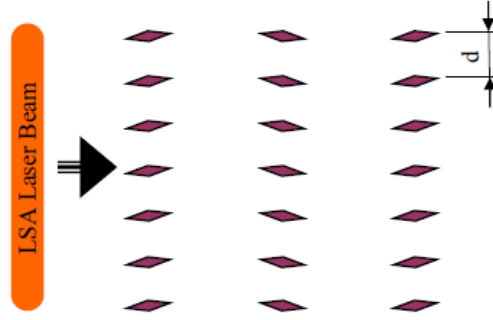


Figure 6.4.1: ADM patterns which are three rows of seven wedge shaped patterns with a pitch of $4\ \mu\text{m}$, and one dimensional laser scanning beam, LSA [76].

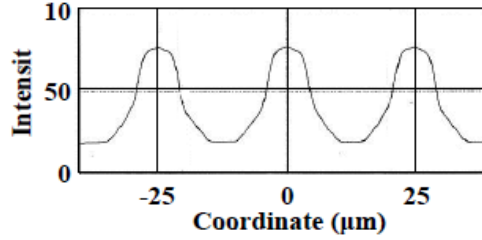


Figure 6.4.2: Waveform of LSA signals [76].

We investigate experimentally the validity of ADM by demonstrating that the precision improves in inverse proportion to the increasing magnification factor. Masks with several different angles are prepared for the experiment. In Fig. 6.4.3, there are two masks with $2\ \mu\text{m}$ line width periodic gratings for the double exposures. The masks are designed to have four different values of K : 10, 16, 20, and 32, corresponding to four different θ of 21.8, 14, 11.3, and 7.1 degrees, respectively.

Regarding $K = 1$, we adapt the conventional ORMS experimental data with the arrangement as shown in Fig. 3.2.2. The ORMS data for $K = 1$ gives 3σ precision as

$$3\sigma_{i\text{-line}, \text{Metrology}} = 14\text{nm}. \quad (6.4.1)$$

$\sigma_{i\text{-line}, \text{Metrology}}$ denotes the precision for the i -line exposure tool and the ORMS metrology.

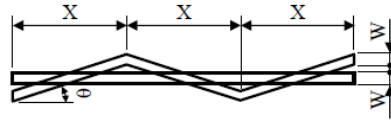


Figure 6.4.3: Design details of ADM patterns [76].

Four different magnification data points for the ORMS of ADM are taken repeatedly. From this data, the measurement precision is shown as five dots in Fig. 6.4.4, where the vertical and the horizontal axes are expressed on a logarithmic scale. According to the theory of ADM method, the precision in the experimental data in Fig. 6.4.4 has a slope of -1 with respect to magnification.

The data fit well to a linear model, with a coefficient of determination R of over 0.93 for five data points, indicating good agreement with the data. The increased multiplication of ADM in our experiments is consistent with the theoretical prediction calculated according to Eq. (6.3.11). At 32x ($\theta=7.1$ degrees), the 3σ precision of ORMS using the ADM for the i -line exposure apparatus is derived from the linear curve in Fig. 6.4.4 as

$$3\sigma_{i\text{-line}, \text{Moiremetrology}} = 0.50nm. \quad (6.4.2)$$

$3\sigma_{i\text{-line}, \text{Moiremetrology}}$ refers to the precision of the i -line tool and ORMS using the ADM metrology.

6.4.2 Verification of Assumptions

Based on the experimental results as shown in Fig. 6.4.4, the errors ϵ_i in Eqs. (6.3.1- 6.3.3) reach a constant value. Thus, the assumption in the former subsection 6.3.2, namely where the error expressed by a normal distribution is not influenced by the shape of the mark, is valid under experimental conditions. The LSA system has a constant measurement error, despite the fact that the wedge shaped mark becomes narrower with an increase in magnification factor K , at least up to $K = 32$.

6.4.3 Future Precision Improvement

The ADM method is a new methodology that combines the moire principle with a conventional sensor to detect a resist wedge mark. This method

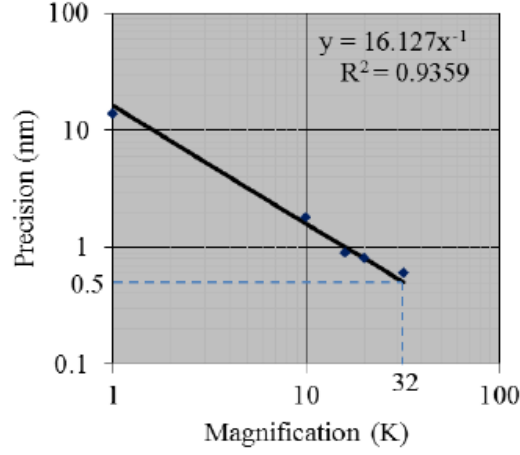


Figure 6.4.4: Relation between the precision and the magnification, K . The ordinate shows the ORMS 3σ precision. The abscissa shows ADM magnification K . Five data points for different K values are plotted. The solid line corresponds to a K^{-1} dependence [76].

would be relevant to an advanced generation immersion ArF exposure apparatus if mark edge measurement employs a non-stage-scanning sensor, such as a two dimensional CMOS image sensor. This assumes that equally precise measurement is possible for CMOS image sensors in comparison to one dimensional laser spot sensors.

For a typical immersion ArF exposure apparatus with a two dimensional CMOS alignment sensor, the 3σ sensor precision is usually assumed as $3\sigma_{ArF,CMOS} = 0.5 \text{ nm}$ [71], where $\sigma_{ArF,CMOS}$ is the sensor precision of the ArF tool with the two dimensional CMOS sensor. Then, from Eq. (6.2.1):

$$3\sigma_{ArF,metrology} = \sqrt{2} \cdot 3\sigma_{ArF,CMOS} = 0.707 \text{ nm}, \quad (6.4.3)$$

where $3\sigma_{ArF,metrology}$ denotes the ORMS precision for the ArF tool. The estimated ORMS precision of a 32x angle pattern using the ADM for the ArF tool is calculated from Eqs. (6.4.1), (6.4.2), and (6.4.3) as

$$\begin{aligned} 3\sigma_{ArF,Moiremetrology} &= (3\sigma_{i-line,Moiremetrology}) \cdot (3\sigma_{ArF,metrology}) / (3\sigma_{i-line,Metrology}) \\ &= 0.025 \text{ nm}, \end{aligned} \quad (6.4.4)$$

where $\sigma_{ArF,Moiremetrology}$ is the ADM precision of ORMS for the ArF exposure apparatus. This precision is less than 0.1 nm which is the target number [109] as desired in the semiconductor manufacturing processes recently.

Furthermore, by selecting an appropriate θ , the sensitivity, K , can be adjusted to provide further improvement if needed. It is notable that the ADM technology can offer a new possibility of the “pico-meter” order optical registration metrology.

6.4.4 Non-Photoresist Application

Our experiment for the ADM has been performed in case of photolithography where the moire pattern is printed on the photoresist. However, the concept of ADM is general because it is fundamentally derived from the bar-in-bar method. Therefore, beyond the case of photolithography, our ADM pattern is one of the candidates for wide area of moire position measuring metrology. To obtain the ADM marks, we require two sequential exposures, G1 and G2, as shown in Fig. 6.2.3.

Thus, it is impossible for the ADM to be applied to real-time moire measurement. However, if the pattern G1 has already been exposed on the wafer plane and the G2 reticle is fixed, real-time measurement is possible by using a camera.

6.5 Summary

We have developed a new concept for the ORMS measurement called the Alternating Direction Moire method, which utilizes the double exposure photoresist marks with opposite phases to create a moire pattern. This measurement method greatly improves precision in inverse proportion to the moire multiplication ratio. The experiments showed a 0.50 nm (3σ) precision using the optical alignment sensor of an i -line exposure apparatus, a precision 32 times higher than conventional methods.

Given an alignment sensor sensitivity adequate for wedge detection using the ADM, this method can be applied to the most advanced immersion ArF scanners, with a possible 25 pm (3σ) precision. As compared with the conventional approach to reduce error, our method does not require a large

number data and slow machine movements. This is because the error is reduced directly inverse proportional to the moire magnification.

The concept of ADM is not limited to the ORMS, but can be applied to other moire metrologies.

Chapter 7

Conclusion

7.1 Summary of Findings

I first summarize the important findings and results presented in each chapter. Table 7.1.1 represents the key research results.

Table 7.1.1: Summary of Findings

Metrology results		
Derived metrology	precision	measuring time
Sidelobe amplification	increase of over x10	less than SEM
Focus wedge mark	increase of over x50	decrease to 1 to 5 s
Double exposure linewidth	equal to SEM	decrease to 1 to 5 s
Alt. Dir. Moire	increase of over x30	equal to bar-in-bar

7.2 Concluding Remarks

In this section, I present the main concluding remarks of the study.

A coma aberration measurement derived from the image intensity amplification method has been created with high precision successfully.

Under monochromatic coherent illumination, the underlying methodology is that the left and right sidelobe amplitudes of the object image which are

transmitted through the projection lens affected by the coma aberration, interfere with the other bright-background amplitude to result in the sidelobe coma image intensity amplification; moreover, the difference of the left sidelobe and the right sidelobe image intensities is linearly proportional to the amount of coma aberration.

This principle of coma aberration measurement offers about ten-time better precision than that of the conventional methods.

The efficacy of the method is proved by theoretical analysis and simulation. The new methodology is easy to use, and it can be extended to other measurement applications.

I next discuss three image position amplification methods and the results via the methods, in which the wedge effect yields the best focus detection.

The tip position of the wedge image transmitted through the projection lens is non-linearly amplified by several tens; this idea which uses an array of wedge marks is termed as Focus Wedge Mark. In this study, by means of using a physical model, the tip position is expressed by the cosine function as the focus change, and the ratio of the tip position amplification is derived via both analyses and simulation.

The linewidth measurement technique which is derived from the image position amplification method is also successfully demonstrated. Two groups of lines and spaces lying at small cross angles are exposed sequentially to result in image position amplification; the efficacy of the method is verified for several angles and linewidths. From the results, it is noteworthy that small linewidth measurement (linewidth less than 20 nm) can easily amplify the linewidth to the large length ratio.

The third technique results in a new moire measurement configurations; the alternating direction moire (ADM) exhibits two pairs of oppositely directed moire fringes in one set that correspond to the positive and negative directional signals. For application to the displacement and the distortion measurements of the optical aligner, the ADM is formed by double exposure on the photoresist to magnify the displacement positions. With this approach, the experimental results show an increase in measurement precision by a factor of over 30 times than that achieved with conventional methods. My proposed methodology can be applied to obtain precision measurement which is required for semiconductor manufacturing.

Acknowledgement

I fully appreciate Dr. Naoshi Baba, Specially Appointed Professor at Hokkaido University to instruct me the thesis and to discuss with me every study.

I greatly appreciate Dr. Shunichi Muto, Specially Appointed Professor and Dr. Ryuji Morita, Professor, both at Hokkaido University who audit my study and improve my manuscript.

I deeply appreciate Dr. Hiroshi Ohki and Mr. Takashi Mori who are my supervisors willingly allowed me to submit manuscripts on several journals. I also deeply appreciate Mr. Koichi Matsumoto, Dr. Tomoya Noda, Dr. Shinichi Nakajima, Mr. Kinya Kato, Mr. Nobutaka Magome, Mr. Shigeru Hirukawa and Dr. Kazuya Okamoto who are in Nikon Corp. occasionally suggested me very helpful opinions.

On study of my thesis, I greatly thank to Mr. Satoshi Yashiki in Nikon research division to discuss the optical theory, and to study with me as a co-researcher. I sincerely thank to Dr. Michael Sogard who was the senior scientist of Nikon Research Corporation of America, for helping me to write manuscripts from the stand points of English language and logicalities.

In the discussion of Fourier Optics, I fully appreciate Dr. Masato Shibuya who is a Professor at Tokyo Polytechnic University to discuss with me.

I deeply thank to Mr. Koji Kaise, Mr. Akihiro Takezawa, Mr. Takashi Kambayashi, Mr. Takashi Ozawa, and Mr. Akira Yonezawa for preparing many conceptual drawings, and thank to Mr. Hiroshi Jimbo and Mr. Masaaki Mochida for keeping me the room space to write the thesis.

In order to complete the thesis, I have spent my time to write several manuscripts for about three years. During this period, I deeply thank to my wife, Toshiko, *Itsumo Arigatou* for supporting me continuously.

Bibliography

- [1] J. Kilby, USP **3138743**, Filed (1959).
- [2] R. Noyce, USP **2981877**, Filed (1959).
- [3] G. Moore, Electronics Magazine, 19 April (1965).
- [4] G. Moore, Proc. SPIE **2440**, p. 2-17 (1995)
- [5] E. Hecht, *Optics 4th.ed.*, Chap. 1 (Addison Wesley, SF, CA, 2002).
- [6] E. Hecht, *Optics 4th.ed.*, Chap. 6 (Addison Wesley, SF, CA, 2002).
- [7] M. Steinbach, “Ernst Abbes Komparatorprinzip”, p. 70-81 *Ernst – Abbe – Sonderband*, (JENAER JAHRBUCH, Verlag Christine Jager KG, Jena, 2005).
- [8] H. H. Hopkins, and P. M. Barham, Proc. Phys. Soc., B **63**, p. 737-742 (1950).
- [9] J. W. Goodman, *Introduction to Fourier Optics, 3rd.ed.*, Chap. 6 (Roberts and Company Publishers, Greenwood Village, CO, 2005).
- [10] M. Born and E. Wolf, *Principles of Optics*, p.517-553 (Cambridge University Press, London, 1999).
- [11] J. H. Bruning, Proc. SPIE **6520**, p. 652004 1-13 (2007).
- [12] J. H. Bruning, Proc. SPIE **3051**, p. 14-27 (1997).
- [13] R. Kingslake, *Lens Design Fundamentals*, p. 320-322 (Academic Press, New York, NY, 1978).

- [14] M. J. Kidger, *Intermediate Optical Design*, Chap. 8 (SPIE Press Bellingham, WA, 2003).
- [15] K. Nakazawa, K. Suwa, and S. Yoshida, Proc.19th Symposium on Semiconductors and Integrated Circuit Technology, Electrical Society of Japan (in Japanese), p. 48-53 (1980).
- [16] K. Nakazawa, K. Suwa, and S. Yoshida, Solid State Technology (in Japanese), p. 16-26 (1980).
- [17] K. Nakazawa, K. Suwa, and S. Yoshida, Kodak Microelectronics Seminar, p. 61-66 (1981).
- [18] A.E. Conrady, Mon. Not. Royal Ast. Soc., **79**, p. 60-66 (1918).
- [19] A.E. Conrady, Mon. Not. Royal Ast. Soc., **79**, p. 384-390 (1918).
- [20] A.E. Conrady, Mon. Not. Royal Ast. Soc., **80**, p. 78-91 (1919).
- [21] H. Kubota, *Kogaku (in Japanese)*, p. 69-79 (Iwanami, Tokyo, 1978).
- [22] H. Kubota, *Kogaku (in Japanese)*, p. 81-84 (Iwanami, Tokyo, 1978).
- [23] D. Malacara, *Optical Shop Testing*, 3rd. ed., Chap. 11 (John Wiley and Sons, Hoboken, New Jersey, 2007).
- [24] F. Twyman, *Prism and Lens Making*, p. 367-372 (Adam Filger, Bristol and New York, 1988).
- [25] P. Hariharan, *Optical Interferometry 2nd. ed.*, Chap. 2 (Academic Press, San Diego, CA, 2003).
- [26] F. Twyman, *Prism and Lens Making*, p. 422-448 (Adam Filger, Bristol and New York, 1988).
- [27] D. Malacara, *Optical Shop Testing*, 3rd.ed., Chap. 2 (WILEY-InterScience, Hoboken, NJ, 2006).
- [28] K. Yamada, *Kogaku no Chishiki (in Japanese)*, p. 254-258 (Tokyo Denki Daigaku Press, Tokyo, 1966).
- [29] J. H. Bruning, D. R. Herriott, J. E. Gallagher, D. P. Rosenfeld, A. D. White, and D. J. Brangaccio, Appl. Opt. **13**, p. 2693-2703 (1974).

- [30] K. Suwa and K. Ushida, Proc. SPIE **922**, p. 270-276 (1988).
- [31] K. Suwa and K. Ushida, *Selected Papers on Optical Microlithography*, ed. H. L. Stover, p. 406-412 (SPIE Press, Bellingham, 1992).
- [32] R.P.Hardin and M.P.Gulden, Proc. SPIE **633**, p. 298-306 (1986).
- [33] K. Murray, Semiconductor International, Dec. p. 69-73 (1982).
- [34] J. P. Kirk, Proc. SPIE **1463**, p. 282-291 (1991).
- [35] J. W. Gemmink, Proc. SPIE **1088**, p. 220-230 (1989).
- [36] K. Toh and A. Neureuther, KTI Microelectronics Seminar, p. 155 (1989).
- [37] J. P. Kirk, Proc. SPIE **4000**, p. 2-8 (2000).
- [38] M. Shibuya, JP **1441789** (in Japanese). Japan Patent Publication No. 62052 (1982).
- [39] M. D. Levenson, N. S. Viswanathan, and R. A. Simpson, IEEE Trans. Electron Devices ED-29, p. 1828-1836 (1982).
- [40] T. A. Brunner, A. L. Martin, R. M. Martino, C. P. Ausschnitt, T. H. Newman, and M. S. Hibbs, Proc. SPIE **2197**, p. 541-549 (1994).
- [41] H. Nomura, Opt. Rev. **8**, 3, 184-190 (2001).
- [42] K. Komatsu, T. Omori, T. Kitamura, Y. Nakajima, A. A. Aiyyer, and K. Suwa, Proc. SPIE **3677**, p. 764-771 (1999).
- [43] T. Kitamura, Y. Nakajima, H. Matsumoto, Takeo Omori, and K. Komatsu, Proc. SPIE **3998**, p. 615-621 (2000).
- [44] K. Fukazawa, K. Endo, K. Yoshino, and Y. Yamazaki, Proc. SPIE **6518**, p. 65182H 1-7 (2007).
- [45] K. Lensing, J. Cain, A. Prabhu, A. Vaid, R. Chong, R. Good, B. LaFontaine, and O. Kritsun, Proc. SPIE **6518**, p. 651804 1-12 (2007).

- [46] T. Wiltshire, D. Corliss, T. Brunner, C. Ausschnitt, R. Young, R. Nielson, E. Hwang, and J. Iannucci. Jr., Proc. SPIE **7272**, p. 72720W 1-12 (2009).
- [47] T. Sato, S. Inoue, K. Kasa, Y. Nakajima, S. Nagai, and K. Sato, Jpn. J. Appl. Phys. **48**, p. 06FA03 1-4 (2009).
- [48] P. Vanoppen, T. Theeuwes, H. Megens, H. Cramer, T. Fliervoet, M. Ebert, and D. Satriasaputra, Proc. SPIE **7640**, p. 764010 1-15 (2010).
- [49] Y. Fujimori, T. Tsuto, Y. Kudo, T. Inoue, K. Suwa, and K. Okamoto, J. Micro/Nanolith. MEMS MOEMS **12**(1), p. 013013 1-9 (2013).
- [50] Y. Fujimori, T. Tsuto, H. Tsukamoto, K. Okamoto, and K. Suwa, J. Micro/Nanolith. MEMS MOEMS **13**(1), p. 011204 1-7 (2014).
- [51] T. Tsuto, Y. Fujimori, H. Tsukamoto, K. Suwa, and K. Okamoto, Trans. Japan Institute of Electron Packaging, 6(1), p. 13-17 (2013).
- [52] P. Yoder Jr., *Mounting Optics in Optical Instruments, 2nd.ed.*, p. 527-528 (SPIE Press, Bellingham, WA, 2008).
- [53] K. B. Doyle, V. L. Genberg, and G. J. Michels, *Integrated Optomechanical Analysis, 2nd ed.*, Chap. 4, p. 81 (SPIE Press, Bellingham, Washington, 2012).
- [54] R. W. Gray, C. Dunn, K. P. Thompson, and J. P. Rolland, Opt. Express **20**, p. 16436-16449 (2012).
- [55] R. Barakat and A. Houston, J. Opt. Soc. Am. **54**(9), p. 1084-1088 (1964).
- [56] R. Barakat and A. Houston, J. Opt. Soc. Am. **55**, p. 1132-1135 (1965).
- [57] R. Barakat and A. Houston, J. Opt. Soc. Am. **56**, p. 1042-1403 (1965).
- [58] J. Kye, M. Dusa, and H. J. Levinson, Proc. SPIE **4346**, p. 1280-1289 (2001).
- [59] A. H. Slocum, *Precision Machine Design*, Chap. 2, p. 58 (Society of Manufacturing Engineers, Dearborn, MI, 1992).

- [60] R. B. Northrop, *Introduction to Instrumentation and Measurements second edition*, p. 9 (Taylor and Francis, Boca Raton, FL, 2005).
- [61] S. T. Smith, D. G. Chetwynd, *Foundation of Ultraprecision Mechanism Design*, p. 6 (CRC Press Boca Raton, FL, 1992).
- [62] O. Taniguchi, Y. Horigome, *Keisoku Kogaku* (in Japanese)", p. 8 (Morikita Shuppan, Tokyo, 1999).
- [63] S. Murakami, T. Matsuura, M. Ogawa, and M. Uehara, Proc. SPIE **538**, p. 9-16 (1985).
- [64] A. H. Slocum, *Precision Machine Design*, Chap. 4 (Society of Manufacturing Engineers, Dearborn, MI, 1992).
- [65] K. Suwa, Kokai **H6/101427**, Open 1987, Kokoku (1994).
- [66] K. Ota, N. Magome, and K. Nishi, Proc. SPIE **1463**, p. 304-314 (1991).
- [67] A. Sugaya, Jpn. J. Appl. Phys. **43**, p. 3999-4003 (2004).
- [68] A. Sugaya, Y. Kanaya, S. Nakajima, T. Nagayama, and N. Shiraishi, Jpn. J. Appl. Phys. **43**, p. 7419-7426 (2004).
- [69] A. Sugaya, T. Nagayama, and T. Masada, Jpn. J. Appl. Phys. **44**, p. 5505-5514 (2005).
- [70] A. Sugaya, Jpn. J. Appl. Phys. **44**, p. 5970-5976 (2005).
- [71] Y. Shirata, Y. Shibazaki, J. Kosugi, T. Kikuchi, and Y. Ohmura, Proc. SPIE **8683**, p. 86831K6 1-10 (2013).
- [72] J. Maas, M. Ebert, K. Bhattacharyya, H. Cramer, A. Becht, S. Keij, R. Plug, A. Fuchs, M. Kubis, T. Hoogenboom, and V. Vaenkatesan, Proc. of SPIE **7985**, 79850H, p. 1-10 (2011).
- [73] K. Suwa, S. Yashiki, S. Hirukawa, and T. Noda, Appl. Opt. **53**, p. 2494-2506 (2014).
- [74] K. Suwa, S. Yashiki, and S. Hirukawa, Jpn. J. Appl. Phys., **53**(9), p. 096601-6 (2014).

- [75] I. Grodnensky, K. Suwa, N. Farrar, E. Johnson, and J. Pan, J. Vac. Sci. Technol. **B17**(6), p. 3285-3290 (1999).
- [76] K. Suwa, K. Koji, H. Tatenno, and N. Magome, Opt. Eng., **53**, p. 084101-7 (2014).
- [77] K. Suwa and S. Hirukawa, JP **3259293**, Filed (1991).
- [78] N. Takatsu, K. Suwa, S. Nakamura, H. Hosokawa, and S. Hirukawa, USP **5434026**, Filed (1992).
- [79] S. Hirukawa, H. Tatenno, and K. Suwa, JP **3117095**, Filed (1991).
- [80] K. Suwa, S. Hirukawa, and H. Tatenno, JP **2580668**, Filed (1988).
- [81] K. Suwa, H. Tatenno, N. Irie, and S. Hirukawa, Proc. SPIE **2440**, p. 712-720 (1995).
- [82] <http://www.drlitho.com/cms/website.php>
- [83] M. A. Narasimham and J. H. Carter. Jr, Proc. SPIE **135**, p. 2-9 (1978).
- [84] C. A. Mack and P. M. Kaufman, Proc. SPIE **1138**, p. 88-105 (1989).
- [85] C. A. Mack, *Field Guide to Optical Lithography*, p. 54 (SPIE Press, Bellingham, 2006).
- [86] T. Matsuyama, N. Kita, and Y. Mizuno, Proc. SPIE **7973**, p. 79731H (2011).
- [87] T. Matsuyama, N. Kita, R. Matsui, and J. Ikeda, Proc. SPIE **8326**, p. 83260K (2012).
- [88] N. Hirayanagi, Y. Mizuno, M. Mori, N. Kita, R. Matsui, and T. Matsuyama, Proc. SPIE **8683**, p. 86830D 1-10 (2013).
- [89] R. Matsui, T. Noda, H. Aoyama, N. Kita, T. Matsuyama, and D. Flagello, Proc. SPIE **8683**, p. 868300 1-7 (2013).
- [90] H. Aoyama, T. Nakashima, T. Ogata, S. Kudo, N. Kita, J. Ikeda, R. Matsui, H. Yamamoto, A. Sukegawa, K. Makino, M. Murayama, K. Masaki, and T. Matsuyama, Proc. SPIE **9052**, p. 90520A 1-12 (2014).

- [91] K. Suwa, H. Tatenno, and S. Hirukawa, JP **2696962**, Filed (1988).
- [92] I. Grodnensky, K. Suwa, K. Ushida, and E. Johnson, USP **6449031B1**, Filed (2000).
- [93] K. Suwa, USP **6855997B2**, grant (2005).
- [94] I. Grodnensky, E. Morita, K. Suwa, and S. Hirukawa, Proc. SPIE **3334**, p. 289-296 (1998).
- [95] O. Kafri and I. Glatt, *The Physics of Moire Metrology*, (Wiley, New York, 1989).
- [96] R. F. Anastasi, No. MTL TR 92-55 Army Lab Command, Material Technology Lab, Watertown, MA (1992).
- [97] R. Chang, J. Sheu, C. Lin, and H. Liu, Optics & Laser Technology **35**, p. 43-47 (2003).
- [98] J. Liu, H. Furuhashi, A. Torii, N. Hayashi, J. Yamada, Y. Uchida, R. Sharma, A. K. Kanjilal, and V. T. Chitnis, Proc. IEEE IECON 21st **2**, p. 968-972 (1995).
- [99] B. Han, Journal of Thermal Stresses **26**, p. 583-613 (2003).
- [100] S. Yokozeki and T. Suzuki, Appl. Opt. **9**, p. 2804-2805 (1970).
- [101] A. Livnat and O. Kafri, Appl. Opt. **22**, p. 3232-3235 (1983).
- [102] J. Dhanotia, S. Prakash, S. Rana, and O. Sasaki, Appl. Opt. **50**, p. 2958-2963 (2011).
- [103] M. Seib and H. Hoefler, Proc. SPIE **1375**, p. 28-38 (1989).
- [104] Y. Cheng, S. Lin, K. Chen, C. Ke, and T. Gau, Proc. SPIE **8681**, p. 86812P 1-6 (2013).
- [105] M. Adel, M. Ghinovker, B. Golovanevsky, P. Izikson, E. Kassel, D. Yaffe, A. M. Bruckstein, R. Goldenberg, Y. Rubner, and M. Rudzsky, Semiconductor Manufacturing **17**, p. 166-179 (2004).
- [106] L. Nugent-Glandorf and T. T. Perkins, Opt. Lett. **29**, p. 2611-2613 (2004).

- [107] K. Suwa, H. Tateno, and S. Hirukawa, JP **3341774**, Filed (1990).
- [108] H. Tateno, K. Kaise, K. Suwa, and Y. Imai, USP **5666205**, Filed (1995).
- [109] W. H. Arnold, Proc. SPIE **8681**, p. 1-12 (2013).